

REVISTA IBERO —AMERICANA

de Educación
de Educação

Nuevos datos, nuevos retos:
Iberoamérica en las últimas
evaluaciones educativas

Novos dados, novos desafios:
Ibero-América nas últimas
avaliações educativas

© Madrid, OEI, 2020

Nuevos datos, nuevos retos: Iberoamérica en las últimas evaluaciones educativas

Novos dados, novos desafios: Ibero-América nas últimas avaliações educativas

Revista Iberoamericana de Educación / *Revista Ibero-americana de Educação*

Vol. 84. Núm. 1

Septiembre-Diciembre / *Setembro-Dezembro*

264 páginas

Revista cuatrimestral / *Revista quadrimestral*

EDITA

Educación Superior, Ciencia y EFTP

Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI)

Bravo Murillo, 38. 28015 - Madrid, España / Tel.: (34) 91 594 43 82

rie@oei.org.es; <https://rieoei.org/RIE>

ISSN: 1022-6508 - ISSNe: 1681-5653

Depósito Legal: BI-1094-1993

Diseño de la cubierta: Leyre Mayendia

Foto de la portada: Shutterstock

TEMAS / TEMAS

evaluaciones internacionales a gran escala; evaluación educativa, políticas educativas, Iberoamérica

avaliações internacionais em grande escala; avaliação educacional, políticas educacionais, Ibero-América

La REVISTA es una publicación indizada en: / *A REVISTA é uma publicação indexada em:*

WOS: <https://clarivate.com/>

DOAJ: <https://doaj.org/>

REDIB: www.redib.org/

LATINDEX: www.latindex.unam.mx

Dialnet: <https://dialnet.unirioja.es/>

IRESE: www.iisue.unam.mx/iresie

ABES SUDOC: www.sudoc.abes.fr

Biblioteca Digital: www.oei.es/bibliotecadigital.php

Qualis - CAPES: <http://qualis.capes.gov.br/webqualis>

Portal periodicos. Capes: <http://www.periodicos.capes.gov.br/>

La Revista evalúa los trabajos mediante el sistema de arbitraje «ciego por pares» a través de especialistas externos a la Redacción.

A Revista avalia os trabalhos mediante o sistema de arbitragem «cego por pares» através de especialistas externos à Redação.

Todos los números MONOGRÁFICOS se encuentran digitalizados en acceso abierto en la página web de la revista. Además, la RIE edita números extraordinarios con investigaciones, ensayos e innovaciones educativas de otras áreas educativas, que complementan a los monográficos

Todos os números TEMÁTICOS podem ser consultados em formato digital no site da revista. Também, a RIE apresenta números especiais não temáticos com investigações, estudos, inovações e ensaios para complementar em outras áreas de interesse educacional

La OEI no se responsabiliza de las opiniones expresadas en los artículos firmados ni comparte necesariamente las ideas manifestadas en los mismos.

A OEI não se responsabiliza pelas opiniões expressas nos artigos assinados nem compartilha necessariamente as ideias manifestadas nos mesmos.



Director / Diretor: Francesc Pedró, UNESCO-IESALC

Secretaría de Redacción / Secretaria de Redação: Ana Capilla

Equipo de Redacción / Equipe de Redação: Bárbara García, Paula Sánchez-Carretero, Andrés Viseras

Traducción (portugués) / Tradução (português): Simone Nascimento

COORDINADORES DE ESTE NÚMERO / COORDENADORES DESTE NÚMERO

Jorge Sainz González, Universidad Rey Juan Carlos

Andrés Sandoval-Hernández, University of Bath

CONSEJO EDITORIAL / CONSELHO EDITORIAL

Otto Granados, ex Secretario de Educación Pública de México y presidente del Consejo Asesor de la OEI

Ángel Gabilondo, Universidad Autónoma de Madrid (España)

Alejandro Jorge Granimian, Universidad de New York (EE.UU.)

Ariel Fiszbein, *Diálogo Interamericano* (Argentina)

Axel Rivas, Universidad de San Andrés (Argentina)

Cecilia María Vélez, Universidad de Bogotá Jorge Tadeo Lozano (Colombia)

Claudia Laura Limón Luna, CONCIUS (México)

Claudia Peirano, Fundación Educacional Oportunidad (Chile)

Emiliana Vegas, Center for Universal Education - Brookings Institution (EE.UU.)

Maria Claudia Uribe Salazar, Oficina Regional de Educación de la UNESCO para América Latina y el Caribe. Representante para Chile

Juan Ernesto Treviño Villareal, Pontificia Universidad Católica de Chile

Fernando M. Reimers, Universidad de Harvard (EE.UU.)

Héctor Valdés Véloz, Corporación Conciencia Educativa de Chile

Victor Hugo Díaz Díaz, Consejo Nacional de Educación, Perú

Jaime Saavedra, Banco Mundial

Jorge Sainz González, Universidad Rey Juan Carlos (España)

José Augusto Britto Pacheco, Instituto de Educação do Minho (Portugal)

José Henrique Paim Fernandes, Centro de Gestão Municipal e Políticas Educacionais (Brasil)

José Joaquín Brunner, Universidad Diego Portales (Chile)

José David Weinstein Cayuela, Universidad Diego Portales (Chile)

Lorenzo Gomez Morin Fuentes, Secretaría de Educación del Gobierno de Baja California (México)

Margarita Peña, Universidad Jorge Tadeo Lozano (Colombia)

Maria Helena Guimarães de Castro, Universidad Estatal de Campinas - UNICAMP (Brasil)

Mariano Fernández Enguita, Universidad Complutense de Madrid (España)

Mariano Narodowski, Universidad Toruato Di Tella (Argentina)

Melina Gabriela Furman, Universidad de San Andrés (Argentina)

Rafael de Hoyos Navarro, Unidad de Educación para América Latina. Banco Mundial

Renato Esteban Operti Belando, Universidad Católica del Uruguay

Ricardo Cuenca Pareja, Instituto de Estudios Peruano - IEP (Perú)

Sergio Cárdenas Denham, Centro de Investigación y Docencia Económicas - CIDE (México)

CONSEJO CIENTÍFICO / CONSELHO CIENTÍFICO

Agustín de la Herrán Gascón, Universidad Autónoma de Madrid, España. ①

Américo Domingos Matindique, Universidade Jean Piaget de Moçambique. ②

Ángel San Martín Alonso, Universidad de Valencia, España. ③

Ascensión Palomares Ruiz, Universidad de Castilla-La Mancha, España. ④

António Manuel Águas Borralho, Universidade de Évora, Portugal. ⑤

Carmen Nieves Pérez Sánchez, Universidad de La Laguna, España. ⑥

Cleci Werner da Rosa, Universidade de Passo Fundo, Brasil. ⑦

Edson Jorge Huairé Inacio, Universidad San Ignacio de Loyola (USIL), Perú. ⑧

Elsa Piedad Cabrera Murcia, Pontificia Universidad de Chile. ⑨

Francisco Ramos Calvo, Loyola Marymount University, EE.UU. ⑩

Gregorio Jiménez Valverde, Universidad de Barcelona, España. ⑪

Isabel María Gallardo Fernández, Universidad de Valencia, España. ⑫

Isabel Patricia Espiro Barrera, Universidad Santo Tomás, Chile. ⑬

Joan Andrés Traver Martí, Universidad Jaime I, España. ⑭

Jorge Bonito, Universidade de Évora, Portugal. ⑮

José Armando Salazar Ascencio, Universidad de La Frontera, Chile. ⑯

José Quintanal Díaz, Universidad Nacional de Educación a Distancia, España. ⑰

Juan José Leiva Olivencia, Universidad de Málaga, España. ⑱

Juan Vicente Ortiz Franco, Fundación Universitaria Los Libertadores, Colombia. ⑲

Liliana Soares Ferreira, Universidade Federal de Santa Maria, Brasil. ⑳

Manuel Ferraz Lorenzo, Universidad de La Laguna, España. ㉑

Márcia Lopes Reis, Universidade Estadual Paulista - UNESP, Brasil. ㉒

Marco Silva, Universidade do Estado do Rio de Janeiro - UERJ, Brasil. ㉓

Maria Ángeles González Galán, Universidad Nacional de Educación a Distancia (UNED), España. ㉔

Maria Amelia Pidello Rossi, IRICE-CONICET, Argentina. ㉕

Maria Clemente Linuesa, Universidad de Salamanca, España. ㉖

Maria Célia Borges, Universidade Federal de Uberlândia (UFU), Brasil. ㉗

Maria del Carmen Lorenzatti, Universidad Nacional de Córdoba, Argentina. ㉘

Maria Eveline Pinheiro Villar de Queiroz, Ministério da Educação, Brasil. ㉙

Maria Inmaculada Egidio Gálvez, Universidad Autónoma de Madrid, España. ㉚

Maria Jesús Vitón de Antonio, Universidad Autónoma de Madrid, España. ㉛

Maria José Bautista-Cerro Ruiz, Universidad Nacional de Educación a Distancia (UNED), España. ㉜

Maria Teresa Gómez del Castillo, Universidad de Sevilla, España. ㉝

Ondina Pena Pereira, Universidade Católica de Brasília, Brasil. ㉞

Rafael Guimarães Botelho, Instituto Federal de Educação, Ciência e Tecnologia do Rio de Janeiro (IFRJ), Brasil. ㉟

Rafael Pérez Flores, Universidad Autónoma Metropolitana, México. ㊱

Rosa Vázquez Recio, Universidad de Cádiz, España. ㊲

Ronilson Freitas de Souza, Universidade do Estado do Pará (UEPA), Brasil. ㊳

Silvia Maria de Oliveira Pávao, Universidade Federal de Santa Maria-UFSM, Brasil. ㊴

Teresita Alzate Yepes, Universidad de Antioquia, Colombia. ㊵

Valentín Martínez-Otero Pérez, Universidad Complutense de Madrid, España. ㊶

William Moreno Gómez, Universidad de Antioquia, Colombia. ㊷

EVALUADORES DE DE ESTE NÚMERO / AVALIADORES DESTA NÚMERO

Alicia Sianes-Bautista, *Universidad de Jaén (UJA), España* 
Adriano de Araujo Santos, *Universidade Federal Rural de Pernambuco (UFRPE), Brasil* 
Artemio Cortez Ochoa, *University of Bristol, United Kingdom* 
Begoña Ladrón de Guevara Pascual, *Centro Universitario Villanueva, España*, 
Carolina Valladares Celis, *University of Bristol, United Kingdom* 
Daniel Miranda, *Pontificia Universidad Católica de Chile (UC)* 
Delia Arroyo Resino, *Universidad Internacional de La Rioja (UNIR), España* 
Eva Expósito, *Universidad Nacional de Educación a Distancia (UNED), España* 
Francisco López Rupérez, *Universidad Camilo José Cela (UCJC), España* 
Ismael Sanz Labrador, *Universidad Rey Juan Carlos (URJC), España* 
Jesús García Laborda, *Universidad Alcalá de Henares (UAH), España* 
María Herrero Herrero, *University of Bath, United Kingdom* 
María Inmaculada Egido Gálvez, *Universidad Autónoma de Madrid (UAM), España* 
Marián Navarro-Beltrá, *Universidad Católica de Murcia (UCAM), España* 

REVISTA IBEROAMERICANA DE EDUCACIÓN REVISTA IBERO-AMERICANA DE EDUCAÇÃO

Volumen 84. Número 1
Septiembre-Diciembre / Setembro-Dezembro 2020

MONOGRÁFICO / MONOGRÁFICO

Nuevos datos, nuevos retos: Iberoamérica en las últimas evaluaciones educativas

Novos dados, novos desafios: Ibero-América nas últimas avaliações educativas

Coordinadores / Coordenadores: Jorge Sainz y Andrés Sandoval-Hernández

Sumario / Sumário

Editorial

Jorge Sainz y Andrés Sandoval-Hernández, Introducción / *Introdução*9

Monográfico

Cristóbal Villalobos, Ignacio Wyman y Ernesto Treviño, Evaluaciones internacionales a gran escala y ciudadanía. Explorando la relación entre políticas y ICCS en Chile (2009-2016)15

David G. Miranda y Patricia Torres Género y etnicidad. Un análisis crítico de las políticas de equidad y los resultados de la prueba PISA en Brasil y México37

Luiz Carlos Gesqui, Desempenho Ibero-americano nos PISA de 2012, 2015 e 2018....67

Abigail Middel, Kalyan Kumar Kameshwara and Andrés Sandoval-Hernandez, Exploring trends in the relationship between child labour, gender and educational achievement in Latin America85

Nurullah Eryilmaz; Mauricio Rivera-Gutiérrez y Andrés Sandoval-Hernández, Should different countries participating in PISA interpret socioeconomic background in the same way? A measurement invariance approach.....109

Soledad Miranda y Ernesto Treviño, Uso de TERCE-UNESCO para informar la práctica educativa: factores que influyen en el aprendizaje en Ciencias en Chile y Paraguay135

Adrián Silveira Aberastury, Determinante del desempeño en lectura en Uruguay: un análisis multinivel a partir de TERCE.....157

Edite Maria Sudbrack e Dora Ramos Fonseca, A regulação supranacional operada pelo PISA: construindo credibilidade internacional177

Francisco López Rupérez e Isabel García García, España vs. Portugal en educación. Una aproximación sistémica193

Camila Ferreira da Silva; Cecília Acácia da Silva Sampaio e Thiago Felipe Paranatinga da Cunha, Do julgamento professoral às avaliações externas, que fazemos com as crianças e jovens na escola? Um estudo de caso na cidade de Manaus-AM217

Juan Rubén Compañ García, La evaluación para el ingreso al servicio educativo y su impacto en el rendimiento de la prueba PISA en México.....237

Nuevos datos, nuevos retos:
Iberoamérica
en las últimas evaluaciones
educativas

Novos dados, novos
desafios: Ibero-América
nas últimas avaliações
educativas

Nuevos datos, nuevos retos: Iberoamérica en las últimas evaluaciones educativas *Novos dados, novos desafios: Ibero-América nas últimas avaliações educacionais*

Jorge Sainz¹  ; Andrés Sandoval-Hernández² 

¹Universidad Rey Juan Carlos (URJC), España; ²University of Bath, United Kingdom

Las evaluaciones internacionales a gran escala (conocidas por sus siglas en inglés como ILSA) han sido parte de la escena educativa global desde 1964, cuando la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA, por sus siglas en inglés) realizó el primer estudio internacional comparativo de matemáticas, en el que participaron 12 países (Cresswell, et al. 2015). Desde entonces, ha habido un importante incremento en el número de estudios globales (e.g. PISA, TIMSS, PIRLS, EGRA) y regionales (e.g. ERCE, SAQMEC, PASEC, SEA-PLM, PILNA) que buscan evaluar y comparar los sistemas educativos de diferentes países. Una revisión de las listas de países participantes en los diferentes ILSAs deja ver que, a la fecha, más de la mitad de los países del mundo y la gran mayoría de los países de Iberoamérica¹ han participado en al menos uno de estos estudios.

La cantidad y calidad de la información que estos estudios han puesto a disposición de los diferentes actores educativos, han hecho que los ILSAs ganen terreno en la agenda política global y de nuestros países. Por ejemplo, la relativamente reciente adopción de los Objetivos de Desarrollo Sostenible (ODS), en particular el Objetivo 4 (Garantizar una educación de calidad inclusiva y equitativa, y promover las oportunidades de aprendizaje permanente para todos), ha contribuido a incrementar el protagonismo de la información estadís-

As avaliações internacionais em grande escala (conhecidas pela sigla em inglês como ILSA) fazem parte do cenário educacional global desde 1964, quando a Associação Internacional para a Avaliação do Desempenho Educacional (IEA, em sua sigla em inglês) realizou o primeiro estudo internacional comparativo de matemática, no qual participaram 12 países (Cresswell, et al. 2015). Desde então, houve um aumento significativo no número de estudos globais (por exemplo PISA, TIMSS, PIRLS, EGRA) e regionais (por exemplo ERCE, SAQMEC, PASEC, SEA-PLM, PILNA) que buscam avaliar e comparar os sistemas educacionais de diferentes países. Uma revisão das listas de países participantes nas diferentes ILSAs revela que, até o momento, mais da metade dos países do mundo e a grande maioria dos países da Ibero-americanos¹ participaram de pelo menos um desses estudos.

A quantidade e a qualidade das informações que esses estudos disponibilizaram dos diferentes atores educacionais fizeram com que as ILSAs ganhassem espaço na agenda política global e nas dos nossos países. Por exemplo, a adoção relativamente recente dos Objetivos de Desenvolvimento Sustentável (ODS), em particular o Objetivo 4 (Garantir uma educação de qualidade inclusiva e equitativa e promover oportunidades de aprendizagem permanente para todos), contribuiu para aumentar o protagonismo da informação estatística sobre educação de forma inédita, tornando

¹ Las excepciones son Andorra y Guinea Ecuatorial.

¹ As exceções são Andorra e Guiné Equatorial.

tica sobre educación de una manera sin precedente, convirtiendo a los ILSAs en una herramienta crucial para el monitoreo y la guía de políticas educativas.

Sin embargo, es indispensable reconocer los límites de los ILSAs para poder generar las oportunidades que nos permitan utilizar esta información de manera eficaz para informar acciones encaminadas a la mejora de nuestros sistemas educativos. La reciente publicación de nuevos datos internacionales (e.g. PISA in 2020) y la inminente publicación de datos regionales (e.g. ERCE in 2021) marca una excelente oportunidad para enfrentar este reto. Al mismo tiempo, este es uno de los objetivos principales de este monográfico. De acuerdo con un reporte de la UNESCO (2019) sobre este tema, son cinco los retos principales que, como comunidad interesada en la educación, debemos atender para asegurar la utilización de estos datos para tomar decisiones fundamentadas que mejoren la enseñanza y el aprendizaje:

- *Limitaciones en cuanto a la conceptualización de la educación y el aprendizaje.* Tradicionalmente, los ILSAs han enfocado sus esfuerzos en un rango limitado de dimensiones de aprendizaje (i.e. matemáticas, lectura, ciencias). No obstante, existen algunas iniciativas que conceptualizan la educación y sus fines de una manera más amplia. Algunos ejemplos son el Estudio Internacional de Educación Cívica y Ciudadanía (ICCS) de la IEA, o las mediciones de ciudadanía o competencia global incluidas en SEA-PLM y la más reciente edición de PISA. El artículo de [Villalobos y otros](#) utiliza datos de ICCS para estudiar la relación entre diferentes aspectos de la formación ciudadana de los estudiantes y tres políticas públicas implementadas en esta área entre 2008 y 2012 en Chile.

as ILSAs uma ferramenta fundamental para o monitoramento e direcionamento das políticas educacionais.

No entanto, é imprescindível reconhecer os limites das ILSAs para poder gerar as oportunidades que nos permitam utilizar essas informações de forma efetiva para subsidiar ações voltadas à melhoria de nossos sistemas educacionais. A recente publicação de novos dados internacionais (por exemplo, PISA em 2020) e a publicação iminente de dados regionais (por exemplo, ERCE em 2021) representam uma excelente oportunidade para enfrentar este desafio. Ao mesmo tempo, este é um dos principais objetivos desta monografia. De acordo com um relatório da UNESCO (2019) sobre este tema, existem cinco desafios principais que, como uma comunidade interessada na educação, devemos enfrentar para garantir o uso desses dados para tomar decisões fundamentadas que melhorem o ensino e a aprendizagem:

- *Limitações quanto à conceitualização de educação e aprendizagem.* Tradicionalmente, as ILSAs concentram seus esforços em uma gama limitada de dimensões de aprendizagem (ou seja, matemática, leitura, ciências). No entanto, existem algumas iniciativas que conceituam a educação e suas finalidades de forma mais ampla. Alguns exemplos são o Estudo Internacional de Educação Cívica e Cidadania (ICCS) da IEA, ou as medições de cidadania ou competência global incluídas no SEA-PLM e na edição mais recente do PISA. O artigo de [Villalobos e outros](#) utiliza dados do ICCS para estudar a relação entre diferentes aspectos da formação para a cidadania dos estudantes e três políticas públicas implementadas nesta área entre 2008 e 2012, no Chile.

- *Limitaciones en cuanto a la conceptualización de la calidad de la educación.* Los resultados del aprendizaje medidos por los ILSAs constituyen un indicador clave de la calidad educativa. Sin embargo, existe un consenso sobre la imposibilidad de interpretar estos resultados de manera aislada o como la única medición válida de la calidad de la educación. Mas bien, como apuntan los autores del reporte, deben concebirse en relación con el contexto en el que tienen lugar las dinámicas de enseñanza y aprendizaje. Tres artículos de este monográfico atienden esta limitación al centrar sus análisis en diferencias de género y etnicidad en Brasil y México ([Miranda y Torres](#)), la relación entre trabajo infantil, género y logro educativo ([Middel y otros](#)), y en la comparabilidad de las escalas para medir nivel socioeconómico entre los países participantes en estos estudios ([Erylmaz y otros](#)).
- *La infrautilización de los datos de las evaluaciones.* Los ILSAs corren el riesgo de convertirse en una inversión inútil de recursos y/o en una pérdida de oportunidades si sus resultados no son analizados adecuadamente y utilizados para informar acciones de política pública de manera oportuna. Este riesgo es aún más presente en nuestros países. Basta hacer una comparación entre el número de artículos y reportes producidos a partir de los datos de ILSAs en los países de Europa o Norteamérica y en nuestros países para darse cuenta de que este riesgo está aún más presente en Iberoamérica. Todos los artículos de este monográfico abordan esta limitación al centrar el foco de sus análisis en los países de la región. De manera particular lo hacen [Miranda y Treviño](#) al utilizar datos de TERCE (un estudio regional) para identificar los factores que influyen el aprendizaje en ciencias
- *Limitações quanto à conceitualização da qualidade da educação.* Os resultados de aprendizagem medidos pelas ILSAs constituem um indicador-chave da qualidade educacional. Porém, há consenso sobre a impossibilidade de interpretar esses resultados isoladamente ou como única medição válida da qualidade da educação. Em vez disso, como apontam os autores do relatório, eles devem ser concebidos em relação ao contexto em que ocorrem a dinâmica de ensino e aprendizagem. Três artigos desta monografia abordam essa limitação, ao concentrar suas análises nas diferenças étnicas e de gênero no Brasil e no México ([Miranda e Torres](#)), a relação entre trabalho infantil, gênero e realização educacional ([Middel e outros](#)), e na comparabilidade das escalas para medir o nível socioeconômico entre os países que participam desses estudos ([Erylmaz e outros](#)).
- *A subutilização dos dados das avaliações.* As ILSAs correm o risco de se tornarem um investimento inútil de recursos e/ou uma perda de oportunidades se seus resultados não forem adequadamente analisados e utilizados para configurar ações de políticas públicas de maneira oportuna. Esse risco está ainda mais presente em nossos países. Basta fazer uma comparação entre o número de artigos e relatórios produzidos a partir dos dados das ILSAs nos países da Europa ou da América do Norte e em nossos países para perceber que esse risco está ainda mais presente na Ibero-América. Todos os artigos desta monografia abordam essa limitação, concentrando sua análise nos países da região. Em particular, o fazem [Miranda e Treviño](#) ao utilizar dados do TERCE (um estudo regional) para identificar os fatores que influenciam a aprendizagem de ciências no Chile

en Chile y Paraguay; así como [Silveira](#) al utilizar estos mismos datos para estudiar los determinantes del logro en lectura en Uruguay.

- *La sobreutilización de los datos de las evaluaciones.* Centrarse demasiado en los resultados numéricos de las evaluaciones u otorgarles un crédito excesivo puede tener un efecto de desviación de las prioridades por parte de las diversas partes interesadas. Esto ocurre cuando la acción política se orienta única o principalmente a mejorar las cifras o indicadores, en lugar de atender las causas que subyacen a los problemas detectados por la evaluación. Del mismo modo, la incorporación de los datos de las evaluaciones en programas de financiación basados en los resultados (cada vez más utilizados por la comunidad de donantes), puede resultar en estrategias orientadas exclusivamente a ganar posiciones en los *rankings*, fomentando de esta manera prácticas de préstamo o emulación de políticas mal fundamentadas. Los artículos de [Sudbrack y Ramos](#) y de [Gesqui](#) abordan este problema; el primero al discutir la regulación supranacional operada por PISA; y, el segundo al hacer un análisis de tendencias de los resultados de la participación de los países iberoamericanos en este mismo estudio. De manera similar, el artículo de [López y García](#) deja la superficialidad de los resultados para hacer un análisis comparativo, profundo y sistemático de los sistemas educativos de España y Portugal, utilizando los datos ofrecidos anualmente por la OCDE en la publicación *Education at Glance*.
- *La combinación de las evaluaciones con la rendición de cuentas.* Finalmente, otro reto importante se da cuando algunos actores, como los docentes, directores de escuela o autoridades educativas son premiados o sancio-

e no Paraguai; assim como [Silveira](#) ao usar esses mesmos dados para estudar os determinantes do aproveitamento em leitura no Uruguai.

- *O uso excessivo de dados das avaliações.* Concentrar-se demais nos resultados numéricos das avaliações ou dar-lhes um crédito excessivo pode ter um efeito desviante das prioridades por parte dos vários interessados. Isso ocorre quando a ação política visa única ou principalmente a melhoria dos números ou indicadores, ao invés de abordar as causas subjacentes aos problemas detectados pela avaliação. Da mesma forma, a incorporação dos dados das avaliações em programas de financiamento baseados nos resultados (cada vez mais utilizados pela comunidade de doadores), pode resultar em estratégias orientadas exclusivamente a ganhar posições nos *rankings*, estimulando assim práticas de empréstimo ou emulação de políticas mal fundamentadas. Os artigos de [Sudbrack e Ramos](#) e de [Gesqui](#) abordam este problema; o primeiro, ao discutir a regulação supranacional operada pelo PISA; e, o segundo, fazendo uma análise de tendências dos resultados da participação dos países ibero-americanos neste mesmo estudo. Da mesma forma, o artigo de [Lopez e Garcia](#) abandona a superficialidade dos resultados para fazer uma análise comparativa, profunda e sistemática dos sistemas educativos da Espanha e Portugal, utilizando os dados disponibilizados anualmente pela OCDE na publicação *Education at Glance*.
- *A combinação das avaliações com a prestação de contas.* Por fim, outro desafio importante ocorre quando alguns atores, como professores, diretores de escolas ou autoridades educacionais, são premiados ou sancionados com

nados en función del desempeño de los estudiantes en ILSAs. Algunos de los efectos no previstos de estas dinámicas incluyen la ‘enseñanza para la prueba’ y la reducción de los contenidos del currículum. Este tema es abordado en el artículo de [Ferrerira y otros](#), que analiza los efectos de los ILSAs en la vida escolar para el caso de la ciudad brasileña de Manaus-AM. El tema es también discutido por [Compañ](#) al analizar la política de ingreso al servicio docente implementada en México y su impacto en los resultados de PISA 2018.

Resulta claro que los ILSAs son una invaluable herramienta para la mejora de nuestros sistemas educativos, pero que conllevan también importantes riesgos. Si bien, las limitaciones y retos discutidos en esta editorial no hacen que los ILSAs sean inherentemente problemáticos, existen motivos para mantenernos atentos al diseño de las evaluaciones y al uso de la evidencia que resulta de ellas. Las evaluaciones internacionales están aquí para quedarse, y por ello es importante determinar de qué manera podemos optimizar el uso de la información que producen. En cualquier caso, nos parece que el conjunto de artículos incluidos en este monográfico contribuye de manera significativa a superar los retos y a aprovechar las oportunidades que ofrecen las evaluaciones internacionales a gran escala.

base no desempenho dos alunos nas ILSAs. Alguns dos efeitos imprevistos dessas dinâmicas incluem o ‘ensinar para a prova’ e reduzir o conteúdo do currículo. Este tópico é abordado no artigo de [Ferreira e outros](#), que analisa os efeitos das ILSAs na vida escolar para o caso da cidade brasileira de Manaus-AM. O assunto também é discutido por [Compañ](#) ao analisar a política de admissão ao serviço de ensino implementada no México e seu impacto nos resultados do PISA 2018.

É claro que as ILSAs são uma ferramenta inestimável para melhorar nossos sistemas educacionais, mas também apresentam riscos significativos. Embora as limitações e desafios discutidos neste editorial não tornem as ILSAs intrinsecamente problemáticas, há razões para ficarmos atentos ao desenho das avaliações e ao uso das evidências que delas resultam. As avaliações internacionais vieram para ficar, por isso é importante determinar como podemos otimizar o uso das informações que elas produzem. Em todo caso, parece-nos que o conjunto de artigos incluídos nesta monografia contribui significativamente para superar os desafios e aproveitar as oportunidades oferecidas pelas avaliações internacionais em larga escala.

REFERENCIAS

- Cresswell, J., Schwantner, U. y Waters, C. (2015). *A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data*. Washington, D.C. / Paris: The World Bank / OECD Publishing. <http://dx.doi.org/10.1787/9789264248373-en>
- UNESCO (2019). *La promesa de las evaluaciones de aprendizaje a gran escala*. Paris: Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. <https://bit.ly/37ROTUe>

Evaluaciones internacionales a gran escala y ciudadanía. Explorando la relación entre políticas y ICCS en Chile (2009-2016)

Cristóbal Villalobos ¹ ; Ignacio Wyman ¹ ; Ernesto Treviño ¹ 

¹ Pontificia Universidad Católica de Chile (UC), Chile

Resumen. El artículo analiza cuatro atributos de la formación ciudadana de los estudiantes (conocimiento cívico, disposición al voto, confianza en los partidos políticos y disposición a compartir con personas de otras clases sociales) considerando la implementación de tres políticas desarrolladas en Chile: i) la Ley de Voto Voluntario (2012); ii) la Ley de Subvención Escolar Preferencial (2008) y; iii) los cambios curriculares a los Planes de Formación Cívica y Ciudadana (2011). Usando los datos de la evaluación ICCS de 2009 (N=5.192 estudiantes; 177 escuelas) y 2016 (5.081 estudiantes; 178 escuelas) y a través de análisis inferenciales y multinivel, los resultados muestran que existen pocas transformaciones en los promedios entre ambos años, aunque se evidencian cambios cuando se desagregan los datos entre niveles socioeconómicos. Estos resultados permiten reinterpretar la relación entre los ILSA y las políticas educativas, de manera que permitan pensar cómo las Evaluaciones Internacionales a Gran Escala podrían aportar al desarrollo de sistemas educativos para potenciar la ciudadanía en los estudiantes.

Palabras claves: evaluaciones internacionales a gran escala; ciudadanía; desigualdades, ICCS; Chile; política.

Avaliações internacionais em larga escala e cidadania. Explorando a relação entre políticas e o ICCS no Chile (2009-2016)

Resumo. O artigo analisa quatro atributos da formação cidadã dos estudantes (conhecimento cívico, disposição para o voto, confiança nos partidos políticos e disposição para compartilhar com pessoas de outras classes sociais), considerando a implementação de três políticas desenvolvidas no Chile: i) a Lei de Voto Voluntário (2012); ii) a Lei de Subsídio Escolar Preferencial (2008) e; iii) as mudanças curriculares nos Planos de Formação Cívica e Cidadã (2011). Usando os dados da avaliação ICCS de 2009 (N=5.192 estudantes; 177 escolas) e 2016 (5.081 estudantes; 178 escolas) e através de análises estatísticas inferenciais e multinível, os resultados mostram que existem poucas alterações nas médias entre os dois anos, embora evidenciem mudanças ao desagregar os dados entre níveis socioeconômicos. Estes resultados permitem reinterpretar a relação entre os ILSA e as políticas educacionais, permitindo pensar como as Avaliações Internacionais em Larga Escala podem contribuir para o desenvolvimento de sistemas educativos para promover a cidadania nos estudantes.

Palavras-chave: avaliações internacionais em larga escala; cidadania; desigualdades, ICCS; Chile; política.

International large-scale assessment and citizenship. Exploring the relationship between politics and ICCS in Chile (2009-2016)

Abstract: This article analyzes four attributes of student citizenship training (civic knowledge, willingness to participate in elections, trust in political parties and willingness to share with people from other social classes) considering the implementation of three policies developed in Chile in recent years: i) the Voluntary Voting Act (2012); ii) the Preferential School Subsidy Act (2008) and; iii) curricular changes in Civic and Citizenship (2011). Using data from Chile from ICCS study in 2009 (N = 5,192 students; 177 schools) and 2016 (5,081 students; 178 schools), and through inferential and multilevel analyzes, the results show that there are few changes in national averages between both years, although changes are evident when the data is disaggregated by socioeconomic level. These results allow us to re-interpret the relationship between ILSAs and educational policies, to think about how International Large-Scale Assessments may contribute to the development of educational systems that empower students for citizenship.

Keywords: International Large-Scale Assessment; Citizenship; ICCS; Inequalities, Chile; Politics.

1. Introducción

En las últimas décadas, las Evaluaciones Internacionales a Gran Escala o *International Large-Scale Assessments* (ILSA) han adquirido una creciente importancia en los sistemas educativos del mundo. Este crecimiento puede relacionarse con tres fenómenos convergentes. En primer término, las ILSA se han promovido y potenciado por instituciones supraestatales (organismos internacionales y empresas multinacionales), que buscan difundir, coordinar y promover; pero también restringir, coaccionar o limitar políticas gubernamentales, a través de los llamados procesos de “gobernanza suave” (Bieber, 2016; Niemann y Martens, 2018). En segundo lugar, las ILSA se han expandido a través de la relevancia de lo que Ozga (2008) ha llamado el gobierno del conocimiento, esto es, la tendencia de parte importante de los países a organizar sus políticas a través de procesos de producción y utilización de datos (Grek, 2009), promoviendo procesos de competencia, emulación y/o aprendizaje entre los países (Parcerisa, Fontdevila, y Verger, 2020). Finalmente, las ILSA se han consolidado debido al consenso (hegemónico) sobre la relevancia de la evaluación educativa como un pilar central del proceso escolar (Gorur, 2017; Wagemaker, 2013), que ha promovido no solo la legitimidad sino también la necesidad de estos instrumentos.

16

Uno de los resultados más llamativos de estos procesos ha sido la multiplicación en intensidad y alcance de estas evaluaciones. Así, en las últimas décadas, las ILSA han incorporado progresivamente a países con menores niveles de desarrollo social (por ejemplo, PISA-D, incorpora a Camboya, Ecuador, Guatemala, Honduras, Panamá, Paraguay, Senegal y Zambia) y a regiones con menos información hasta comienzos del siglo, como Latinoamérica y África (LLECE, SACMEQ y PASEC). En términos de las áreas, las ILSA han involucrado tópicos educativos tan diversos como matemáticas, ciencias, lectura, escritura, habilidades computacionales, competencias digitales, educación financiera y educación cívica y ciudadanía, levantando gran cantidad de información sobre niños, niñas y jóvenes, pero también sobre docentes (en formación y en ejercicio), directores y población adulta.

Este crecimiento explosivo ha motivado una fértil discusión académica sobre los impactos de las ILSA en la política educativa, el currículo, los procesos de enseñanza y las capacidades de medición en distintos países. Aunque se está lejos de alcanzar un consenso sobre los aportes, dificultades y tensiones que generan estos mecanismos, lo cierto es que gran parte de los actores han destacado que las ILSA serían un agente activo de cambio de las políticas (Heyneman y Lee, 2014; Wagemaker, 2013), especial (pero no exclusivamente) en áreas como matemáticas y lectura, que son las áreas más sensibles para la rendición de cuentas (también llamado *accountability*) de los sistemas educativos. En estos casos, se ha mostrado que los ILSA son

entendidos como una fuente resumida y comparada de la calidad general del sistema escolar, siendo interpretada como una medida del nivel y prestigio educativo de los países (Carnoy, Khavenson e Ivanova, 2013; Sellar, Lingard, Rutkowski y Takayama, 2018).

Sin embargo, la relación entre las políticas gubernamentales y las ILSA que no están directamente relacionados con el *accountability* escolar ha sido mucho menos explorada, surgiendo distintas interrogantes a este respecto. ¿Qué influencia pueden tener las ILSA en ámbitos no centrales de la rendición de cuentas escolar? ¿Cuál es el uso que los gobiernos hacen de las ILSA no relacionados con el *accountability* escolar? ¿Cómo, y en qué medida, la implementación de políticas nacionales ha influido en los resultados de evaluaciones internacionales sobre formación cívica y ciudadanía juvenil? Para aproximarnos a estos interrogantes, este artículo se focaliza en una evaluación internacional particular, el *International Civic and Citizenship Education Study* (ICCS) en un país, Chile. Así, el estudio explora los resultados de ICCS en 2009 y 2016 en cuatro variables de la formación cívica y ciudadana de los estudiantes (conocimiento cívico, disposición al voto, confianza en los partidos políticos y disposición a compartir con personas de otras clases sociales) considerando la implementación de tres políticas que podrían haber impactado en estas mediciones: i) la Ley de Voto Voluntario (2012); ii) la Ley de Subvención Escolar Preferencial (2008) y; iii) los cambios curriculares a los Planes de Formación Cívica y Ciudadana (2011). Particularmente, el estudio tiene tres objetivos específicos: i) comparar los resultados promedio entre 2009 y 2016 en las cuatro variables mencionadas; ii) desagregar los resultados de estas variables en términos socioeconómicos y; iii) analizar la variación del efecto del nivel socioeconómico, pero considerando la influencia de una serie de otras variables, entre los dos años.

Para dar cuenta de estos objetivos, el artículo se estructura en seis apartados además de esta introducción. El segundo apartado describe el ICCS y el impacto de esta prueba en los sistemas educativos. El tercero, se focaliza en describir las características del sistema chileno, con un énfasis en el rol de las ILSA de educación cívica y ciudadanía. El cuarto apartado describe las políticas seleccionadas, argumentando cómo podrían relacionarse con las mediciones de la ICCS. En la quinta sección se describen las variables utilizadas, el método de estimación y análisis realizados. La sexta sección expone los resultados, focalizándose en el cambio de las variables, su distribución y el peso de distintos factores en el tiempo. Finalmente, la séptima sección desliza las principales conclusiones, aportes y limitaciones del estudio.

2. El ICCS y su impacto en el desarrollo de políticas educativas

Administrado por la *International Association for the Evaluation of Educational Achievement* (IEA), el ICCS puede ser considerado como la ILSA de mayor alcance sobre formación cívica y ciudadana en el mundo¹. Implementada originalmente en 2009, el estudio vuelve a aplicarse en 2016 y se espera una tercera ronda para 2022², el ICCS recolecta información sobre estudiantes de 13,5 años en promedio. El instrumento recoge información sobre conocimiento, actitudes y acciones de los estudiantes sobre la ciudadanía y la vida cívica (Schulz, Fraillon, Ainley, Losito y Kerr, 2008). Además, incluye información sobre el desarrollo de la educación cívica y ciudadana en la escuela, las experiencias, clima escolar, habilidades y prácticas docentes.

Al contrario de lo que ocurre con otras pruebas como PISA, TALIS o TIMSS, el análisis de la relación entre ICCS y la implementación de políticas educativas y sociales es un ámbito de análisis poco explorado. Las investigaciones existentes al respecto han señalado que el ICCS y el CIVED ha permitido entender y gestionar los procesos de desafección de la identidad nacional, así como los problemas de identidad cultural y étnica de los países (Wagemaker, 2011). Asimismo, se ha destacado que los resultados de las ILSA de educación cívica y ciudadanía han motivado la priorización educativa sobre temas cívicos (por ejemplo, a través de la creación de un curso específico de educación cívica y ciudadana) en ex países comunistas como Estonia, Lituania, Polonia y Eslovenia (Klemencic, 2010). Finalmente, y en términos curriculares, los resultados de ICCS y CIVED motivaron cambios o nuevos planes de estudio en Eslovaquia y Lituania (Pelgrum, Blahova, Dukynaitė, Pavešić y Kangro, 2011) pero también influyeron en otros contextos, como Australia (Wagemaker, 2011) y México (Cox, en prensa).

3. El sistema escolar chileno. Marketización, accountability e ILSA

Dentro del concierto internacional, el sistema escolar chileno es considerado como un caso extremo de marketización, privatización y consolidación de un sistema universal de *vouchers* (Villalobos y Quaresma, 2015). Una de sus características más sobresalientes de este sistema es la existencia

¹ En 2009 participaron 35 países, mientras que en 2016 fueron 24. Para la ronda de 2022 participarán 25 países. En todas ellas, el instrumento incluye mayoritariamente países de Europa, y algunos países de América Latina y Asia.

² En 1971 y 1999, la misma IEA administró la prueba CIVED (*Civic Education Study*), que también versó sobre aspectos relacionados con la ciudadanía, la democracia, la identidad nacional y la cohesión social, y que puede considerarse como antecedente directo del ICCS.

de un sistema de rendición de cuentas ligado a pruebas estandarizadas, que premia y castiga a las escuelas según sus resultados (Parcerisa y Falabella, 2017). Por ello, no extraña que Chile sea un terreno fértil para el desarrollo de las ILSA. Actualmente, Chile participa de siete ILSAs: PISA, ICILS, PIRS, TIMMS, ERCE, ECES y ICCS.

La relación entre las ILSA y las políticas educativas en Chile ha sido recientemente analizada. Así, por ejemplo, se ha evidenciado como los resultados de PISA han sido utilizados tanto para entender y describir los “avances” y “estancamientos” del sistema escolar durante las tres primeras décadas desde el retorno a la democracia (Cox y Meckes, 2016). Adicionalmente, los datos de PISA (y en menor medida de TIMSS) han servido como fundamento para el desarrollo de reformas relevantes, como la Ley General de Educación en 2009 o la Ley de Inclusión Escolar en 2015 (Alarcón, 2020; Parcerisa y Verger, 2019). Finalmente, los datos de PISA han sido utilizados para discutir sobre “sociedades de referencia”, esto es, debatir en torno a los cambios que el sistema educativo chileno debiera desarrollar, teniendo como puntos de referencia países con mejor desempeño, como Finlandia (Parcerisa y Verger, 2019) u otros con modelos de gobernanza semejantes al chileno, como Holanda (Alarcón, 2020).

Respecto de la relación entre las ILSA de educación cívica y ciudadanía y las políticas educativas chilenas, la literatura ha identificado tres impactos. Por una parte, se ha indicado que los marcos de las pruebas internacionales (CIVED e ICCS) han sido una herramienta para apoyar cambios curriculares sobre este tópico en el país durante la década del 2000 (Cariola, Covacevich, Gubler, Lagos y Ortiz, 2011; Cox, en prensa). En segundo término, la existencia de estas pruebas ha fomentado una explotación gubernamental de los datos, permitiendo mejorar la comprensión sobre la materia, desarrollar comisiones de expertos (la Comisión Formación Ciudadana en el 2004) y promover la cooperación internacional (Cariola *et al.*, 2011) sobre estos tópicos. Finalmente, se ha indicado que los resultados de ICCS estarían relacionados con reformas curriculares recientes, como la creación de los Planes de Formación Ciudadana y el Curso de Formación Cívica y Ciudadana para los últimos dos cursos de la secundaria (Cox, en prensa).

Respecto de la relación entre las ILSA de educación cívica y ciudadanía y las políticas educativas chilenas, la literatura ha identificado tres impactos. Por una parte, se ha indicado que los marcos de las pruebas internacionales (CIVED e ICCS) han sido una herramienta para apoyar cambios curriculares sobre este tópico en el país durante la década del 2000 (Cariola, Covacevich, Gubler, Lagos y Ortiz, 2011; Cox, en prensa). En segundo término, la existencia de estas pruebas ha fomentado una explotación gubernamental de los datos, permitiendo mejorar la comprensión sobre la

materia, desarrollar comisiones de expertos (la Comisión Formación Ciudadana en el 2004) y promover la cooperación internacional (Cariola *et al.*, 2011) sobre estos tópicos. Finalmente, se ha indicado que los resultados de ICCS estarían relacionados con reformas curriculares recientes, como la creación de los Planes de Formación Ciudadana y el Curso de Formación Cívica y Ciudadana para los últimos dos cursos de la secundaria (Cox, en prensa).

4. Tres políticas en torno a la ciudadanía de la juventud de la última década

Desde inicios del 2000, en Chile se han desarrollado tres políticas que han buscado transformar el desarrollo cívico y ciudadano de los estudiantes del país. La Tabla 1 describe la política, el ámbito de acción en el que se desenvuelve y las variables del ICCS desde las cuales se podría analizar este cambio, dando cuenta así del marco de análisis desde el que se organiza este artículo.

Tabla 1. Ámbito de acción, política y variables ICCS (2009-2016)

Ámbito	Política	Variable ICCS
Actitudes cívicas hacia la institucionalidad política	Ley de Voto Voluntario (2012)	Disposición a votar Confianza en los partidos políticos
Actitudes ciudadanas hacia la convivencia con otros	Ley de Subvención Escolar Preferencial (2008)	Disposición a convivir con otros de distinto nivel socioeconómico
Conocimiento cívico y Ciudadano	Reformas Curriculares (2009 y 2012)	Conocimiento cívico

Fuente: Elaboración propia.

En primer lugar, la Ley de Voto Voluntario (2012) buscó enfrentar el creciente proceso de abstención electoral de la juventud (Carlin, 2006) y la creciente desconfianza en los partidos políticos (INJUV, 2009). Para ello, la Ley transformó el sistema de inscripción desde un registro voluntario y voto obligatorio a uno de registro automático y voto voluntario, derribando así las supuestas “barreras de acceso” para la participación juvenil. Aunque estudios empíricos han mostrado que este cambio no ha generado transformaciones en los patrones de votación de estas nuevas generaciones (Corvalán y Cox, 2013), no se ha explorado suficientemente cómo esta política podría haber afectado las actitudes hacia la institucionalidad política de los jóvenes que aún no tienen el derecho legal a ejercer el voto, como sí permite el ICCS en sus mediciones 2009 y 2016.

En segundo lugar, y considerando los altos niveles de segregación y polarización del sistema educativo (Valenzuela, Bellei y de los Ríos, 2014) la Ley de Subvención Escolar Preferencial, SEP (2008) entregó más recursos a las escuelas de acuerdo a la vulnerabilidad de sus estudiantes, bajo la condición de eliminar las barreras de ingreso existentes (Weinstein y Villalobos, 2016)³. Aunque la evidencia ha mostrado que esta ley ha mejorado los niveles de heterogeneidad sólo en los grupos medios-bajos (Valenzuela, Villaroel y Villalobos, 2013), no se ha explorado mayormente cómo estos cambios podrían estar motivando transformaciones en las actitudes de los jóvenes hacia los otros, lo que es posible realizar gracias a los datos del ICCS.

Finalmente, en el ámbito curricular, y tras importantes cambios que buscaron anclar el currículum escolar hacia ámbitos como la institucionalidad política (Cox y García, 2017), durante 2009 y 2012 se realizaron una serie de modificaciones al currículum en el área de educación cívica y ciudadanía (MINEDUC, 2013). En ambos años, se buscó reforzar la asignatura de Ciencias Sociales e Historia, alineando los contenidos curriculares a los desafíos evidenciados por CIVED 1999 en aspectos como los riesgos de la democracia y los nuevos desafíos de la ciudadanía (Cox, en prensa). En este sentido, estas transformaciones podrían haber promovido cambios en los resultados de conocimiento cívico y en la capacidad de entender los problemas cívicos y ciudadanos de los estudiantes, lo que es medido a través del conocimiento cívico en las pruebas ICCS de 2009 y 2016.

5. Metodología

5.1 Datos

En este estudio utilizamos los datos para Chile de las versiones de ICCS 2009 y 2016. ICCS es un estudio internacional que recoge información sobre conocimiento y entendimiento de conceptos relacionados con la ciudadanía, así como con las creencias, actitudes y comportamientos de los estudiantes (Schulz, Fraillon, Ainley, Losito y Kerr, 2008). Además, recoge información contextual sobre la enseñanza, organización y prácticas escolares relacionadas con la educación cívica y ciudadana en la escuela. El cuestionario lo compone un módulo general de preguntas y otro específico para las regiones de los países participantes. La recolección de datos ocurre en dos etapas, dentro de cada país. En la primera, se selecciona una muestra representativa de las escuelas con una probabilidad basada en el tamaño de su matrícula.

³ Estas prohibiciones se extendieron para el conjunto del sistema en el 2015, con la aprobación de la Ley de Inclusión Escolar (LIE).

En un segundo momento, dentro de cada escuela se selecciona al azar una clase del grado de interés (octavo grado, en el caso chileno). El cuestionario es respondido por estudiantes, directores de escuela y una muestra fija de docentes seleccionados azarosamente entre aquellos que imparten clases al grado de interés. Para el caso chileno, la muestra se compone en 2009 de 5.192 estudiantes pertenecientes a 177 establecimientos, mientras que en 2016 se compone por 5.081 estudiantes de 178 escuelas.

5.2 Variables de interés

a) *Variables dependientes:* La investigación se focaliza en cuatro variables de interés. Por una parte, y relacionado con la implementación de la Ley del Voto Voluntario, se utilizan dos variables: i) la **confianza en partidos políticos**, que señala el grado de confianza declarado por los estudiantes en estas organizaciones, con un rango de respuesta que va de 0 (no confía) a 3 (confía completamente) y; ii) la **participación electoral esperada**, que es un índice normalizado para cada país (media 50 y desviación estándar 10 dentro de cada país) creado a partir de tres indicadores: la disposición a votar en futuras elecciones locales, la disposición a votar en elecciones nacionales, y la disposición a informarse sobre los candidatos antes de participar de una elección, donde mayores valores indican una disposición más alta de participación. La tercera variable es la **disposición a compartir con otros de distinta clase social**, que indica el porcentaje de estudiantes que señalan no tener inconvenientes con compartir en un mismo barrio con personas de distinta clase social⁴, y se utiliza considerando los objetivos de la Ley de Subvención Escolar Preferencial (SEP). Finalmente, para estudiar el posible cambio de las modificaciones curriculares en las áreas de formación cívica y ciudadana, se utiliza el **conocimiento cívico**, que se mide a través de 87 ítems y se reporta a través de cinco valores plausibles con media 500 y desviación estándar 100 (ambos valores considerando el total de países participantes en el estudio). Puntajes más altos en el indicador muestran mayor conocimiento cívico de los estudiantes.

⁴ Este indicador se construye siguiendo un enfoque de “análisis de datos integrados” (Rioux y Little, 2020), el cual avala la reconstrucción de indicadores que, difiriendo en la forma exacta de medición entre años, presenta información superpuesta. Así, para el año 2009 el estudiante debía responder entre las siguientes alternativas: “me gustaría”, “no me importaría” y “me disgustaría”. En el 2016, en cambio: “sí me molesta” y “no me molesta”. Para identificar la disposición a compartir con otros de distinta clase social entre estudiantes encuestados el 2009 se considera las alternativas “me gustaría” y “no me molestaría”, y “no me molesta”, para 2016.

b) Variable independiente: Las variables dependientes son analizadas en referencia al nivel socioeconómico del hogar del estudiante. Diversos estudios internacionales han señalado al nivel socioeconómico como uno de los factores más relevantes para explicar el conocimiento y comportamiento cívico de estudiantes (Isac, Maslowski, Creemers y van der Werf, 2014; Treviño, Béjares, Villalobos y Naranjo, 2017). Además, es clave cuando la pregunta -como es este caso- se centra en comprender las relaciones entre políticas y evaluaciones de gran escala. Como señalan algunos autores, la atención a las medias nacionales en resultados educativos puede conducir a inferencias erradas sobre los efectos que tienen distintas políticas (Torney-Purta y Amadeo, 2013). Por lo mismo, desagregar estos resultados entre grupos sociales es clave, pues de lo contrario no se estaría reconociendo el altamente probable efecto diferenciado que muchas políticas (si no todas) tienen en sub-grupos poblacionales (Carnoy *et al.*, 2013). En el caso del ICCS, el nivel socioeconómico se construye como un índice a partir de información del estatus ocupacional, el nivel educativo de los padres, y el número de libros en el hogar. Para diferenciar claramente grupos que puedan presentar escenarios disímiles, para esta investigación se decidió trabajar con quintiles de esta variable.

c) Variables de control: Estas variables representan un conjunto de características relacionadas a los individuos, al contexto y a los procesos educativos (Torney-Purta y Amadeo, 2013) que han sido anteriormente incluidas en modelos teórico-conceptuales (Knowles, Torney-Purta y Barber, 2018) y empíricos sobre efectividad escolar en el contexto de la educación cívica (Isac *et al.*, 2014; Treviño *et al.*, 2017; Treviño, Villalobos, Béjares y Naranjo, 2019). Así, se incluyen variables críticas a nivel estudiante (género, el interés de los padres en política y la frecuencia con que los jóvenes discuten asuntos políticos fuera de la escuela) y de la escuela (nivel socioeconómico y dependencia pública o privada) y variables relacionadas con procesos educativos (apertura de la clase para la discusión), que han sido señaladas repetidamente como factores críticos para evaluar distintas variables cívicas (Knowles *et al.*, 2018)

5.3 Estrategia de análisis

El análisis se desarrolla a través de un proceso organizado en tres fases. En primer lugar, y a través de análisis inferencial, se discuten los cambios en el tiempo de las medias de las cuatro variables en 2009 y 2016. En un segundo momento, las medias se desagregan según el nivel socioeconómico de los estudiantes. De esta forma, se busca explorar posibles diferencias asociadas al capital socioeconómico de los jóvenes en las distintas variables estudiadas, buscando entender el posible efecto diferenciado de la

política según posición social (Carnoy *et al.*, 2013). Finalmente, la tercera fase presenta modelos multinivel⁵, considerando la estructura de los datos de la prueba (estudiantes anidados a escuelas) que buscan analizar -para cada variable y para cada año- la variación del efecto del nivel socioeconómico, controlando por una serie de variables críticas tanto a nivel de estudiante como de escuela. De esta forma, se busca explorar si la magnitud del efecto podría estar cambiando entre 2009 y 2016 en los cuatro indicadores analizados. Para cada variable dependiente se especifican dos modelos por año. En su primera especificación sólo se incluye como predictor el nivel socioeconómico y el género del estudiante. En la segunda, se incluyen otras variables de control a nivel del estudiante y de la escuela. En todos los casos, los análisis se realizan usando el software R, usando los pesos muestrales correspondientes.

6. Resultados

La Tabla 2 presenta la media de las cuatro variables dependientes para 2009 y 2016. Como se puede observar, en dos variables (conocimiento cívico y disposición a participar en elecciones) no se observan diferencias estadísticamente significativas entre 2009 y 2016. Respecto del conocimiento cívico, esto podría tener relación con que la implementación de reformas curriculares no es automática, sino que implica procesos de apropiación por parte de los docentes, que deben re-significar los cambios curriculares en sus propias prácticas (Ball, 2003). En el caso de la disposición al voto, la mantención de la disposición está en línea con la actuación que los patrones de votación que han mostrado los jóvenes, que se han mantenido inalterados a pesar del cambio en la legislación (Corvalán y Cox, 2013).

⁵ Los modelos multinivel permiten analizar el impacto de una o un conjunto de covariables en una variable dependiente considerando el anidamiento de estudiantes en escuelas (Hox, Moerbeek y Van de Schoot, 2017). De esta manera, se controla el efecto de pertenencia de un grupo de unidades a una de mayor escala, lo que en caso de omisión, podría llevar a inferencias incorrectas (error tipo 1). La especificación general del modelo es:

$$Y_{ijk} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij} \quad (1)$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + v_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{01} \quad (3)$$

donde Y es la variable dependiente y X el conjunto de covariables a nivel de estudiante (1). En (2), W representa las variables a nivel de escuela, mientras que (3) señala que los modelos estimados consideran pendientes fijas.

Tabla 2. Diferencias de las medias entre 2009 y 2016

Variable	Año medición	Media	Límite superior	Límite inferior
Conocimiento Cívico	2009	483,02	485,41	480,64
	2016	482,45	485,06	479,85
Disposición a participar en elecciones	2009	49,83	50,17	49,49
	2016	50,17	50,46	49,89
Confianza en partidos políticos	2009	2,80*	2,82	2,78
	2016	2,86*	2,88	2,83
Disposición a compartir con personas de distinta clase social	2009	95,05%*	95,64%	94,46%
	2016	90,41%*	91,23%	89,60%

Nota: *p<.05. Límite inferior y Límite superior de los intervalos al 95% de confianza⁶. Fuente: elaboración propia.

Por otra parte, se evidencia un aumento estadísticamente significativo (pero mínimo) en el nivel de confianza con los partidos políticos. Este aumento llama la atención, ya que distintos autores han mostrado la baja legitimidad de estas instituciones. Sin embargo, es posible hipotetizar que este aumento podría deberse a los cambios en los clivajes políticos experimentados en la última década que han generado el surgimiento de nuevas fuerzas y una mayor presencia de figuras jóvenes en la política nacional (Alvarado, Rivera y Morales, 2019). Finalmente, los datos revelan que la disposición a compartir con otros de distinta clase social ha disminuido entre 2009 y 2016. Hipotéticamente, podría pensarse que el aumento de la discusión sobre las desigualdades escolares promovidas por la SEP (pero también por otras legislaciones, como la LGE, y por intensos procesos de movilización social en 2006 y 2011) hayan profundizado temores subterráneos, disminuyendo la propensión a compartir con “otros”, discurso especialmente presente en los grupos medios (Canales, Bellei y Orellana, 2016).

Para indagar en las diferencias según nivel socioeconómico, la Tabla 3 describe las medias y límites (inferiores y superiores) para cada quintil obtenidos mediante regresiones con errores estándares robustos, mostrando importantes variaciones “escondidas” en la media poblacional.

⁶ Cuando ambos límites se encuentran en el mismo cuadrante (mantienen el mismo signo), los efectos son estadísticamente significativos al 95% de confianza.

Tabla 3. Medias según quintil.

Variable	NSE	ICCS 2009			ICCS 2016		
		Media	Límite superior	Límite inferior	Media	Límite superior	Límite inferior
Conocimiento cívico	1	442,62	453,67	431,57	435,93	444,93	426,93
	2	461,80	474,58	449,02	465,23	474,81	455,65
	3	483,78	495,87	471,69	483,66	493,70	473,62
	4	510,26	522,12	498,40	510,10	521,04	499,16
	5	548,36	562,32	534,40	544,07	555,93	532,21
Disposición a participar en elecciones	1	49,36	50,33	48,38	47,96	48,72	47,20
	2	49,03	50,25	47,81	49,08	50,08	48,08
	3	48,95	50,34	47,56	50,21	51,17	49,25
	4	49,90	51,39	48,41	51,21	52,29	50,13
	5	52,98	54,41	51,55	53,70	54,64	52,76
Confianza en partidos políticos	1	2,69	2,76	2,63	2,77	2,85	2,70
	2	2,75	2,83	2,67	2,85	2,95	2,75
	3	2,85	2,95	2,75	2,87	2,95	2,79
	4	2,90	3,00	2,80	2,92	3,02	2,82
	5	2,82	2,92	2,72	2,88	3,00	2,76
Disposición a compartir con personas de distinta clase social	1	93,3%*	95,01%	91,58%	87,88%*	90,17%	85,59%
	2	95,60%	97,68%	93,52%	90,85%	93,93%	87,77%
	3	95,86%	97,98%	93,74%	92,29%	95,21%	89,37%
	4	95,44%	97,67%	93,21%	92,11%	95,01%	89,21%
	5	93,80%	96,23%	91,37%	91,67%	95,06%	88,28%

Nota: *p<.05. Límite inferior y Límite superior de los intervalos al 95% de confianza
Fuente: Elaboración propia a partir de datos ICCS 2009 y 2016.

Así, pareciera ser que el cambio en la Ley de Voto Voluntario podría haber afectado de manera distinta la disposición a participar en elecciones en el futuro. La disposición a esta acción cívica habría disminuido en el quintil 1 y aumentado en la clase media y alta (quintiles 3, 4 y 5) aunque no de manera estadísticamente significativa. Esto podría estar indicando, tal como han mostrado otros estudios, que la generación de mayores niveles de voluntariedad podría estar afectando las posibilidades de participación política de los jóvenes más vulnerables, que no verían en el voto un acto que pudiera transformar la realidad social (Corvalán y Cox, 2013).

En segundo lugar, la desagregación por quintiles muestra que el aumento promedio en la confianza en los partidos políticos está explicado principalmente por un incremento estadísticamente significativo entre los

estudiantes del primer quintil. Tal hallazgo es complementario al anterior, ya que se visualiza que, en el quintil más bajo existiría, al mismo tiempo, una disminución de la disposición a votar y un aumento en la confianza en los partidos políticos. Se trata de una compleja combinación que podría incentivar el clientelismo entre los partidos políticos para promover el voto y que nos parece que es necesario explorar con mayor profundidad en el futuro.

Por otra parte, la disposición a compartir con otros de distinta clase social disminuyó en forma generalizada entre los quintiles 1 al 4, y con especial fuerza en el primer quintil. En contraste, en el quintil 5 no hay diferencias significativas en torno a esta variable. Estos resultados podrían estar dando cuenta de la plausibilidad de la hipótesis de “reacción” en un contexto de polarización del sistema escolar (Villalobos y Valenzuela, 2012), ya que es precisamente en el quintil 5 donde los cambios impulsados por la SEP (y luego profundizados por la LIE) no operarían en su totalidad, ya que estas regulaciones no afectan a los establecimientos privados sin subvención estatal que están concentrados en este quintil. En contraste, es en los sectores más vulnerables (quintil 1) donde la política podría haber desarrollado más transformaciones en la convivencia con sectores distintos, lo que en parte podría explicar esta disminución a la disposición a compartir con otros sectores sociales.

Finalmente, y para explorar el peso relativo de los factores asociados a distintos resultados cívicos entre 2009 y 2016 se presentan los resultados de los modelos multinivel. Las Tablas 4 y 5 presentan los resultados de las variables relacionadas con el cambio en la Ley de Voto Voluntario: disposición a participar en las elecciones y confianza en los partidos políticos. La Tabla 4 se focaliza en la disposición a participar en las elecciones. Se puede observar, en primer término, que las variables individuales adquieren significancia en 2016 en los modelos que presentan sólo las variables de nivel socioeconómico, además de género (Modelos 1 y 3). Sin embargo, tanto en 2009 como en 2016 los modelos que controlan por características individuales y de la escuela (Modelos 2 y 4) muestran que el género y nivel socioeconómico no son estadísticamente significativas. En contraste, la mayoría de las variables de control (interés de padres en política, discutir sobre política fuera de la escuela, discusión abierta en la sala de clases) son significativas en ambos años y mantienen magnitudes relativamente iguales. Finalmente, se puede observar que el nivel socioeconómico a nivel de la escuela adquiere significancia sólo en 2016, lo que podría estar indicando que la composición socioeconómica podría ser una variable que haya adquirido mayor relevancia en los últimos años.

Tabla 4. Modelos multinivel. Disposición a participar en elecciones

	2009				2016			
	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Coef (SE)	LI, LS	Coef (SE)	LI, LS	Coef (SE)	LI, LS	Coef (SE)	LI, LS
Intercepto	49.22* (0.55)	[48.14, 50.31]	51.03* (0.68)	[49.7, 52.36]	47.4* (0.42)	[46.59, 48.22]	50.53* (0.54)	[49.47, 51.58]
Género (Fe- menino)	0.33 (0.46)	[-0.56, 1.22]	-0.25 (0.45)	[-1.12, 0.62]	1.11* (0.4)	[0.33, 1.89]	0.74 (0.38)	[-0.01, 1.49]
NSE quin- til 2	-0.39 (0.63)	[-1.64, 0.85]	-1.09 (0.64)	[-2.34, 0.16]	1.09* (0.51)	[0.09, 2.08]	-0.3 (0.57)	[-1.41, 0.81]
NSE quin- til 3	-0.4 (0.71)	[-1.8, 0.99]	-1.39 (0.65)	[-2.67, -0.11]	2.27* (0.49)	[1.31, 3.22]	0.34 (0.57)	[-0.79, 1.46]
NSE quin- til 4	0.52 (0.77)	[-0.99, 2.02]	-1.16 (0.76)	[-2.64, 0.32]	3.3* (0.55)	[2.23, 4.38]	-0.23 (0.61)	[-1.43, 0.97]
NSE quin- til 5	3.59* (0.73)	[2.16, 5.02]	0.68 (0.89)	[-1.07, 2.43]	5.72* (0.47)	[4.8, 6.64]	0.62 (0.61)	[-0.58, 1.82]
Interés de padres en política			2.02* (0.18)	[1.67, 2.36]			2.14* (0.21)	[1.73, 2.55]
Discutir so- bre política fuera de la escuela			2.28* (0.22)	[1.84, 2.71]			1.67* (0.22)	[1.24, 2.09]
Dependen- cia (privado)			-0.4 (0.58)	[-1.53, 0.74]			-0.74 (0.4)	[-1.52, 0.04]
NSE Es- cuela			0.18 (0.4)	[-0.6, 0.97]			1.21* (0.27)	[0.68, 1.73]
Discusión abierta en clases			0.94* (0.27)	[0.42, 1.47]			0.85* (0.21)	[0.44, 1.26]
ICC	0.06		0.05		0.06		0.06	

Nota: *p<.05; LI y LS son los límites inferior y superior de los intervalos al 95% de confianza. Fuente: elaboración propia

Por otra parte, la Tabla 5, que presenta los resultados para la variable confianza en partidos políticos, muestra que tanto el género como el nivel socioeconómico presentan resultados estadísticamente significativos en ambos años, en todos los casos a favor de las mujeres y de los niveles socioeconómicos más altos, tanto en 2009 como en 2016. Es interesante notar que, en términos socioeconómicos, las mayores brechas están entre el quintil 1 (de referencia) y los quintiles que representan las clases medias y medias-altas (quintiles 3 y 4) y que además la magnitud del cambio tiende a disminuir entre 2009 y 2016, aunque de manera no significativa.

Tabla 5. Modelo multinivel. Confianza en partidos políticos

	2009				2016			
	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Coef (se)	LI, LS	Coef (se)	LI, LS	Coef (se)	LI, LS	Coef (se)	LI, LS
Intercepto	2.66* (0.04)	[2.58, 2.73]	2.64* (0.04)	[2.57, 2.72]	2.71* (0.04)	[2.63, 2.79]	2.68* (0.05)	[2.58, 2.78]
Género (Femenino)	0.07* (0.02)	[0.02, 0.11]	0.09* (0.02)	[0.05, 0.13]	0.12* (0.03)	[0.06, 0.18]	0.14* (0.03)	[0.08, 0.2]
NSE quintil 2	0.07 (0.04)	[-0.02, 0.15]	0.08 (0.04)	[0, 0.16]	0.08 (0.05)	[-0.02, 0.18]	0.04 (0.05)	[-0.06, 0.14]
NSE quintil 3	0.16* (0.05)	[0.07, 0.26]	0.18* (0.05)	[0.08, 0.28]	0.1* (0.05)	[0.01, 0.19]	0.07 (0.05)	[-0.02, 0.17]
NSE quintil 4	0.21* (0.05)	[0.12, 0.31]	0.25* (0.05)	[0.14, 0.35]	0.16* (0.05)	[0.07, 0.25]	0.12* (0.06)	[0.01, 0.23]
NSE quintil 5	0.13* (0.05)	[0.02, 0.24]	0.19* (0.06)	[0.07, 0.31]	0.11 (0.06)	[0, 0.23]	0.06 (0.07)	[-0.09, 0.21]
Interés de padres en política			-0.14* (0.02)	[-0.17, -0.11]			-0.16* (0.02)	[-0.19, -0.12]
Discutir sobre política fuera de la escuela			-0.07* (0.01)	[-0.1, -0.04]			0 (0.02)	[-0.03, 0.03]
Dependencia (privado)			-0.03* (0.05)	[-0.13, 0.07]			0.1* (0.04)	[0.02, 0.18]
NSE Escuela			0.05 (0.02)	[0, 0.1]			0.07* (0.03)	[0.02, 0.12]
Discusión abierta en clases			-0.03 (0.02)	[-0.07, 0.01]			-0.02 (0.02)	[-0.06, 0.01]
ICC	0.03		0.03		0.03		0.03	

Nota: *p<.05; LI y LS son los límites inferior y superior de los intervalos al 95% de confianza. Fuente: elaboración propia

La Tabla 6 presenta los resultados sobre disposición a compartir con otros de distinta clase. Como se puede ver, al controlar por otros factores, el peso del nivel socioeconómico es prácticamente la única variable significativa del modelo, tanto para el 2009 como el 2016. En ambos años, la diferencia entre el quintil y los grupos medios (quintiles 2, 3 y 4) es significativa y positiva, lo que podría estar confirmando la idea anterior: que la disposición a compartir con otros sería menor en los quintiles extremos (quintil 1 y 5). Adicionalmente, y de manera diferenciada con otras variables, ninguna de las variables de la escuela ni de los procesos escolares es estadísticamente significativa para los años analizados.

Tabla 6. Modelo multinivel. Disposición a compartir con otros de distinta clase

	2009				2016			
	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Coef (se)	LI, LS	Coef (se)	LI, LS	Coef (se)	LI, LS	Coef (se)	LI, LS
Intercepto	0.93* (0.01)	[0.91, 0.95]	0.93* (0.01)	[0.91, 0.95]	0.86* (0.01)	[0.84, 0.89]	0.86* (0.02)	[0.83, 0.89]
Género (Femenino)	0.01 (0.01)	[-0.01, 0.02]	0.01 (0.01)	[-0.01, 0.02]	0.03* (0.01)	[0.01, 0.05]	0.03* (0.01)	[0.01, 0.05]
NSE quintil 2	0.02 (0.01)	[0, 0.04]	0.03* (0.01)	[0.01, 0.05]	0.03 (0.02)	[0,00 0.06]	0.03 (0.02)	[0,00 0.06]
NSE quintil 3	0.03 (0.01)	[0, 0.05]	0.03* (0.01)	[0.01, 0.05]	0.04* (0.01)	[0.01, 0.07]	0.04* (0.02)	[0.01, 0.07]
NSE quintil 4	0.02 (0.01)	[0, 0.04]	0.03 (0.01)	[0, 0.05]	0.04* (0.01)	[0.01, 0.07]	0.04* (0.02)	[0.01, 0.07]
NSE quintil 5	0.01 (0.01)	[-0.02, 0.03]	0.01 (0.02)	[-0.02, 0.05]	0.04 (0.02)	[0.00, 0.07]	0.03 (0.02)	[-0.01, 0.07]
Interés de padres en política			-0.01 (0.00)	[-0.01, 0.00]			0.00 (0.00)	[-0.01, 0.01]
Discutir sobre política fuera de la escuela			0.01 (0.00)	[0,00 0.01]			0.00 (0.01)	[-0.01, 0.01]
Dependencia (privado)			-0.01 (0.01)	[-0.03, 0.01]			0.01 (0.01)	[-0.02, 0.03]
NSE Escuela			0.00 (0.01)	[-0.01, 0.01]			0.00 (0.01)	[-0.02, 0.01]
Discusión abierta en clases				[-0.01, 0.01] 0 (0)			0.01 (0.00)	[0,00 0.02]
ICC	0.01		0.01		0.02		0.02	

Nota: *p<.05; LI y LS son los límites inferior y superior de los intervalos al 95% de confianza. Fuente: elaboración propia.

Finalmente, la Tabla 7 despliega los resultados de las estimaciones de conocimiento cívico⁷, pudiéndose destacar tres aspectos. En primer lugar, se puede observar que en ambos años el género es estadísticamente significativo en todos los modelos, siempre a favor de las mujeres. Además, hay un crecimiento de este efecto en el tiempo, ya que mientras en 2009 las mujeres tenían una ventaja promedio de 0.09 desviaciones estándar respecto a los hombres (controlando por todas las variables del modelo), en 2016, dicha diferencia era de 0.22 desviaciones estándar. Esto podría estar indicando una temprana diferencia a favor de las mujeres respecto de la capacidad de

⁷ Para el modelamiento multinivel el indicador de conocimiento cívico se estandarizó (media 0 y desviación estándar 1), de manera de hacer comparables las magnitudes de los coeficientes entre años.

comprender los problemas sociales y entender el entorno político y cívico que les rodea, tendencia que es internacional y no se limita solo al conocimiento cívico (Sandoval-Hernández, Isac y Miranda, 2018).

En segundo lugar, se observa que el conocimiento cívico está fuertemente influenciado por el nivel socioeconómico promedio de la escuela y del estudiante. Así, cuanto mayor es el nivel socioeconómico mayor es el conocimiento cívico, aunque, a diferencia de lo que ocurre con el género, la diferencia en el nivel socioeconómico se mantiene estable entre 2009 y 2016. Finalmente, otras variables ya relevadas por estudios anteriores, como el rol de la discusión en la sala de clases y fuera de la escuela (Carrasco, Bannerjee, Treviño y Villalobos, 2020; Treviño *et al.*, 2017) son estadísticamente significativos en ambos años, aunque con magnitudes similares.

Tabla 7. Modelo multinivel. Conocimiento Cívico

	2009				2016			
	Modelo 1		Modelo 2		Modelo 3		Modelo 4	
	Coef (se)	LI, LS	Coef (se)	LI, LS	Coef (se)	LI, LS	Coef (se)	LI, LS
Intercepto	-0.66 (0.07)	[-0.8, -0.52]	-0.19 (0.06)	[-0.31, -0.07]	-0.69* (0.05)	[-0.79, -0.59]	-0.32* (0.06)	[-0.44, -0.2]
Género (Femenino)	0.16* (0.04)	[0.07, 0.25]	0.09* (0.03)	[0.03, 0.16]	0.24* (0.04)	[0.18, 0.31]	0.22* (0.03)	[0.15, 0.29]
NSE quintil 2	0.22* (0.07)	[0.08, 0.36]	0.03 (0.06)	[-0.09, 0.14]	0.3* (0.05)	[0.2, 0.4]	0.08 (0.05)	[-0.02, 0.18]
NSE quintil 3	0.46* (0.07)	[0.33, 0.6]	0.16* (0.06)	[0.05, 0.26]	0.5* (0.05)	[0.4, 0.61]	0.19* (0.05)	[0.09, 0.3]
NSE quintil 4	0.76* (0.07)	[0.63, 0.9]	0.28* (0.05)	[0.18, 0.39]	0.78* (0.06)	[0.67, 0.9]	0.28* (0.06)	[0.16, 0.4]
NSE quintil 5	1.18* (0.08)	[1.03, 1.34]	0.43* (0.06)	[0.31, 0.55]	1.12* (0.06)	[1, 1.24]	0.35* (0.07)	[0.21, 0.48]
Interés de padres en política			0.01 (0.01)	[-0.02, 0.04]			0.03 (0.02)	[-0.01, 0.06]
Discutir sobre política fuera de la escuela			0.07* (0.02)	[0.04, 0.1]			0.06* (0.01)	[0.03, 0.09]
Dependencia (privado)			-0.07 (0.06)	[-0.18, 0.05]			0.08 (0.06)	[-0.05, 0.2]
NSE Escuela			0.36* (0.03)	[0.3, 0.42]			0.34* (0.04)	[0.27, 0.42]
Discusión abierta en clases			0.16* (0.03)	[0.1, 0.22]			0.14* (0.03)	[0.09, 0.19]
ICC	0.28		0.28		0.29		0.28	

Nota: *p<.05; LI y LS son los límites inferior y superior de los intervalos al 95% de confianza. Fuente: elaboración propia

7. Conclusiones

Las ILSA han sido motivo de un creciente debate académico durante la última década. Aunque poco tematizado en este debate, el ICCS permite observar cómo un ILSA puede informar sobre un aspecto integral de la educación como es la ciudadanía. Considerando esto, el estudio analizó las tendencias de cambio en cuatro variables relacionadas con la educación cívica y ciudadana en Chile (conocimiento cívico, disposición a votar en las elecciones, confianza en los partidos políticos y disposición a compartir con otros de distinta clase social) entre 2009 y 2016, considerando los cambios en programas y políticas buscaron transformar algunos de estos elementos durante este tiempo.

Aunque los resultados generales mostraron pocos cambios entre los jóvenes al comparar entre años, el análisis por nivel socioeconómico y el análisis multinivel evidenciaron que algunos cambios (p.ej., el cambio en la disposición a compartir con otros o en la disposición a votar en las elecciones) se producían de manera diferenciada entre grupos sociales; o con cambios relevantes en la magnitud de otras variables, como en el caso del género; aunque otras variables (p.e., la discusión en la sala de clases) se mostraron estables entre 2009 y 2016, entregando un panorama heterogéneo de las tendencias en el tiempo.

32

Aunque particularizados en un país, este conjunto de resultados permite re-discutir la relación entre las ILSA y las políticas educativas, pudiendo destacarse tres aspectos. En primer lugar, es interesante observar que, aún en contextos de implementación de políticas y programas enfocados a promover transformaciones, los resultados del ICCS entre 2009 y 2016 muestran gran estabilidad. Esto sugiere que los resultados de ciudadanía no parecen haber “permeado” las discusiones sobre políticas -como si ocurriría, por ejemplo, en TIMSS o PISA (Parcerisa y Verger, 2019)-, ya que no se evidencian cambios que hayan permitido formar estudiantes con mayores niveles de conocimiento cívico, más dispuestos a compartir con otros y a participar de manera más activa en los procesos de democracia participativa. Evidentemente, esta estabilidad puede deberse a múltiples factores -como la baja capacidad transformativa de las políticas, el peso de los factores culturales o la potencia del mercado educativo del país, entre otras- pero es interesante mostrar la baja capacidad de cambio que ha mostrado el sistema, aún en posesión de evidencia de resultados poco alentadores para la formación ciudadana de los estudiantes.

En segundo término, esta estabilidad contrasta con los cambios evidenciados cuando se desagrega por nivel socioeconómico. Esto permite discutir ya no tanto la existencia misma de la información, sino su proceso de presentación y difusión. El énfasis entregado a los promedios y el pro-

cesamiento mediático que se hace a los datos de las ILSA (Pizmony-Levy y Torney-Purta, 2018) reduce la riqueza de los datos, no permitiendo entender los alcances de las políticas ni tampoco entendiendo cómo estas políticas afectan de manera diferenciada a los estudiantes según las distintas posiciones en la estructura social que estos ocupan. Asimismo, estos análisis generales tampoco permiten ver efectos que aumentan la desigualdad, como el aumento en las brechas a la disposición a votar evidenciadas en Chile entre 2009 y 2016. Por este motivo, se requieren análisis más sofisticados que den cuenta del comportamiento de las relaciones a lo largo de la distribución de la variable de interés.

Finalmente, y algo que se puede colegir a partir de la interpretación de los resultados de los modelos multinivel, es interesante pensar en cómo los datos de las ILSA pueden orientar no sólo políticas dentro de la estructura del sistema educativo, sino también iluminar otros aspectos de las sociedades. Esto refiere, en último término, a una discusión sobre el uso que se hace de estas evaluaciones. Así, creemos que es posible pensar y promover que las ILSA no sean usados como “barómetros de calidad” de los sistemas educativos (Carnoy *et al.*, 2013), que generen una competencia entre sistemas y que orienten cambios radicales siguiendo tendencias extranjeras (Parcerisa *et al.*, 2020; Simola, 2013) sino utilizar los datos para orientar cambios en políticas sociales, educativas, económicas que permitan mostrar espacios de mejora de los estudiantes, sin “culpar” inmediatamente a los sistemas educativos por estos resultados. Así, por ejemplo, en el caso chileno, los resultados analizados sobre voto, disposición a compartir con otros y confianza en los partidos políticos parecen ser un reflejo de problemas estructurales de la sociedad, que no se limitan necesariamente a cambios particulares del sistema educativo, sino que requieren transformaciones en los cimientos (constitucionales) del país.

8. Agradecimientos

El artículo se desarrolla en el marco del proyecto Fondecyt Regular N° 1180667. Adicionalmente, los autores agradecen el apoyo brindado por el proyecto CONICYT PIA CIE160007. Ambos proyectos son financiados por Agencia Nacional de Investigación y Desarrollo (ANID) de Chile. Finalmente, los autores agradecen los comentarios y sugerencias de Diego Carrasco y Lluís Parcerisa, así como los comentarios y sugerencias de los/as revisores/as anónimos.

Referencias

- Alarcón, C. (2020). An Improbable Identification? The Netherlands as a “Reference Society” within the Chilean Educational Policy-Making Debate (2014–2015). *European Education*, 52(1), 48–67. <https://doi.org/10.1080/10564934.2019.1694418>
- Alvarado, E., Rivera, P., y Morales, R. (2019). Radicalizing democracy from social movements. The compared cases of Podemos in Spain and Frente Amplio in Chile. *Izquierdas*, 48, 87–105. <https://doi.org/10.4067/S0718-50492019000400087>
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215–228. <https://doi.org/10.1080/0268093022000043065>
- Bieber, T. (2016). *Soft governance, international organizations and education policy convergence: comparing PISA and the Bologna and Copenhagen processes*. Amsterdam: Springer.
- Canales, M., Bellei, C., y Orellana, V. (2016). ¿Por qué elegir una escuela privada subvencionada? Sectores medios emergentes y elección de escuela en un sistema de mercado. *Estudios Pedagógicos*, 42(3), 89–109. <https://doi.org/10.4067/s0718-07052016000400005>
- Cariola, L., Covacevich, C., Gubler, J., Lagos, E., y Ortiz, M. (2011). Chilean participation in IEA Studies. In C. Papanastasiou, T. Plomp, y E. Papanastasiou (Eds.), *IEA 1958–2008: 50 Years of Experiences and Memories*. Nicosia: Cultural Center of the Kykkos Monastery (pp. 373-387).
- Carlin, R. E. (2006). The decline of citizen participation in electoral politics in post-authoritarian Chile. *Democratization*, 13(4), 632–651. <https://doi.org/10.1080/13510340600791921>
- Carnoy, M., Khavenson, T., y Ivanova, A. (2013). Using TIMSS and PISA results to inform educational policy: a study of Russia and its neighbours. *Compare*, 45(2), 248–271. <https://doi.org/10.1080/03057925.2013.855002>
- Carrasco, D., Banerjee, R., Treviño, E., y Villalobos, C. (2020). Civic knowledge and open classroom discussion: explaining tolerance of corruption among 8th-grade students in Latin America. *Educational Psychology*, 40(2), 186–206. <https://doi.org/10.1080/01443410.2019.1699907>
- Corvalán, A., y Cox, P. (2013). Class-Biased Electoral Participation: The Youth Vote in Chile. *Latin American Politics and Society*, 55(03), 47–68. <https://doi.org/10.1111/j.1548-2456.2013.00202.x>
- Cox, C. (en prensa). IEA Civic Education Studies in Latin America: paths of influence and critique in policy and research. In J Torney-Purta y B. Malak-Minkiewicz (Eds.), *Civic and Citizenship Education Studies of IEA: Influences on Practice, Policy and Research*. Amsterdam: IEA.
- Cox, C., y García, C. (2017). Evolution of citizenship education in Chile: Recent curricula compared. In García-Cabrero, B., Sandoval-Hernández, A., Treviño, E., Diazgrandos-Ferrand, S. y Perez, G. (Eds.). *Civics and Citizenship. Theoretical Models and Experiences in Latin America*. Rotterdam/Boston/Taipei: Sense Publisher (pp. 85-103).
- Cox, C., y Meckes, L. (2016). International large-scale assessment studies and educational policy-making in Chile: contexts and dimensions of influence. *Research Papers in Education*, 31(5), 502–515. <https://doi.org/10.1080/02671522.2016.1225349>
- Gorur, R. (2017). Towards productive critique of large-scale comparisons in education. *Critical Studies in Education*, 58(3), 341–355. <https://doi.org/10.1080/17508487.2017.1327876>

- Grek, S. (2009). Governing by numbers: The PISA “effect” in Europe. *Journal of Education Policy*, 24(1), 23–37. <https://doi.org/10.1080/02680930802412669>
- Heyneman, S., y Lee, B. (2014). The impact of international studies of academic achievement on policy and research. In Rutkowski, L., von Davier, M. y Rutkowski, R. (Eds.). *Handbook of International Large-Scale Assessment: Background, technical issues and methods of data analysis*. New York: CRC Press (pp. 37–72).
- Hox, J., Moerbeek, M., y Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.
- INJUV (2009). *Sexta encuesta nacional de juventud*. Santiago, Chile.
- Isac, M. M., Maslowski, R., Creemers, B., y van der Werf, G. (2014). The contribution of schooling to secondary-school students’ citizenship outcomes across countries. *School Effectiveness and School Improvement*, 25(1), 29–63. <https://doi.org/10.1080/09243453.2012.751035>
- Klemencic, E. (2010). The impact of international achievement studies on national education policymaking: The case of Slovenia - how many watches do we need? *International Perspectives on Education and Society*, 13, 239–266. [https://doi.org/10.1108/S1479-3679\(2010\)0000013012](https://doi.org/10.1108/S1479-3679(2010)0000013012)
- Knowles, R. T., Torney-Purta, J., y Barber, C. (2018). Enhancing citizenship learning with international comparative research: Analyses of IEA civic education datasets. *Citizenship Teaching and Learning*, 13(1), 7–30. https://doi.org/10.1386/cti.13.1.7_1
- MINEDUC. (2013). *Bases Curriculares 2013. Historia, Geografía y Ciencias Sociales*. Santiago, Chile.
- Niemann, D., y Martens, K. (2018). Soft governance by hard fact? The OECD as a knowledge broker in education policy. *Global Social Policy*, 18(3), 267–283. <https://doi.org/10.1177/1468018118794076>
- Ozga, J. (2008). Governing Knowledge: Research Steering and Research Quality. *European Educational Research Journal*, 7(3), 261–272. <https://doi.org/10.2304/eej.2008.7.3.261>
- Parcerisa, L., y Falabella, A. (2017). La consolidación del estado evaluador a través de políticas de rendición de cuentas: trayectoria, producción y tensiones en el sistema educativo chileno. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, 25, 1–24. <https://doi.org/10.14507/epaa.25.3177>
- Parcerisa, L., Fontdevila, C., y Verger, A. (2020). Understanding the PISA influence on national education policies: a focus on policy transfer mechanisms. In S. Jörnitz y A. Wilmers (Eds.), *International perspectives on school settings, education policy and digital strategies. A transatlantic discourse in education research*. Leverkusen.
- Parcerisa, L., y Verger, A. (2019). PISA projections in Chile: The selective use of league leaders in the enactment of recent education reforms. In Wladow, F. y Steiner-Khamsi, G. (Eds.). *Understanding PISA's attractiveness: Critical analyses in comparative policy studies*. Bloomsbury Academic (pp. 25–48).
- Pelgrum, H., Blahova, V., Dukynaitė R., Pavešić B. J., y Kangro, A. (2011). IEA experiences from Latvia, Lithuania, Slovak Republic and Slovenia. In C. Papanastasiou, T. Plomp, y E. Papanastasiou (Eds.), *IEA 1958–2008: 50 Years of Experiences and Memories*. Nicosia: Cultural Center of the Kykkos Monastery (pp. 447–468).
- Pizmony-Levy, O., y Torney-Purta, J. (2018). How journalists and researchers communicate results of international large-scale assessments. *Cadmo*, 26(1), 51–65. <https://doi.org/10.3280/CAD2018-001007>

- Rioux, C., y Little, T. D. (2020). Underused Methods in Developmental Science to Inform Policy and Practice. *Child Development Perspectives*, 14(2), 97–103. <https://doi.org/10.1111/cdep.12364>
- Sandoval-Hernández, A., Isac, M., y Miranda, D. (Eds.) (2018). *Teaching tolerance in a globalized world*. Amsterdam: Springer-IEA.
- Schulz, W., Fraillon, J., Ainley, J., Losito, B., y Kerr, D. (2008). *International Civic and Citizenship Education Study: Assessment Framework*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).
- Sellar, S., Lingard, B., Rutkowski, D., y Takayama, K. (2018). Student preparation for large-scale assessments: a comparative analysis. In Madoxx, B. (Ed.). *International Large-Scale Assessments in Education. Insider Research Perspective*. London, UK: Bloomsbury. (pp. 137–155).
- Simola, H. (2013). El milagro finlandés de PISA: observaciones históricas y sociológicas sobre la enseñanza y la formación del profesorado. *Revista de Currículum y Formación Del Profesorado*, 17(2), 153–169.
- Torney-Purta, J., y Amadeo, J.-A. (2013). International Large-Scale Assessments: Challenges in Reporting and Potentials for Secondary Analysis. *Research in Comparative and International Education*, 8(3), 248–258. <https://doi.org/10.2304/rcie.2013.8.3.248>
- Treviño, E., Béjares, C., Villalobos, C., y Naranjo, E. (2017). Influence of teachers and schools on students' civic outcomes in Latin America. *The Journal of Educational Research*, 110(6), 604–618. <https://doi.org/10.1080/00220671.2016.1164114>
- Treviño, E., Villalobos, C., Béjares, C., y Naranjo, E. (2019). Forms of youth political participation and educational system: The role of the school for 8th grade students in Chile. *Young*, 27(3), 279–303. <https://doi.org/10.1177/1103308818787691>
- Valenzuela, J. P., Bellei, C., y De los Ríos, D. (2014). Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of Education Policy*, 29(2), 217–241. <https://doi.org/10.1080/02680939.2013.806995>
- Valenzuela, J. P., Villarroel, G., y Villalobos, C. (2013). Ley de Subvención Escolar Preferencial (SEP): algunos resultados preliminares de su implementación. *Revista Pensamiento Educativo*, 50(2), 113–131. <https://doi.org/10.7764/PEL.50.2.2013.7>
- Villalobos, C., y Quaresma, M. L. (2015). Sistema escolar chileno: Características y consecuencias de un modelo orientado al mercado. *Convergencia*, 22(69), 63–84. <https://doi.org/10.29101/crcs.v22i69.3634>
- Villalobos, C., y Valenzuela, J. P. (2012). Polarización y cohesión social del sistema escolar chileno. *Revista de Análisis Económico*, 27(2), 145–172. <http://dx.doi.org/10.4067/S0718-88702012000200005>
- Wagemaker, H. (2011). IEA: International studies, impact and transition. In C. Papanastasiou, T. Plomp, y E. Papanastasiou (Eds.), *IEA 1958-2008: 50 years of experiences and memories*. Nicosia: Cultural Center of the Kykkos Monastery (pp. 257–272).
- Wagemaker, H. (2013). International large-scale assessments: From research to policy. In Rutkowski, L., von Davier, M. y Rutkowski, R. (Eds.). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. New York: CRC Press (pp. 11-36).
- Weinstein, J., y Villalobos, C. (2016). Chile: la experiencia de la Subvención Escolar Preferencial. In Souto, M. (Ed.). *Transferencias directas a escuelas: reflexiones sobre prácticas en América Latina*. IPPE: Buenos Aires (pp. 58–100).

Género y etnicidad. Un análisis crítico de las políticas de equidad y los resultados de la prueba PISA en Brasil y México

David G. Miranda¹  ; Patricia Torres¹ 

¹Universidad SEK (USEK), Chile

Resumen. Históricamente, género y etnicidad han sido factores claves para la generación de inequidades socioeconómicas profundas a escala global, una realidad que nos lleva a la necesidad de analizar la contribución de políticas públicas de equidad relacionadas con ambos factores. En este estudio se abordan Brasil y México, los países más poblados de América Latina, se realizó un ejercicio comparativo de políticas públicas relacionadas con género y etnicidad, triangulando sus posibles efectos en resultados PISA, e indicadores socioeconómicos seleccionados desde datos paramétricos censales. Se concluye que, aunque los principios de equidad se declaran formalmente en sus políticas públicas, no se observa evidencia empírica suficiente del impacto de dichas políticas en los niveles de desempeño educativo desagregado por género ni en los niveles de equidad por género y factor étnico.

Palabras clave: género; etnicidad; Brasil; México; políticas públicas; PISA.

Gênero e etnicidade. Uma análise crítica das políticas de equidade e dos resultados da prova PISA no Brasil e no México

Resumo. Historicamente, gênero e etnicidade têm sido fatores fundamentais para a geração de desigualdades socioeconômicas profundas em escala global, uma realidade que nos leva à necessidade de analisar a contribuição de políticas públicas de equidade relacionadas com os dois fatores. Neste estudo sobre o Brasil e o México, os países mais populosos da América Latina, realizou-se um exercício comparativo de políticas públicas relacionadas com gênero e etnicidade, triangulando seus possíveis efeitos nos resultados PISA, e indicadores socioeconômicos selecionados a partir de testes paramétricos. Conclui-se que, embora os princípios de equidade se declarem formalmente em suas políticas públicas, não se observa evidência empírica suficiente do impacto destas políticas nos níveis de desempenho escolar desagregado por gênero, nem nos níveis de equidade por gênero e fator étnico.

Keywords: gênero; etnicidade; Brasil; México; políticas públicas; PISA.

Gender and Ethnicity. A critical analysis of equity policies and the results of the PISA test in Brazil and Mexico

Abstract. Historically, gender and ethnicity have been key factors for the generation of deep socioeconomic inequities on a global scale, a reality that leads us to analyze its contribution in terms of public equity policies related to both factors, in this case, about Brazil and Mexico, the most populous countries in Latin America. A comparative exercise of public policies was carried out, triangulating its possible effects into PISA results in each country, and selected socioeconomic indicators from census parametric data. It is concluded that, although the principles of equity are formally declared in their public policies, there is no empirical evidence of the impact of these policies on the levels of educational performance disaggregated by gender or on the equity levels by gender or ethnic factor.

Keywords: gender; ethnicity; Brazil; Mexico; public policies; PISA.

37

1. Introducción

La discriminación a causa del género y la etnicidad han sido factores esenciales para la consolidación de las bases de la desigualdad en el mundo. En el caso de América Latina, está altamente documentado que mujeres y pueblos originarios sufren hasta nuestros días fuertemente la desigualdad en términos de derechos políticos, equidad social y situación económica (Bello

y Rangel, 2002; CEPAL, 2013). Sobre el concepto de etnicidad, ha sido definido por autores como Barth (1976), como un fundamento de razonamiento antropológico que, a su vez agrupa y divide una comunidad de individuos en conjuntos constituidos como unidades discretas, y que generalmente “son el fundamento mismo sobre el cual están construidos los sistemas sociales que la contienen” (1976, p.10), mediante dinámicas de autopropagación, preservación de valores, generación de un campo de interacción y procesos de identificación y/o adscripción a partir de características consideradas por los propios individuos como significativas (Rojas, 1990), y cuyos límites pueden o no estar asociados a concomitantes territoriales. A partir de dichos elementos, el factor étnico puede estar sujeto entonces a la tensión entre uno o más grupos constituidos étnicamente, en relación al estado, entidad que puede delimitar dichos colectivos bajo un prisma ideológico (Guidieri, Pellizi y Stanley, 1988), albergando el potencial de evolucionar hacia un concepto de tipo *clasista* (Díaz, 1985). En el caso del factor género, en América Latina se manifiesta no solamente en el ámbito urbano, sino fuertemente en el sector campesino e indígena, tal como lo señala respecto del caso mexicano Pilar Alberti (1995), quien acusa una falta de protagonismo de mujeres en la caracterización (pese a su activismo) de los movimientos sociales vinculados a la ruralidad. Dado que desde las políticas públicas se formulan estrategias para hacer frente a la desigualdad, resulta indispensable examinar si estamos en presencia de políticas robustas con resultados concretos, o éstos aún no son posibles de visualizar en virtud de la evidencia disponible.

Por su parte, existe un amplio consenso sobre la relevancia del factor educativo como eje de la movilidad social (Yaschine, 2015; Espinoza, Gonzalez y Uribe, 2009; Sanhueza *et al.*, 2015) no solo en la región, sino a escala global. Más aún cuando nos enfrentamos ante una realidad compleja donde las economías del conocimiento, economías creativas, circulares, y los sistemas de innovación y emprendimiento se conciben como los ejes más relevantes de los nuevos modelos de desarrollo (OCDE, 2016; Schleicher, 2019). En dicho contexto cambiante, entidades como la OCDE han adquirido especial relevancia como catalizadoras para el proceso de políticas públicas orientadas al desarrollo económico, generando desde una perspectiva crítica, un proceso de transnacionalización y mercantilización (De Sousa, 2001) que afecta al sistema educativo, y que opera también a través de un sistema de “producción de ausencias” sobre población periférica y sobre aquellos que Boaventura de Sousa denomina *saberes sometidos* (De Sousa, 2013). Sin embargo, es necesario señalar que, a partir de la incorporación de índices y variables específicas, la prueba PISA se ha transformado al mismo tiempo, de forma gradual, en un instrumento que permite contrastar elementos estructurales que subyacen en la estructura social, y que tienen la oportunidad de proyectarse a la realidad a partir de su articulación efectiva de sus resultados con las políticas públicas de cada estado.

Es a partir de dichas constataciones que el presente estudio se orienta a un ejercicio comparativo de políticas públicas en Brasil y México, relacionadas con género y etnicidad, triangulando sus posibles efectos con los resultados PISA en cada país, y con algunos indicadores socioeconómicos seleccionados desde datos paramétricos censales. Para dar inicio a éste, revisaremos de forma breve las principales políticas en cada país, para plantear un marco metodológico y un análisis de resultados que permitan generar una discusión y conclusiones.

1.1 El contexto brasileño

En Brasil es posible constatar en su entramado jurídico-institucional que la igualdad de oportunidades es uno de los pilares del derecho constitucional. Tal como se observa a continuación:

Constituyen objetivos fundamentales de la **de vuestro artículo** República Federativa de Brasil: I – construir una sociedad libre, justa y solidaria; II – garantizar el desarrollo nacional; III – erradicar la pobreza y la marginación y reducir las desigualdades sociales y regionales; IV – promover el bien de todos, sin prejuicios de origen, raza, sexo, color, edad o cualquier otra forma de discriminación (Constitución de Brasil, 1988, art. 3°).

No obstante, las alarmantes cifras de desigualdad, y el desarrollo histórico de su estructura social con un fuerte clivaje racial nos hace cuestionar la relación entre la dimensión política *declarativa* y la realidad de la población. Así, Brasil cuenta con un conjunto de leyes basadas en el principio de *acción afirmativa* (Bergmann, 1996; Vieira, 2003; Cruz, 2005; Moehlecke, 2005) las cuales reconocen el derecho a una diferencia en el tratamiento legal para grupos históricamente discriminados, el cual se observa en su constitución, mediante un reconocimiento en sus garantías de la *no-discriminación* racial (art 4°.8; art. 43). En dicho marco se implementa la Ley de Directivas y Bases de Educación Nacional (Ley 9.394 / 1996), cuyo objetivo es estructurar el sistema educativo incorporando dicho principio (Vieira, 2003), en su artículo 78:

[...]el sistema educativo (...), desarrollará programas integrados de docencia e investigación, ofreciendo educación escolar bilingüe e intercultural a los pueblos indígenas, con las siguientes metas:

- I. brindar a los indígenas, sus comunidades y pueblos, la recuperación de su memoria histórica; la reafirmación de sus identidades étnicas; la apreciación de sus lenguas y ciencias;
- II. garantizar a los indígenas, sus comunidades y pueblos, acceso a información, conocimiento, aspectos técnicos y científicos de la sociedad nacional y otras sociedades indígenas y no indígenas.

En términos de legislación de equidad e inclusión educativa y política destacan las siguientes iniciativas:

- Ley no. 10.639 / 2003, enseñanza obligatoria de historia y cultura afrobrasileña y africana.
- Ley. 11.645 / 2008, sobre temática indígena en el currículo.
- Ley 12.711 / 2012 reserva del 50% de la matrícula para estudiantes de escuelas públicas y universidades en una proporción de color o “raza” igual a la población local¹ (derogado por Jair Bolsonaro)

Otros ejemplos de estas acciones se observan a partir de 1997, con la promulgación de la Ley 9.504 que en su artículo 10.2 establece cuotas para mujeres candidaturas políticas. (también derogado por Bolsonaro). Posteriormente, en 2001, el Ministerio de Desarrollo Agrario y el Instituto Nacional de Colonización y Reforma Agraria fueron las primeras entidades oficiales en institucionalizar un *programa de acción afirmativa*, adoptando cuotas del 20% en el cumplimiento de las funciones de alta gerencia y asesoría, y para promover la igualdad en acuerdos de cooperación técnica. Por su parte, el Ministerio de Justicia optó por capacitar al 45% del personal femenino, de color o que padeciera algún grado de discapacidad. Además de ello, desde 2002 se estableció que todas las empresas que brindan servicios a estos organismos deben reservar el 20% de sus vacantes para funcionarios afrodescendientes.

Así, es posible observar diversas iniciativas orientadas a compensar desigualdades sociales, económicas y culturales bajo un enfoque de derecho. Todo ello bajo preceptos constitucionales, que establecen la educación como primer derecho social (art. 6), así como garantiza entre sus derechos sociales la “prohibición de diferencias salariales, ejercicio de funciones y criterios de admisión por sexo, edad, color o estado civil” (art.7). No obstante, las fuertes diferencias se observan hasta nuestros días.

1.2 El contexto mexicano

La constitución mexicana, pionera en el establecimiento de derechos sociales, (dos años antes de la Constitución de Weimar de 1919), contiene un marcado espíritu liberal, e incluye garantías de protección social, formando parte del denominado *constitucionalismo social*. Sin embargo, no ha estado exenta de críticas, habiendo sido tildada de formar una sociedad “homogénea

¹ Ya en mayo de 1996 el Programa Nacional de Derechos Humanos recomendaba mejorar el acceso de población negra a las universidades.

y monocultural” incorporando el mundo indígena al denominado mundo mestizo (Singer, 2014), más aún cuando para interpretar los derechos indígenas, se establecen normas particulares de cada estado, generando divergencias significativas. Así, la constitución mexicana establece en su artículo 1° el principio de *no discriminación* de la siguiente forma:

Queda prohibida toda discriminación motivada por origen étnico o nacional, el género, la edad, las discapacidades, la condición social, las condiciones de salud, la religión, las opiniones, las preferencias sexuales, el estado civil o cualquier otra que atente contra la dignidad humana y tenga por objeto anular o menoscabar los derechos y libertades de las personas. (1917, art. 1°).

Por otra parte, dicha constitución fue la primera en consagrar mecanismos claros respecto a la *educación popular*. Además, entre dichos principios para una nueva educación aseguran que ésta debe ser *democrática*, no solo estructuralmente, sino también desde una perspectiva política y social, garantizando equidad de acceso en todos sus niveles, con un fuerte sistema de educación pública.

En las décadas recientes, destacan como políticas orientadoras los denominados Planes Nacionales de Desarrollo (PND) 2007-2012 y 2013-2018, han tenido como propósito ejecutar acciones para el desarrollo, y avanzar en la disminución de brechas estructurales. En relación a educación, entre sus objetivos destacan “Reducir las desigualdades regionales, de género y entre grupos sociales en las oportunidades educativas” (PND, 2007), e “Incorporar la perspectiva de igualdad de género en las políticas públicas, programas, proyectos e instrumentos compensatorios como acciones afirmativas” (PND, 2013). Por otra parte, en lo que respecta a género y etnicidad, destacan entre su legislación y programas, las siguientes:

- Ley del Instituto Nacional de las Mujeres (2001).
- Ley de derechos y cultura indígena (2002).
- Ley para Prevenir y Eliminar la Discriminación (2003).
- Ley general de derechos lingüísticos de los pueblos indígenas (2003).
- Ley General para la Igualdad entre Mujeres y Hombres (2006).
- Ley de la Comisión Nacional de los Derechos Humanos (2006).
- Programa Transversalidad de Género (2008).
- Programa de Corresponsabilidad Social para Beneficio de las Mujeres (2008).

- Gasto Etiquetado para las Mujeres y la Igualdad de Género (2008)
- Programa Nacional de Igualdad entre Hombres y Mujeres (2007-2012).
- Programa de Capacitación y Especialización en Género, Derechos Humanos de las Mujeres y Política Pública (2014).

1.3 A modo de contraste

En el aspecto normativo, tanto Brasil como México declaran igualdad formal, siendo signatarios de diversos acuerdos internacionales sobre pueblos originarios y protección a las mujeres (p.e. Convenio 169 OIT, Convención sobre la eliminación de todas las formas de discriminación contra la mujer, CEDAW), elevando en su discurso político los valores de la dignidad y la erradicación de toda forma de discriminación. Sin embargo, la brecha entre la realidad social y el orden normativo (Durham, 2003; Maggie y Fry, 2002; Vidal, 2014; Miranda, Freire y Jervis, 2019) resulta preocupante luego de décadas de procesos políticos que incorporan entre sus ejes la equidad social. Dicha inconsistencia podría explicarse aparentemente en una conducta general de conservadurismo que es convergente con el argumento de la denominada *colonialidad* de sus estructuras sociales (Castro-Gómez, 1998; Quijano, 2000, Dussel, 2005; Mignolo *et al.*, 2006). A partir de la interpretación crítica de estos y otros autores, dicho rasgo estaría presente históricamente en la región a raíz de sus procesos intensivos de hibridación de culturas, religiones y razas, donde buena parte de su población se aferra a un *habitus* poscolonial (Mignolo *et al.*, 2006) entendida como aquello que permite la reproducción social de las condiciones materiales de la sociedad, y que es caracterizado por autores como Lugones como un “sistema moderno-colonial de género” (Lugones, 2008), basado en factores de dominancia social (Montes-Berges y Del Prado, 2014).

Por otra parte, en los años recientes, las evaluaciones educativas internacionales han sido objeto discusión a la hora de modelar las políticas públicas de equidad. En Brasil y México, a dos décadas de incorporación a PISA, resulta de alto interés realizar un análisis de resultados desagregados por sexo, recopilados durante las siete ediciones de PISA (2000 a 2018) en tres áreas de conocimiento, y realizar una mirada crítica a las políticas públicas sobre género y etnicidad, triangulando marco jurídico-político, resultados educativos estandarizados, y brechas socioeconómicas sobre género y etnicidad.

2. Marco metodológico

El enfoque metodológico del estudio se configura a partir de dos elementos, el primero es una perspectiva general sobre políticas públicas específicas sobre género y etnicidad, con un marcado enfoque social (Roth, 2009) que permita realizar una aproximación sobre sus efectos (Feinstein, 2012) en dos países que tienen en su proceso de políticas públicas un énfasis de planificación que ha sido acompañado por diferentes organizaciones internacionales (Feinstein, 2012) en su proceso de institucionalización y evaluación. Así, en el apartado de contexto se enuncian sus principales políticas en el ámbito, durante el período. El segundo elemento de análisis se configura como un contraste de evidencia empírica en relación con los datos de PISA de 2000 a 2018 en lectura, matemáticas y ciencias, mediante la comparación estadística de medias entre países, y por sexo. Se realizaron pruebas ANOVA, prueba T para muestras independientes, y coeficiente de correlación de Chi cuadrado. Las fuentes utilizadas para el análisis fueron diversas: Organización para la Cooperación y el Desarrollo Económico (OCDE), National Center for Education Statistics (NCES), Instituto Nacional de Estudios e Investigaciones Anísio Teixeira (INEP, Brasil), Instituto Nacional para la Evaluación de la Educación (INEE, México), y la División de Población de CEPAL (CELADE). Una vez obtenidos los datos desde las fuentes primarias, las herramientas utilizadas para el análisis estadístico fueron: SPSS², IDB Analyzer, NCES International Data Explorer (IDE-PISA), entre otros.

43

Para el estudio de datos socioeconómicos sobre etnicidad y sexo, de acuerdo con lo observado por Abramo y Valenzuela (2006), se consideró pertinente para evaluar las brechas de equidad, analizar factores específicos: tasa de participación económica, ocupación/desempleo, en este caso por sexo y factor étnico, analizando datos censales, a partir de las categorías y forma de medición de cada país.

Asimismo, resulta pertinente señalar que inicialmente este estudio proyectaba la recolección y análisis de datos PISA sobre las categorías de género y etnicidad. Para ello se solicitó información específica a los organismos responsables de aplicar la evaluación en ambos países. El Instituto Nacional de Evaluación y Educación de México (INEE) informó que no pudo encontrar en ningún informe datos por etnia de cada edición ni área de conocimiento. Por su parte, el Instituto Nacional de Estudios e Investigación Educativa

² Asimismo, cabe señalar que para el procesamiento de datos PISA, se utilizó la metodología descrita en sus manuales, aplicando pesos replicados, y macros, para evitar el sesgo analítico producto del factor de ponderación poblacional en el cálculo de desviación estándar, error estándar, y análisis de varianzas.

Anísio Teixeira (INEP-Brasil), organismo responsable de la aplicación de PISA, informó que, según la información proporcionada al Instituto, los datos por origen étnico solo se desglosan para Estados Unidos.

De esta manera, los datos recopilados y analizados sobre resultados PISA sólo se refieren a las diferencias de medias entre países, y su respectiva segmentación por sexo. De acuerdo al marco metodológico descrito, y a los antecedentes presentados, se pretende analizar los resultados obtenidos en función de las siguientes preguntas de investigación:

- a) ¿En qué medida las políticas públicas sobre género y etnicidad en Brasil y México se han visto reflejadas en los resultados PISA?
- b) ¿Existen efectos visibles de las políticas públicas sobre género y etnicidad, en los indicadores analizados sobre las brechas de equidad en la población entre 2000 y 2010, de acuerdo a los últimos datos censales?

3. Análisis y resultados

44

La muestra para el estudio de resultados PISA está compuesta por los siguientes grupos. Si bien se observan diferencias de magnitud, el número de casos en cada país permite realizar la comparación con niveles adecuados de confianza (95%).

Tabla 1. Datos de escuelas y alumnos en pruebas PISA, Brasil y México

PISA /año	Brasil		México	
	Escuelas	Estudiantes	Escuelas	Estudiantes
2018	598	10.691	297	7.299
2015	841	23.141	275	7.568
2012	767	18.598	1.471	33.806
2009	950	20.127	1.535	38.250
2006	625	9.295	1.140	33.706
2003	229	4.452	1.124	29.983
2000	250	4.893	183	5.276

Fuente: PISA Data explorer, elaboración propia.

Otro factor a considerar es que tanto en México como en Brasil se realizaron cambios en la definición de edad de los estudiantes. En el caso de México, se realiza en estudiantes de 15 años, mientras en Brasil, con 14. Dicho factor etario podría explicar una buena parte de las brechas en resultados entre ambos países (ver tabla 2). Sin embargo, para alcanzar una visión más acabada sobre dicho fenómeno, habría que ahondar mediante un estudio específico de tipo factorial (Araújo, 2014).

3.1 Resultados Totales por país

Como es posible observar (tabla 2), los resultados para lectura muestran una clara tendencia de rendimiento en tres niveles (Brasil: 395-410 México: 400-425, OCDE: 490-505) que se mantiene a lo largo todas las ediciones de PISA. Se observa una diferencia donde Brasil se encuentra por debajo de lo obtenido por México y ambos países se encuentran muy por debajo de la media OCDE. Dicha estructura de resultados se observa consistentemente en las tres áreas medidas³.

Tabla 2. Puntajes promedio para lectura PISA. 2000-2015 (Brasil, México, OCDE)

Año	Jurisdicción	Todos los estudiantes	
		Promedio	Error Estándar
2018	Promedio OCDE	492	(1,3)
	México	420	(2,7)
	Brasil	412	(2,1)
2015	Promedio OCDE	494	(1,3)
	México	423	(2,6)
	Brasil	407	(2,8)
2012	Promedio OCDE	502	(1,2)
	México	424	(1,5)
	Brasil	407	(2,0)
2009	Promedio OCDE	500	(1,3)
	México	425	(2,0)
	Brasil	412	(2,7)
2006	Promedio OCDE	495	(0,6)
	México	410	(3,1)
	Brasil	393	(3,7)
2003	Promedio OCDE	496	(1,2)
	México	400	(4,1)
	Brasil	403	(4,6)
2000	Promedio OCDE	506	(0,7)
	México	422	(3,3)
	Brasil	396	(3,1)

Fuente: PISA, elaboración propia.

A continuación, observamos los resultados de la prueba ANOVA de comparación de medias entre ambos países y la media de la OCDE, para cada disciplina (véase tablas 3, 5 y 7). Los resultados confirman la existencia de diferencias significativas en todas las medias comparadas entre cada país y la media de la OCDE (sig. < 0,01) con un 95% de confianza. Dado que no se asume el supuesto de igualdad de varianzas (corroborado en prueba de Levene), hemos corregido el análisis mediante una prueba de Welch.

³ Es necesario señalar que en 2000 y 2003 no se realizó la prueba de ciencia en Brasil ni México, ni en 2000 en matemáticas.

Tabla 3. Prueba ANOVA para medias en lectura, PISA (Brasil, México, OCDE).

Año	Estadístico ^a	gl1	gl2	Sig.
2018	877693,42	2	3042282	0,00**
2015	1086079,04	2	3177682	0,00**
2012	1509647,83	2	3004016	0,00**
2009	1073934,63	2	2764450	0,00**
2006	983151,24	2	2599027	0,00**
2003	988861,37	2	2255428	0,00**
2000	1999496,33	2	2480836	0,00**

a. Asymptotically F distributed. (Welch) ** $p < 0,01$ alfa: 0,05.
Fuente: PISA, elaboración propia.

Tabla 4. Puntajes promedio en matemáticas PISA.

Año	Jurisdicción	Todos los estudiantes	
		Promedio	Error Estándar
2018	Promedio OCDE	486	(1,2)
	México	409	(2,4)
	Brasil	383	(2,0)
2015	Promedio OCDE	486	(1,2)
	México	408	(2,2)
	Brasil	377	(2,9)
2012	Promedio OCDE	495	(1,3)
	México	413	(1,4)
	Brasil	389	(1,9)
2009	Promedio OCDE	498	(1,3)
	México	419	(1,8)
	Brasil	386	(2,4)
2006	Promedio OCDE	492	(1,3)
	México	406	(2,9)
	Brasil	370	(2,9)
2003	Promedio OCDE	499	(0,6)
	México	385	(3,6)
	Brasil	356	(4,8)

Fuente: PISA, elaboración propia.

Tabla 5. Prueba ANOVA para medias en matemáticas, PISA (Brasil, México, OCDE)

Año	Estadístico ^a	gl1	gl2	Sig.
2018	1541567,34	2	3067980	0,00**
2015	1821375,44	2	3193068	0,00**
2012	2023578,08	2	3127445	0,00**
2009	1891237,37	2	2874688	0,00**
2006	1746007,15	2	2451682	0,00**
2003	2271056,99	2	2334324	0,00**

a. Asymptotically F distributed. (Welch) ** $p < 0,01$ alfa: 0,05
Fuente: PISA, elaboración propia.

Tabla 6. Puntajes promedio para ciencias. PISA (Brasil, México, OCDE)

Año	Jurisdicción	Todos los estudiantes	
		Promedio	Error Estándar
2018	Promedio OCDE	494	(1,2)
	México	419	(2,6)
	Brasil	403	(2,1)
2015	Promedio OCDE	497	(1,2)
	México	416	(2,1)
	Brasil	401	(2,3)
2012	Promedio OCDE	506	(2,0)
	México	415	(1,3)
	Brasil	402	(2,1)
2009	Promedio OCDE	507	(1,3)
	México	416	(1,8)
	Brasil	405	(2,4)
2006	Promedio OCDE	499	(1,3)
	México	410	(2,7)
	Brasil	390	(2,8)

Fuente: PISA, elaboración propia.

Otro punto relevante para analizar es que dichas diferencias significativas se establecen en Brasil y México entre los niveles 1 y 2 establecidos por la OCDE, mientras que el promedio OCDE se encuentra sistemáticamente en el nivel 3. Estos niveles se asocian a grados de complejidad en la comprensión de la realidad, es decir, el instrumento señala que aquellos estudiantes que se encuentran en el nivel 2 pueden establecer análisis en-

tre elementos y comprender fenómenos, pero un nivel 3 existirían mayores elementos analíticos para la comprensión con un grado de elaborado más elevado (OCDE, 2017).

Tabla 7. Prueba ANOVA para medias en ciencias, PISA (Brasil, México, OCDE).

Año	Estadístico ^a	gl1	gl2	Sig.
2018	1273156,93	2	3119971	0,00**
2015	1548809,81	2	3318286	0,00**
2012	2097748,74	2	3155131	0,00**
2009	1699772,86	2	2906790	0,00**
2006	1547950,72	2	2564435	0,00**

a. Asymptotically F distributed. (Welch) **p<0,01 alfa: 0,05
Fuente: PISA, elaboración propia.

3.2 Resultados PISA desagregados por sexo

Los resultados muestran diferencias significativas en las disciplinas, por sexo, en todos los casos, con pequeñas oscilaciones. En definitiva, confirma que, tal como afirman otros estudios (De la Rica y González, 2013), las mujeres obtienen mejores resultados en lenguaje, sin embargo, la brecha en matemáticas y ciencias favorece a los hombres, la cual resultaría determinada por factores culturales asociadas a la productividad, y que termina por jugar un rol preponderante en la proyección de las mujeres en el mundo laboral, a partir de la asociación de rendimientos económicos más altos para actividades intensivas en conocimiento (García-Holgado, Camacho y García-Peñalvo, 2019; Miranda *et al.*, 2019).

Tabla 8. Puntajes promedio para lectura PISA por sexo.

Año	Jurisdicción	Mujer		Hombre	
		Promedio	Error Estándar	Promedio	Error Estándar
2018	Promedio OCDE	504	(1,5)	480	(1,5)
	México	425	(2,9)	415	(3,0)
	Brasil	425	(2,4)	400	(2,4)
2015	Promedio OCDE	506	(1,5)	484	(1,5)
	México	431	(2,9)	416	(2,9)
	Brasil	419	(3,0)	395	(3,1)
2012	Promedio OCDE	519	(1,2)	486	(1,4)
	México	435	(1,5)	411	(1,7)
	Brasil	421	(2,0)	390	(2,3)

Año	Jurisdicción	Mujer		Hombre	
		Promedio	Error Estándar	Promedio	Error Estándar
2009	Promedio OCDE	517	(1,3)	484	(1,5)
	México	438	(2,1)	413	(2,1)
	Brasil	425	(2,8)	397	(2,9)
2006	Promedio OCDE	513	(1,4)	476	(2,1)
	México	427	(2,9)	393	(3,5)
	Brasil	408	(3,6)	375	(4,6)
2003	Promedio OCDE	513	(1,3)	480	(1,4)
	México	410	(4,6)	389	(4,6)
	Brasil	419	(4,1)	384	(5,8)
2000	Promedio OCDE	521	(2,0)	491	(2,4)
	México	431	(3,8)	411	(4,1)
	Brasil	404	(3,4)	387	(3,9)

Fuente: PISA, elaboración propia.

Dado que se trata un ítem clave para la proyección de condiciones de igualdad, se considera como un asunto crítico, lo que ha conducido a la propia OCDE a profundizar en el análisis de la brecha de género (OCDE, 2015; 2016), incorporando preguntas sobre interés y autoconcepto de estudiantes en relación a disciplinas que tienen mayor impacto económico en los actuales modelos productivos (OECD, 2015).

Tabla 9. Puntajes promedio en matemáticas PISA por sexo.

Año	Jurisdicción	Mujer		Hombre	
		Promedio	Error Estándar	Promedio	Error Estándar
2018	Promedio OCDE	481	(1,2)	490	(1,4)
	México	403	(2,7)	414	(2,8)
	Brasil	379	(2,0)	387	(2,5)
2015	Promedio OCDE	481	(1,3)	490	(1,4)
	México	404	(2,4)	412	(2,7)
	Brasil	370	(3,0)	385	(3,2)
2012	Promedio OCDE	489	(1,3)	501	(1,4)
	México	406	(1,4)	420	(1,5)
	Brasil	380	(2,2)	397	(2,1)
2009	Promedio OCDE	490	(1,4)	505	(1,4)
	México	412	(1,9)	425	(2,1)
	Brasil	379	(2,6)	394	(2,4)

Año	Jurisdicción	Mujer		Hombre	
		Promedio	Error Estándar	Promedio	Error Estándar
2006	Promedio OCDE	486	(1,4)	497	(1,4)
	México	401	(3,1)	410	(3,4)
	Brasil	361	(3,3)	379	(3,5)
2003	Promedio OCDE	494	(1,3)	504	(1,4)
	México	380	(4,1)	391	(4,3)
	Brasil	348	(4,4)	365	(6,1)

Fuente: PISA, elaboración propia.

Tabla 10. Puntajes promedio en ciencias PISA por sexo.

Año	Jurisdicción	Mujer		Hombre	
		Promedio	Error Estándar	Promedio	Error Estándar
2018	Promedio OCDE	494	(0,5)	495	(0,5)
	México	414	(2,3)	424	(2,6)
	Brasil	404	(2,4)	403	(2,5)
2015	Promedio OCDE	491	(0,5)	495	(0,5)
	México	412	(2,3)	420	(2,6)
	Brasil	399	(2,4)	403	(2,5)
2012	Promedio OCDE	493	(1,4)	497	(2,5)
	México	411	(1,3)	418	(1,5)
	Brasil	401	(2,2)	402	(2,3)
2009	Promedio OCDE	505	(1,3)	508	(1,5)
	México	412	(2,1)	419	(2,1)
	Brasil	403	(2,6)	407	(2,5)
2006	Promedio OCDE	498	(1,4)	500	(1,7)
	México	406	(2,7)	413	(3,2)
	Brasil	386	(2,9)	395	(3,2)

Fuente: PISA, elaboración propia.

3.2.1 Comparación de medias por sexo en cada país

Más allá de la simple observación, para analizar si esta diferencia es estadísticamente significativa, comparamos estos promedios mediante una prueba *t* para muestras independientes. En términos de rendimiento académico por cada disciplina, se aprecian diferencias significativas según sexo (ver tablas 8, 9 y 10), y de acuerdo a la prueba de igualdad de varianzas de Levene, deberán asumirse varianzas distintas en cada una de las mediciones. Los resultados fueron los siguientes:

Tabla11. Diferencia de Medias en lectura según sexo, PISA.

Year	CNT	t	df	Sig. (2-tailed)
2018	BRA	185,49	2032825	0,00**
	MEX	81,08	1466784	0,00**
	OCDE	397,15	12981100	0,00**
2015	BRA	180,34	2393339	0,00**
	MEX	114,83	936420	0,00**
	OCDE	401,81	13078255	0,00**
2012	BRA	286,33	2311184	0,00**
	MEX	173,06	1313555	0,00**
	OCDE	601,48	13088917	0,00**
2009	BRA	221,05	2018946	0,00**
	MEX	170,02	1295951	0,00**
	OCDE	605,75	12594275	0,00**
2006	BRA	213,17	1794804	0,00**
	MEX	194,09	1170511	0,00**
	OCDE	519,84	9315734	0,00**
2003	BRA	217,89	1841666	0,00**
	MEX	117,34	1058882	0,00**
	OCDE	601,26	18978309	0,00**
2000	BRA	168,15	2903391	0,00**
	MEX	114,83	936420	0,00**
	OCDE	492,91	11424795	0,00**

Fuente: PISA, elaboración propia. ** $p < 0,01$ alfa 0,0.5

Tabla 12. Diferencia de Medias en matemáticas según sexo, PISA 2000-2015.

Año	CNT	t	df	Sig. (2-tailed)
2018	BRA	-70,24	2024620	0,00**
	MEX	-92,29	1456714	0,00**
	OCDE	-162,76	12988907	0,00**
2015	BRA	-135,49	2392621	0,00**
	MEX	-57,40	1388440	0,00**
	OCDE	-176,33	13076038	0,00**
2012	BRA	-164,25	2316618	0,00**
	MEX	-108,73	1311697	0,00**
	OCDE	-221,23	13122101	0,00**
2009	BRA	-137,77	2026537	0,00**
	MEX	-99,02	1296745	0,00**
	OCDE	-276,12	12608542	0,00**

Año	CNT	t	df	Sig. (2-tailed)
2006	BRA	-139,25	1814959	0,00**
	MEX	-57,46	1170196	0,00**
	OCDE	-214,02	12911853	0,00**
2003	BRA	-113,21	1850990	0,00**
	MEX	-65,99	1057815	0,00**
	OCDE	-221,94	19017506	0,00**

Fuente: PISA, elaboración propia. ** $p < 0,01$ alfa 0,05.

Tabla 13. Diferencia de Medias en ciencias según sexo, PISA 2000-2015.

Año	CNT	t	df	Sig. (2-tailed)
2018	BRA	-1,04	2027691	0,34
	MEX	-76,73	1469358	0,00**
	OCDE	-29,65	12983755	0,00**
2015	BRA	-35,59	2383801	0,00**
	MEX	-67,97	1386975	0,00**
	OCDE	-100,25	13063615	0,00**
2012	BRA	-2,01	2320518	0,06
	MEX	-51,24	1311524	0,00**
	OCDE	-50,61	13109653	0,00**
2009	BRA	-28,35	2026420	0,00**
	MEX	-47,68	1295986	0,00**
	OCDE	-72,56	12602880	0,00**
2006	BRA	-68,33	1802415	0,00**
	MEX	-44,93	1174991	0,00**
	OCDE	-44,25	12906061	0,00**

Fuente: PISA, Elaboración propia. ** $p < 0,01$ alfa 0,05.

Según la evidencia disponible, el factor sexo se asocia significativamente con las diferencias de medias en país y en cada disciplina, con un 95% de confianza. Sin embargo, no hay que perder de vista que estas comparaciones se basan sobre estimaciones estadísticas, y que el valor de los resultados radica, en este caso, en su posible variabilidad respecto al objeto del estudio, las políticas públicas. Resulta destacable la obtención de resultados favorables en ciencias en Brasil en 2012 y 2018, donde las diferencias por sexo dejaron de ser significativas. Es de esperar que se transforme en un dato duradero.

3.3 La problemática de la etnicidad

A partir de la constatación sobre el panorama en las brechas por género y país observadas en los resultados PISA, nos abocamos a datos censales para poder triangular los datos, y de esa forma proyectar un análisis sobre estas variables.

3.3.1 El panorama étnico en México

De acuerdo a datos censales del año 2000, la población indígena⁴ representaba un 6,4% del total, mientras que la económicamente activa sólo llegaba a un 5,8%. De dicho porcentaje, tan solo un 29,5% eran mujeres, pese a que éstas representaban un 52,2% del grupo étnico⁵ (CELADE). Si observamos los datos del censo 2010 (Tab. 14) , podemos constatar que la participación de mujeres de *habla indígena* tiene niveles más elevados que las mujeres de *habla no indígena*, aunque su proporción respecto de los hombres se mantenga en niveles bajos.

Tabla n°14. México: población económicamente activa, mayor de 12 años, por sexo y habla indígena.

		Total	Habla lengua indígena	%	No habla lengua indígena	%
Ocupada	Total	42.669.675	2.729.354	96,8%	39.804.980	95,4%
	Hombre	28.447.257	2.106.719	96,3%	26.254.490	94,6%
	Mujer	14.222.418	622.635	98,4%	13.550.490	97,0%
Desocu- pada	Total	2.031.369	90.628	3,2%	1.934.365	4,6%
	Hombre	1.597.881	80.704	3,7%	1.512.354	5,4%
	Mujer	433.488	9.924	1,6%	422.011	3,0%
Total	Total	44.701.044	2.819.982	100,0%	41.739.345	100,0%
	Hombre	30.045.138	2.187.423	100,0%	27.766.844	100,0%
	Mujer	14.655.906	632.559	100,0%	13.972.501	100,0%

Fuente: INEGI, Censo de Población y Vivienda 2010. Elaboración propia.

⁴ En este punto nos encontramos con ciertas variables categoriales que ameritan reflexión, ya que la variable de etnicidad se encuentra definida en función del habla indígena, y no necesariamente de color de piel, ni de autorreconocimiento subjetivo, definición acorde a lo señalado por Rojas (1990) sobre rasgos considerados relevantes por los propios sujetos.

⁵ Por otra parte, de acuerdo a CELADE, la tasa de desempleo en jóvenes entre 15 y 29 años, alcanzaba un promedio de 1,32% en población indígena, 1% en mujeres indígenas, y 1,93% para el resto de la población.

Asimismo, es posible observar otras tendencias. En términos generales, a mayor tasa de escolaridad, mayor participación económica (ver Tab.15), tanto para personas de habla indígena para quienes no hablan lengua indígena, y en mujeres se observa con mayor claridad que en hombres.

Tabla 15. México: Tasa de participación económica por sexo y habla indígena (%).

		Total	Habla lengua indígena	No habla lengua indígena	Diferencia por lengua	p (chi)
Total	Total	52,63	48,75	53,01	4,26%	
	Hombre	73,37	77,29	73,23	-4,06%	0,000**
	Mujer	33,32	21,41	34,23	12,82%	0,000**
Sin Escolaridad	Total	37,23	38,21	36,96	-1,25%	
	Hombre	68,21	78,48	65,59	-12,89%	0,000**
	Mujer	15,72	15,57	15,78	0,21%	0,000**
Ed. Básica	Total	48,69	50,43	48,56	-1,87%	
	Hombre	72,48	77,73	72,04	-5,69%	0,000**
	Mujer	26,2	20,93	26,61	5,68%	0,000**
Ed. Superior	Total	70,96	72,59	70,96	-1,63%	
	Hombre	79,98	78,23	80,04	1,81%	0,206
	Mujer	61,98	64,18	61,97	-2,21%	0,146

Fuente: INEGI, Censo de Población y Vivienda 2010, elaboración propia ** $p < 0,01$ alfa: 0,05

Otro dato llamativo es que, en personas sin escolaridad, los hombres indígenas tienen un 12,9% más de participación económica que sus pares sin habla indígena. Por otra parte, si observamos el cálculo de diferencias por lengua⁶ las mujeres de *habla no indígena* tienen una participación mayor que las mujeres de *habla indígena* (12,8%). Las mujeres de *habla indígena* sólo disminuyen la brecha de participación económica por género, y por factor étnico cuando poseen educación superior (-2,2%). Los datos sugieren que, para mejorar su participación económica, las mujeres de *habla indígena* requieren mayor preparación que las mujeres no indígenas. Dicha interpretación se refuerza al aplicar la prueba de Chi-cuadrado en las diferencias por lengua (respecto del promedio por categoría), donde dicha correlación sólo deja de ser significativa en personas con educación superior, independiente de su habla (ver tabla 15).

Para observar mejor que ocurre en este nivel de logro educativo, realizamos el cruce con tramos de edad en personas con educación superior (véase tablas 16 y 17), y obtenemos un grado de asociación significativa que

⁶ Diferencia por lengua = tasa de participación económica de no habla indígena - tasa de participación económica de habla indígena.

se manifiesta a partir de los 45 años en personas de no habla indígena, y a partir de los 55 años en personas de habla indígena, dato que nos sugiere diferencias en la incorporación al sistema de educación superior en cada grupo, en décadas precedentes.

Tabla 16. Tasa de Participación Económica (%) por sexo y edad, con educación superior, no habla indígena, 2010.

Edad	Hombres	Mujeres	Total	p(chi)
15 - 24	29,485	25,455	27,29	0,58
25 - 34	92,55	75,89	83,73	0,20
35 - 44	98,25	75,185	86,5	0,08
45 - 54	95,665	68,985	83,565	0,04*
55 - 64	75,765	41,815	62,385	0,00**
65 - 74	44,425	18,91	34,735	0,00**
75 - 84	25,75	9,79	19,355	0,01*
> 85	17,62	7,21	12,99	0,04*

Fuente: INEGI, Censo de Población y Vivienda 2010, elaboración propia. * $p<0,05$ ** $p<0,01$

Tabla17. Tasa de Participación Económica (%) por sexo y edad, con educación superior, habla indígena, 2010.

Edad	Hombres	Mujeres	Total	p(chi)
15 - 24	27,165	24,515	25,835	0,71
25 - 34	90,78	78,255	85,28	0,33
35 - 44	97,865	86,79	93,465	0,41
45 - 54	92,375	69,175	84,875	0,06
55 - 64	64,56	39,705	58,36	0,01*
65 - 74	40,5	21,815	36,485	0,01*
75 - 84	25,96	9,995	22,45	0,01*
> 85	15,73	9,24	14,11	0,17*

Fuente: INEGI, Censo de Población y Vivienda 2010, elaboración propia. * $p<0,05$

En síntesis, los resultados en México sugieren que el factor de habla indígena no es un impedimento para insertarse en el mercado laboral, sino más bien que podría tener un impacto en su nivel de ingresos, que se reduce a medida que aumenta el logro educativo. En virtud de dichos resultados, podríamos interpretar que el factor de *habla indígena* opera como un diferenciador negativo, al igual que el factor sexo, en lo que respecta a participación económica, mediado por la variable educacional, tal como lo sugieren otros estudios que realizan análisis de regresión multivariante (Araújo, 2015; Monsueto, *et al.*, 2005 y Yabrudy, 2011).

3.3.2 El panorama étnico en Brasil

Según datos censales del año 2000, la población indígena (y económicamente activa) representaba un 0,4% del total, de las cuales un 39,6% eran mujeres, (y 49,9% del total de acuerdo a CELADE). En el caso de la población negra o *parda* económicamente activa, representaba un 45% del total, porcentaje del cual un 38,6% eran mujeres (48,5% del total)⁷. Al abordar los datos del Censo 2010 (Ver Tabla n°18), podemos observar una cierta segmentación por factor de etnicidad y sexo. Si comparamos la columna de población (%) versus la de población con ingresos (%), se observa un efecto de sobrerrepresentación de la población blanca (48% a 50,7%) en la población con ingresos, así como en la población negra (8% a 8,2%). Dicha segmentación en población negra favorece a hombres (4,14% a 4,62%) en desmedro de las mujeres negras (3,87% a 3,63%).

Tabla 18. Brasil: personas de 10 años o más, con ingresos, según sexo y color o raza (2010).

Color o raza	Población Total	(%)			Población con ingresos	(%)		
		Total	Hombre	Mujer		Total	Hombre	Mujer
Blanco	77.787.902	48,02	22,83	25,19	51.656.119	50,69	26,71	23,98
Negro	12.974.794	8,01	4,14	3,87	8.407.968	8,25	4,62	3,63
Amarillo	1.824.789	1,13	0,51	0,62	1.153.945	1,13	0,56	0,57
Parda	68.779.712	42,46	20,96	21,5	40.390.634	39,64	21,46	18,18
Indígena	616.927	0,38	0,19	0,19	290.375	0,28	0,15	0,14
Total	161.990.266	100	48,62	51,38	101.900.929	100	53,5	46,5

Fuente: IBGE - Censo demográfico 2010, elaboración propia.

Al analizar la población *parda*, pese a quedar subrepresentada, la situación se repite: varones pasan de 20,9% a 21,4% y mujeres de 21,5% a 18,18%. A partir de dichos datos, podríamos interpretar que, en el caso de Brasil, los factores de etnicidad y sexo actúan como factor que incide en la inserción laboral y nivel de ingresos, al igual que en México.

Dicha interpretación es posible de reafirmar al analizar el valor de ingresos medios por sexo y color o raza, donde se observa además una brecha total de ingresos por género importante (29,3%). Para analizar dicha estructura, se utiliza un coeficiente simple (Dif. H/M) el cual representa la proporción entre valor medio total de la categoría y su respectivo género. De forma análoga, el valor (Dif. Raza) representa la proporción entre valor medio

⁷ El año 2000, la tasa de desempleo en jóvenes entre 15 y 29 años, alcanzaba un promedio de 20,4% en población afro, 26,8% en mujeres negras o pardas, y 17,5% para el resto de la población joven. En población indígena, la tasa de desempleo en jóvenes entre 15 y 29 años, alcanzaba un promedio de 18%, 25% en mujeres indígenas. (CELADE).

total con el valor medio por raza. En definitiva, podemos observar que todos los varones obtienen ingresos superiores a su media por raza, siendo los menos favorecidos los varones negros (+13%) y los varones *pardos* (+15%). Por su parte, en mujeres, la población con mayor diferencia respecto de sus pares por raza, son las mujeres *amarillas* (-22%), pese a que los varones *amarillos* obtienen los ingresos medios más altos de la escala.

Tabla 19. Brasil: personas de 10 años o más. Valor del ingreso promedio mensual nominal según sexo y color o raza (2010).

Color o raza	Total	Hombres	Mujer	Dif. H (%)	Dif. M (%)	Dif. Raza (%)
Total (Media)	1201,47	1390,99	983,37	16%	-18%	
Blanco	1535,47	1795,87	1245,43	17%	-19%	28%
Negro	832,25	943,05	691,2	13%	-17%	-31%
Amarillo	1573,68	1915,21	1233,7	22%	-22%	31%
Parda	843,87	973,28	691,08	15%	-18%	-30%
Indígena	732,93	858,5	602,44	17%	-18%	-39%

Fuente: IBGE - Censo demográfico 2010, elaboración propia

Finalmente, el coeficiente por raza nos muestra que tan solo las poblaciones blanca o amarilla obtiene ingresos medios por sobre la media: 28% y 31% respectivamente. La población negra obtiene ingresos en promedio un 31% más bajo, la población *parda* un 30%, y la población indígena, un 39%.

Por otra parte, para correlacionar los factores de escolaridad y participación económica, se realizaron pruebas de Chi-cuadrado, pudiendo observar que, en población blanca no es una asociación significativa, salvo en educación básica. En población *parda* el resultado es similar, aunque se añade la correlación a nivel de educación superior completa.

Tabla 20. Tasa de Participación Económica (%) por sexo y escolaridad, población blanca, 2010.

	Hombres	Mujeres	Promedio Total Hombres	Promedio Total Mujeres	p (chi)
Sin Escolaridad	5,45	3,28	6,11	4,45	0,538
Básica Completa	2,87	2,11	9,23	8,6	0,002**
Media Completa	4,95	4,55	3,07	3,92	0,263
Superior Completa	2,28	2,79	5,45	3,28	0,166
Total	15,62	12,79	14,06	8,01	0,082

Fuente: IBGE - Censo demográfico 2010, elaboración propia. **p<0,01

Tabla 21. Tasa de Participación Económica (%) por sexo y escolaridad, población parda, 2010.

	Hombres	Mujeres	Promedio total hombres	Promedio total mujeres	p (chi)
Sin Escolaridad	6,98	3,83	6,11	4,45	0,647
Básica Completa	2,63	1,9	9,23	8,6	0,002**
Media Completa	3,43	3,28	3,07	3,92	0,702
Superior Completa	0,60	0,89	5,45	3,28	0,014*
Total	2,83	1,96	14,06	8,01	0,000**

Fuente: IBGE - Censo demográfico 2010, elaboración propia. * $p<0,05$ ** $p<0,01$

Sin embargo, la fragmentación de resultados se produce en la población negra, la cual arroja correlación positiva entre nivel de escolaridad y participación económica en todos los niveles.

Tabla 22. Tasa de Participación Económica (%) por sexo y escolaridad, población negra, 2010.

	Hombres	Mujeres	Promedio total hombres	Promedio total mujeres	p (chi)
Sin Escolaridad	1,44	0,77	6,11	4,45	0,01*
Básica Completa	0,54	0,37	9,23	8,6	0,00**
Media Completa	0,72	0,65	3,07	3,92	0,033*
Superior Completa	0,12	0,16	5,45	3,28	0,004**
Total	2,83	1,96	14,06	8,01	0,00**

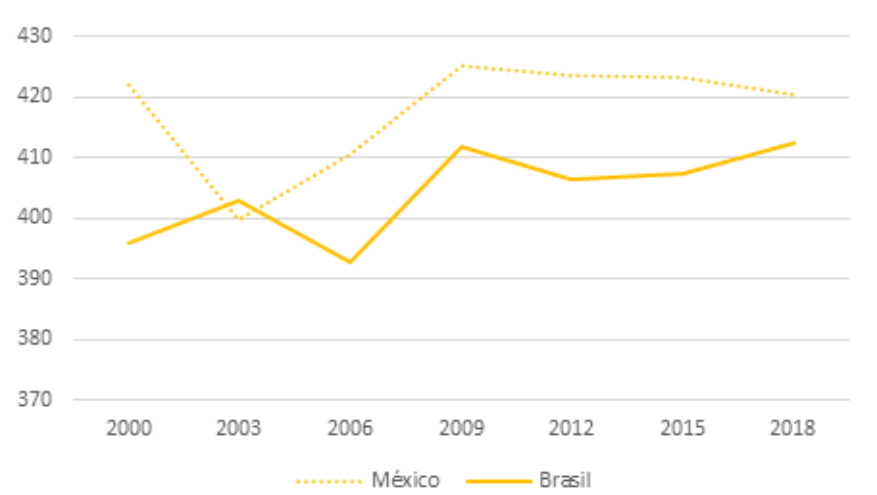
Fuente: IBGE - Censo demográfico 2010, elaboración propia. * $p<0,05$ ** $p<0,01$

58

3.4 Discusión sobre resultados

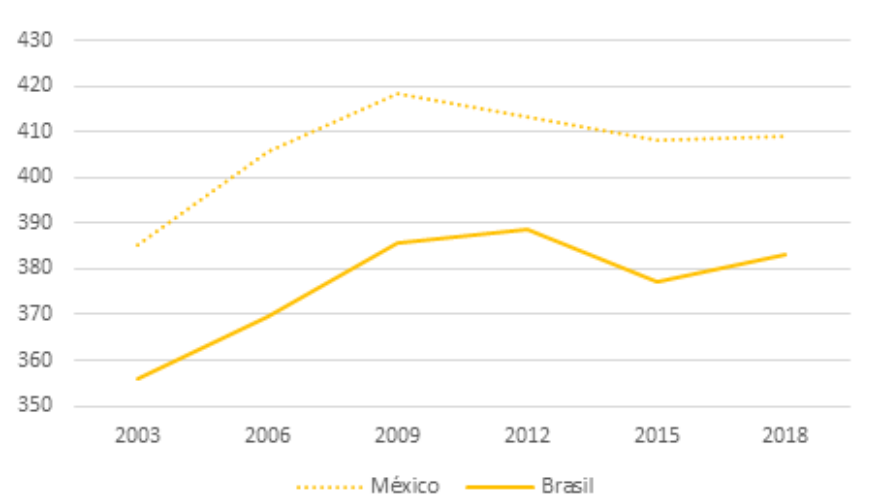
De acuerdo a los resultados obtenidos, es posible realizar algunas comparaciones de interés entre Brasil y México. Ambos países han mostrado progresos en PISA (salvo México en el caso de Lectura), a diferencia de lo que ocurre con el promedio de países OCDE, que tiende sorprendentemente a la baja, pese a que estamos hablando de niveles significativamente distintos de logro, como vimos en la sección 3.1.

Figura1. Puntajes promedio en lectura, Brasil y México



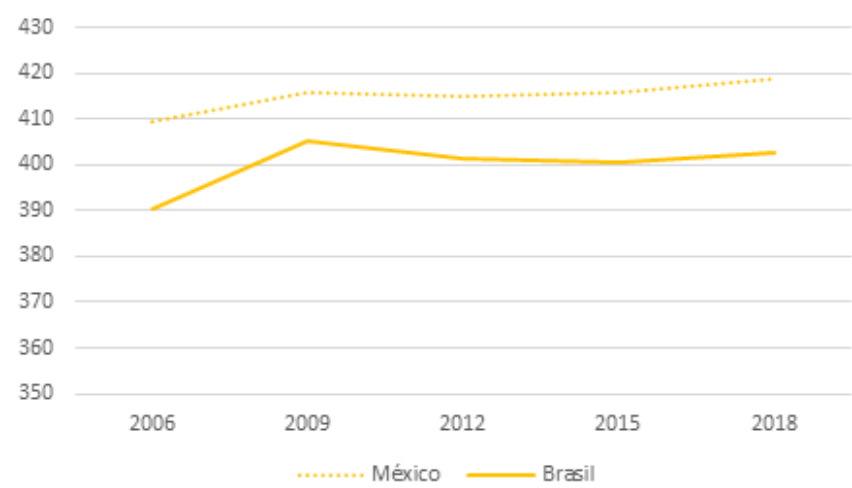
Fuente: PISA. Elaboración propia.

Figura 2. Puntajes promedio en matemáticas, Brasil y México



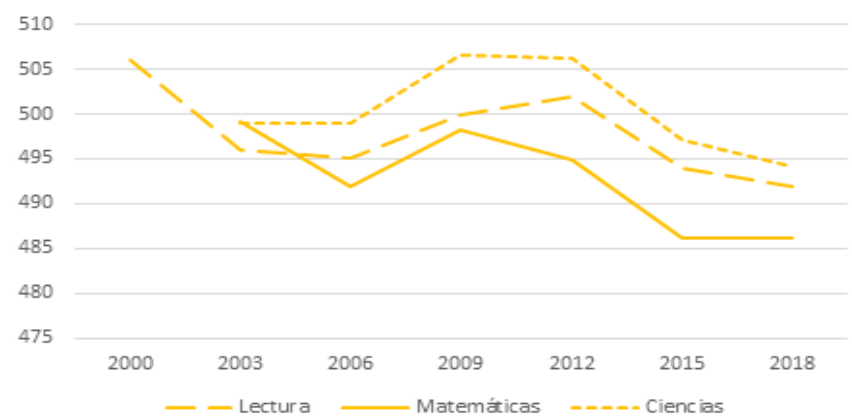
Fuente: PISA. Elaboración propia.

Figura 3. Puntajes promedio en ciencias, Brasil y México



Fuente: PISA. Elaboración propia.

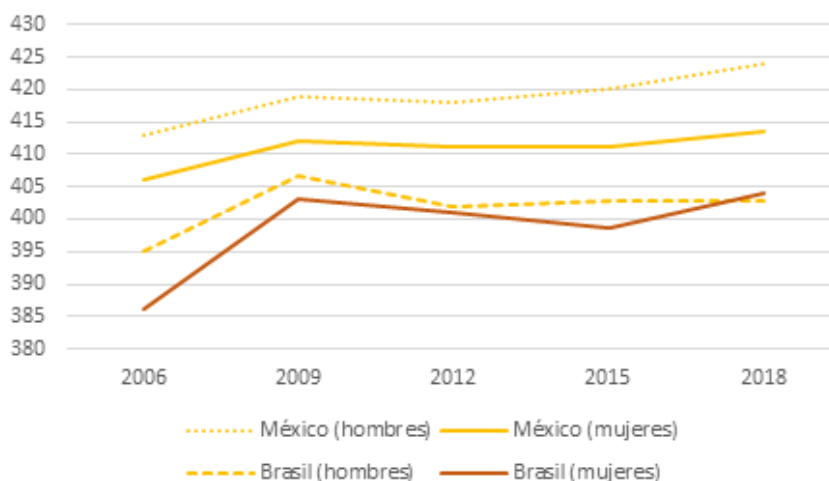
Figura 4. Puntajes promedio en lectura, matemáticas y ciencias, OCDE



Fuente: PISA. Elaboración propia.

Otro punto a destacar es que, a nivel de brecha de rendimiento por género, se aprecian resultados importantes para Brasil, particularmente en el caso de ciencias, que se explica debido a un incremento en los rendimientos de mujeres al punto eliminar la brecha de género los años 2012 y 2018 (figura 5).

Fig. 5 Puntajes PISA promedio en ciencias, por sexo (Brasil, México)



Fuente: PISA. Elaboración propia.

Asimismo, resultan relevantes ciertos resultados sobre correlación positiva entre nivel de escolaridad y participación económica. Mientras en México las mujeres de habla indígena sólo disminuyen su brecha de participación económica por género, y por factor étnico cuando poseen educación superior; en Brasil la fragmentación de resultados se produce en la población negra en todos los niveles de escolaridad, y la educación superior se correlaciona significativamente con participación económica de igual forma en población *parda* y *negra*, afectando en mayor medida a las mujeres. Ello quiere decir que, pese a los esfuerzos en políticas públicas en las últimas décadas, y a las mejoras obtenidas en ámbito educativo, en ambos casos, no ha sido suficiente para equiparar la situación en términos socioeconómicos por género y/o etnia, en relación a la población en general.

4. Conclusiones

En la búsqueda de una perspectiva compleja sobre esta problemática, este estudio nos ha permitido visualizar bajo la óptica de la equidad, los resultados obtenidos en dos países que concentran buena parte de la población latinoamericana y que modelan sus políticas públicas desde nociones clásicas de igualitarismo, sin alcanzar a revertir la problemática de la desigualdad de género y/o etnicidad de forma relevante en el marco temporal analizado. Al observar longitudinalmente los datos, es posible interpretar que

nos encontramos ante estructuras profundas que se manifiestan de forma persistente en sus resultados educativos y brechas socioeconómicas, pese a un marco jurídico-político que intenta compensar dichos desequilibrios sociales.

Asimismo, la evidencia permite constatar una fragmentación de resultados académicos por país, disciplina y sexo de los estudiantes, así como su comparación con el promedio OCDE. Además, su triangulación como datos socioeconómicos confirma ciertos diagnósticos internacionales sobre brechas de género constituidas principalmente por los factores educativos, políticos, de salud, y de participación económica (World Economic Forum, 2016, 2017, 2018, 2020). De acuerdo a datos recientes (WEF, 2020), en Brasil se observa una reducción de la brecha de género en educación y salud, aunque no mejora en su brecha de participación económica (2020), algo que se ve aún más lejano desde la derogación en 2019 de las cuotas de matrícula por “raza” o color en educación superior (y de candidaturas políticas en mujeres). En el caso de México, dicho estudio señala que, si bien se observan mejoras en empoderamiento femenino, su principal problema radica en la participación económica e incorporación al mercado laboral (2020), perpetuando la desigualdad. Este ejemplo nos alerta sobre el riesgo de atender la problemática del género desde una óptica prevalentemente formal y no desde un proceso educativo y de cambio cultural desde las bases, que permita integrar la problemática de la etnicidad como factor relevante, evitando la persistencia conceptos de *asimilación* que terminan reproduciendo estructuras sociales de dominación económica basada en categorías que tienen su origen, en este caso, en el período colonial.

Por otra parte, resulta llamativo que, en un instrumento tan ampliamente difundido como PISA, no se apliquen categorías de análisis sobre etnicidad (tan sólo en EEUU) que permitan mejorar el proceso de evaluación y formulación de políticas públicas en el ámbito educativo de forma más pertinente en relación a variables multiculturales. Si bien se trata de un hallazgo tangencial, resulta del todo relevante dar aviso a las entidades nacionales, y a la propia OCDE, sobre la posibilidad de mejorar su instrumento en una dimensión tan relevante, que genera fracturas sociales en todo el mundo hasta el día de hoy.

Una posible respuesta podría ser la generación de un enfoque *interseccional* en el análisis, formulación, y evaluación de políticas públicas, así como en instrumentos de evaluación educativa como PISA. Ello permitiría abordar factores clave de desigualdad (como género y etnicidad) que se manifiestan desde el nivel escolar al mundo del trabajo, determinando una estructura de brechas que no es posible eliminar mediante políticas sectoriales que operan de forma disociada y/o bajo un enfoque asistencial. Cabe recordar que son precisamente dichas políticas las que intentan paliar

problemáticas que han sido señadas por autores como David Harvey como factores relevantes en el proceso de *acumulación por desposesión* (Harvey, 2003, p.117). Finalmente, el objetivo de la política pública, y ciertamente, de la política educativa, debe acercarnos hacia la construcción de un orden social más equitativo, y para ello resulta ineludible que los ejes de exclusión institucionalizados y socialmente configurados deban sufrir profundas reestructuraciones, proceso que pasa necesariamente por el desafío de *deconstruir* los ejes y las categorías poscoloniales de exclusión social, como lo son el género y la etnicidad, aunque esta vez, de forma indisoluble.

Referencias

- Abramo, L. y Valenzuela, M. (2006). Inserción laboral y brechas de equidad de género en América Latina, *Trabajo decente y equidad de género en América Latina*, 29.
- Alberti, P. (1995). Mujeres indígenas en organizaciones campesinas. En Barceló, R.; Portal, M. Sánchez, J. (coord.) (1995). *Diversidad Étnica y Conflicto en América Latina*. UNAM. Plaza y Valdés. México D.F. 81-106.
- Araújo, A. (2015). La desigualdad salarial de género medida por regresión cuantílica: el impacto del capital humano, cultural y social, *Revista Mexicana de Ciencias Políticas y Sociales*, 223(60).
- Araújo, A. (2014). Desigualdad Salarial de Género *Revista Mexicana de Ciencias Políticas y Sociales*. UNAM, 223(15) 287-316.
- Barth, F. (1976). *Los grupos étnicos y sus fronteras. La organización social de las diferencias culturales*. Fondo de la Cultura Económica: México.
- Bello, A. y Rangel, M. (2002). La equidad y la exclusión de los pueblos indígenas y afrodescendientes en América Latina y el Caribe. *Revista de la CEPAL*, 76. 39-54.
- Bergmann, B. (1996). *In defense of affirmative action*. New York: BasicBooks.
- Castro-Gómez, S. (1998). *Teorías sin disciplina* (latinoamericanismo, poscolonialidad y globalización en debate). México: Miguel Ángel Porrúa.
- CELADE. División de Población. Pueblos Indígenas y Afrodescendientes de la CEPAL. Recuperado de <https://bit.ly/3cmPU7m>
- CEPAL (2013). *Mujeres indígenas en América Latina: dinámicas demográficas y sociales en el marco de los derechos humanos*. Centro Latinoamericano y Caribeño de Demografía (CELADE)-División de Población y División de Asuntos de Género de la CEPAL: Santiago de Chile.
- Constitución Política de los Estados Unidos Mexicanos (2020). de 05 de febrero de 1917. Secretaría de Servicios Parlamentarios.
- Constituição da República Federativa do Brasil (1988). Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. Brasília DF.

- Cruz, A. (2005). *O direito à diferença. As ações afirmativas como mecanismo de inclusão social de mulheres, negros, homossexuais e pessoas portadoras de deficiência*. 2. ed. Belo Horizonte: Del Rey.
- De la Rica, S. y A. González. (2013). *Brechas de género en los resultados de PISA: El impacto de las normas sociales y la transmisión intergeneracional de las actitudes de género*. Madrid: Fedea.
- De Sousa, B. (Ed.). (2001). *Globalização: fatalidade ou utopia?* Porto: Editora.
- De Souza (2013). *Descolonizar el saber, reinventar el poder*. Santiago de Chile: LOM.
- Decreto nº 91.144/85. *Ministério da Cultura, Gaceta Oficial* - Sección 1 - 15/3/1985, Folio 4703. Brasília, DF.
- Díaz, H. (1985). *La cuestión étnico nacional*. Primera Línea: México.
- Durham, E. R. (2003). *Desigualdade educacional e quotas para negros nas universidades*. NUPES/USP: São Paulo.
- Dussel, E. (2005). *Transmodernidad e interculturalidad*. Ciudad de México: Universidad Autónoma de México.
- Espinoza, O., González, L. E. y Uribe, D. (2009). Movilidad social en Chile: El caso del gran Santiago urbano. *Revista de Ciencias Sociales*, 15(4), 586-606.
- Feinstein, O. (2012). La institucionalización de la evaluación de políticas públicas. *América Latina. Presupuesto y Gasto Público* 68, 41-52.
- García-Holgado, A., Camacho Díaz, A. y García-Peñalvo, F. J. (2019). La brecha de género en el sector STEM en América Latina: Una propuesta europea. In M. L. Sein-Echaluze Lacleta, Á. Fidalgo-Blanco, & F. J. García-Peñalvo (Eds.), *Actas del V Congreso Internacional sobre Aprendizaje, Innovación y Competitividad*. CINAIC 2019 Madrid. 704-709.
- Guideri, R., Pellizzi, F. y Stanley, J. (eds). (1988) *Ethnicities and Nations, Processes of Interethnic Relations in Latin America, Southeast Asia, and the Pacific*. Texas: The Rothko Chapel.
- Harvey, D. (2003). *The New Imperialism*. Oxford University Press.
- Lei Federal nº.: 9.394/96 das Diretrizes e Bases da Educação Nacional. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. Brasília, DF, 1996.
- Lei Federal nº.:11.494/07. Regulamenta o FUNDEB. Presidência da República Casa Civil. Subchefia para Assuntos Jurídicos. Brasília, DF, 2007.
- Lei Federal nº.:13.005/14. Aprova o PNE – Plano Nacional de Educação. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. Brasília, DF, 2014.
- Lei nº 10.683/03. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. Congresso Nacional, Brasília, DF.2003
- Lei nº 12.711/12. Presidência da República. Casa Civil. Subchefia para Assuntos Jurídicos. Congresso Nacional, Brasília, DF. 2012
- Ley de derechos y cultura indígena del estado de Querétaro. Comisión Nacional de Derechos Humanos. 30 de noviembre de 2016.

- Ley de la Comisión Nacional de los Derechos Humanos (1992). *Diario Oficial de la Federación* México D.F., 29 de junio de 1992
- Ley del Instituto Nacional de las Mujeres (2001), *Diario Oficial de la Federación*, México D.F. el 12 de enero de 2001
- Ley general de derechos lingüísticos de los pueblos indígenas (2003). *Diario Oficial de la Federación*. México D.F., 13 de marzo de 2003
- Ley General para la Igualdad entre Mujeres y Hombres. Secretaría de Servicios Parlamentarios. *Diario Oficial de la Federación*, México D.F. 2 de agosto de 2006
- Ley para Prevenir y Eliminar la Discriminación (2003). *Diario Oficial de la Federación*, México D.F., 11 de junio de 2003
- Lugones, M. (2008). Colonialidad y Género. *Tabula Rasa*, 9, 73-101.
- Maggie, Y. y Fry, P. (2002). O debate que não houve: a reserva de vagas para negros nas universidades brasileiras *Revista Brasileira de Educação. Enfoques*, 1(1), 93-117.
- Mignolo, W., Lugones, M., Jiménez-Lucena y Madina, T. (2006). *Género y descolonialidad*. Buenos Aires: Signo.
- Miranda, D., Freire, M. y Jervis, M. (2019). Asimetría, poder y construcción de género, ¿un camino para conseguir igualdad? *Revista de Estudios Políticos y Estratégicos*, 7(1), 104-121.
- Moehlecke, S. (2005). A igualdade que perturba a justiça no mundo moderno: o discurso sobre a ação afirmativa. *GT: Movimentos Sociais e Educação*, 3, 1- 7.
- Monsueto, S., Machado, A. y Golgher, A. (2005). Desigualdades de remuneraciones en Brasil: regresiones por cuantiles y descomposición de las brechas. *Revista de la Cepal*, 90(6), 171-189
- Montes-Berges, B; Del Prado, M (2014). Social dominance and neosexism: Relationship with attitudes towards affirmative action programs. *International Journal of Social Psychology*, 19(3), 275-285.
- National Center for Education Statistics (NCES) PISA Program. Recuperado de <https://bit.ly/3iW3TmT>
- OCDE (2015). *The ABC of Gender Equality in Education: Aptitude, Behaviour, Confidence*. Program for International Students Assessment (PISA): OECD Publishing. Recuperado de <https://bit.ly/2RKlq5E>
- OCDE (2016). *Gender Equality in the Pacific Alliance: Promoting Women's Economic Empowerment*, Paris: OECD Publishing. <https://doi.org/10.1787/9789264262959-en>.
- OCDE (2017). *Marco de Evaluación y de Análisis de PISA para el Desarrollo: Lectura, matemáticas y ciencias*, Paris: OECD Publishing.
- OECD International Development Statistics (database). Recuperado de: <http://dx.doi.org/10.1787/data-00728-en>
- Plan Nacional de Desarrollo (2007). 2007-2012. Diario Oficial. México D.F. 31 de mayo de 2007
- Plan Nacional de Desarrollo (2013). 2013-2018. Diario Oficial. México D.F. 20 de Mayo de 2013.

- Quijano, A. (2000). Colonialidad del Poder y Clasi? cacion Social. *Journal of world-systems research*, 342-386. <https://doi.org/10.5195/jwsr.2000.228>
- Rojas, F. (1990). *Etnicidad: Teoría y Praxis. La revolución cultural de 1990*. Guatemala: Editorial Cultura.
- Roth, A. (2009). La evaluación de políticas públicas en Colombia: una mirada crítica a partir de las prácticas evaluativas oficiales de los programas de la “Red de Apoyo Social” *Revista del CLAD Reforma y Democracia*, 45, 161-186.
- Sanhueza, F., Cornejo, P. y Leyton, J. (2015). La educación como agente de movilidad social. *Convergencia Educativa*, 6, 95-106.
- Schleicher, A. (2019). PISA 2018. Insights and results. Paris: OCDE Publishing.
- Singer, M. (2014). ¿Exclusión o inclusión indígena? *Estudios Políticos*, 31, 87-106.
- Vidal, F. (2014). La participación política en México: entendiendo la desigualdad entre hombres y mujeres. *Revista Mexicana de Ciencias Políticas y Sociales*, 223(15), 317-356.
- Vieira, A. L. da C. (2003). Políticas de educação, educação como política: observações sobre a ação afirmativa como estratégia política. In: Silva, P. B.G. & Silvério, V. R. (Orgs.). *Educação e ações afirmativas: entre a injustiça simbólica e a injustiça econômica*. INEP, pp. 81-98.
- World Economic Forum (2016). *Global Gender Gap Report 2016*. Recuperado de <https://bit.ly/3kHVrbd>
- World Economic Forum (2017). *Global Gender Gap Report 2017*. Recuperado de <https://bit.ly/32TXpiZ>
- World Economic Forum (2018). *Global Gender Gap Report 2018*. Recuperado de <https://bit.ly/2G1JvSB>
- World Economic Forum (2020). *Global Gender Gap Report 2020*. Recuperado de <https://bit.ly/3kGAugO>
- Yabrudy, J. (2011). Discriminación étnica e ingresos en la Isla de San Andrés, Colombia. *Revista de Economía del Rosario*, 2(14), 229-261.
- Yaschine, Iliana (2015). ¿Alcanza la educación para salir de la pobreza? Análisis del proceso de estratificación ocupacional de jóvenes rurales en México. *Revista Mexicana de Ciencias Políticas y Sociales*, 223(60), 377-405.

Desempenho Ibero-americano nos PISA de 2012, 2015 e 2018

Luiz Carlos Gesqui 

Universidade de Araraquara (UNIARA), Brasil.

Resumo. Os resultados das avaliações em larga escala ocupam centralidade nas políticas educacionais de todos os países, em especial quanto ao que expressam sobre a aprendizagem dos alunos. Este artigo, de base documental, tem como objetivo central analisar os indicadores dos países da Organização dos Estados Ibero-americanos para Educação, Ciência e Cultura (OEI) nas edições de 2012, 2015 e 2018 do Programa Internacional de Avaliação de Alunos (Pisa), bem como apresentar e comparar sua evolução histórica em relação aos principais expoentes do programa. Os resultados obtidos apontam que os países da OEI avaliados pelo Pisa em 2012, 2015 e 2018 podem ser classificados em dois grupos: o primeiro, formado por Portugal e Espanha, ocupa e mantém, nas três edições, posição intermediária entre todos os países avaliados, com pontuação, nos três domínios avaliados – Leitura, Ciências, Matemática –, de no máximo 500 pontos. O segundo grupo, formado pelos demais países da OEI, ocupa e alterna, nas três edições, as vinte últimas posições entre todos os países avaliados, com pontuação, nos três domínios avaliados, de no máximo 475 pontos. Independentemente do grupo, os países da OEI apresentam em 2018 um maior distanciamento em relação aos demais países avaliados se comparados a 2012.

Palavras-chave: avaliações em larga escala; ibero-americano; Pisa.

Desempeño iberoamericano en PISA 2012, 2015 y 2018

Resumen. Los resultados de las evaluaciones a gran escala son fundamentales para las políticas educativas de todos los países, particularmente en lo relativo al aprendizaje de los estudiantes. Este artículo de carácter documental tiene como objetivo analizar los indicadores de los países de la Organización de Estados Iberoamericanos para la Educación, la Ciencia y la Cultura (OEI) en las ediciones de 2012, 2015 y 2018 del Programa Internacional de Evaluación de Estudiantes (PISA), así como presentar y comparar su evolución histórica en relación con los principales exponentes del programa. Los resultados obtenidos indican que los países de la OEI evaluados por Pisa en 2012, 2015 y 2018 pueden clasificarse en dos grupos: el primero, formado por Portugal y España, que ocupa y mantiene, en las tres ediciones, una posición intermedia entre todos los países evaluados, con puntuaciones en los tres ámbitos evaluados (Lectura, Ciencias, Matemáticas) de un máximo de 500 puntos. El segundo grupo, formado por los demás países de la OEI, ocupa alternando, en las tres ediciones, los últimos veinte puestos entre todos los países evaluados, con una puntuación máxima de 475 puntos en las tres áreas evaluadas. Independientemente del grupo, los países de la OEI muestran en 2018 un distanciamiento mayor con respecto al resto de países evaluados en comparación con 2012.

Palabras clave: evaluaciones a gran escala; Iberoamericano; Pisa.

Ibero-American Performance in PISA 2012, 2015 and 2018

Abstract. The results of large-scale assessments are central to the educational policies of all countries, especially when they express about student learning. This paper, of documental basis, aims to analyze the indicators of the countries of the Organization of Ibero-American States for Education, Science and Culture (OEI) in the 2012, 2015 and 2018 editions of the International Student Assessment Program (Pisa), as well as to present and compare its historical evolution in relation to the main exponents of the program. The results obtained show that the OEI countries evaluated by Pisa in 2012, 2015 and 2018 can be classified into two groups: the first, formed by Portugal and Spain, occupies and maintains, in the three editions, an intermediate position among all the countries evaluated, with scores, in the three domains evaluated - Reading, Science, Mathematics, of a maximum of 500 points. The second group, formed by the other OEI countries, occupies and alternates, in the three editions, the last twenty positions among all the countries evaluated, with a score, in the three domains evaluated, of a maximum of 475 points. Regardless of the group, the OEI countries presented in 2018 a greater distance from the other countries evaluated compared to 2012.

Keywords: Large-scale assessments; Ibero-American; Pisa.

1. Introdução

A utilização das avaliações educacionais em larga escala como objeto de estudo, bem como de seus resultados, se tornou recorrente nas últimas décadas. É com base nos resultados dessas avaliações – geralmente expressos em indicadores estatísticos – que as instâncias administrativas responsáveis pela elaboração, definição, implantação, acompanhamento e avaliação de políticas educacionais balizam suas principais decisões e prestam contas à sociedade sobre essas políticas. Tome-se como exemplo o atual Plano Nacional de Educação (PNE) (Brasil, 2014), que estabelece datas e metas a serem alcançadas nos mais variados segmentos da educação, bem como o amplo acompanhamento do cumprimento ou não desses parâmetros.

Atualmente se encontram em uso diversos programas ou sistemas de avaliação em larga escala no âmbito dos estados ou do país (Bauer, 2019), cujos resultados subsidiam pesquisas relacionadas com objetos de estudo variados, como a premiação financeira a profissionais da educação em função de resultados obtidos em avaliações externas (Vicino, 2013); os desdobramentos dos resultados das avaliações externas no trabalho pedagógico das escolas (Cosso, 2013) ou a utilização dos resultados de avaliações externas na cobrança, via judiciário, de direitos educacionais (Franco, 2018).

68

No âmbito da América Latina, o Estudo Regional Comparativo e Explicativo (Erce) investiga, desde 1997, a qualidade da educação do ensino fundamental dos países da América Latina e Caribe (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020a). Contudo, Vaillant e Zidan (2016) enfatizam que, apesar das importantes transformações pelas quais os sistemas educacionais latino-americanos têm passado nas últimas décadas, práticas pautadas mais em experiências do que em conhecimento científico – como a atuação dos diretores de escola – ainda são recorrentes em alguns desses países.

No que se refere ao âmbito mundial, o parâmetro mais utilizado é o Programa Internacional de Avaliação de Alunos (Pisa), por sua abrangência, longevidade e bancos de dados disponíveis (Organização para a Cooperação e Desenvolvimento Econômico, 2020a). Dada sua importância e constatada ascensão, influencia diretamente nos rumos das políticas educacionais de diversos países.

Apesar da contestação de que tem sido alvo, por razões de ordem epistemológica, metodológica, ou mesmo política, o PISA obteve uma consagração expressiva no campo educacional, onde vem servindo para legitimar medidas políticas, instigar pesquisas e animar debates públicos. É possível dizer que se trata de um objeto ao qual políticos, profissionais da comunicação

social, investigadores, peritos e altos quadros da administração recorrem com frequência para escrutinar, problematizar ou apontar soluções para o setor da educação (Carvalho, 2016, p.671).

Segundo Lingard (2016), os resultados do Pisa possibilitam inúmeras análises, dentre elas a comparação entre países com diferentes realidades econômicas, culturais e sociais; em outras palavras, é possível elaborar um ranqueamento desses países por meio de seus indicadores. Pizarro e Lopes Júnior (2017) constatam que, em muitas produções científicas, os indicadores do Pisa são utilizados quase que exclusivamente para ranquear e selecionar os participantes de determinada produção acadêmica ou simplesmente para delimitar o campo empírico de variadas pesquisas em diferentes áreas.

O ranqueamento obtido por meio dos resultados do Pisa também é objeto de estudo de Klein (2011), ao relatar que diversas peculiaridades de cada país não são captadas nessas avaliações e que a interpretação superficial desses resultados pode conduzir os responsáveis pelas políticas educacionais à tomada de decisões pautadas em informações ainda não analisadas em profundidade, diminuindo assim as possibilidades de sucesso das ações propostas. Pinto, Silva e Bixirão Neto (2016), em estudo comparativo entre alunos brasileiros e portugueses, identificam que fatores como a organização e estrutura do sistema educativo, os contextos sociais e econômicos e as características das escolas de cada país impactam com graus de intensidade diferente cada grupo investigado.

No mesmo sentido, Salokangas e Kauko (2015), ao citarem o exemplo dos expressivos resultados da Finlândia no Pisa 2012, enfatizam que o sucesso, a estagnação ou o fracasso de uma política educacional estão relacionados a variáveis – em sua maioria não captadas pelas avaliações em larga escala – que necessitam de investigação, uma vez que foram forjadas a partir dos diferentes contextos de cada país; portanto, a simples transposição de fórmulas de sucesso de um país para outro não corresponde necessariamente à garantia dos mesmos resultados no Pisa.

Gorur (2016) discorre que análises secundárias que considerem as peculiaridades, realidades ou variáveis de cada país são fundamentais para subsidiar os responsáveis pelas políticas educacionais na tomada de decisões que visem alavancar os resultados do Pisa. Essa conclusão é corroborada por Villani e Oliveira (2018), que buscam, em função do fortalecimento do Pisa como parâmetro da educação mundial, possíveis relações deste com um indicador de qualidade da educação básica nacional, no caso o Índice de Desenvolvimento da Educação Básica (Ideb). Baccheto (2016) contribui com o tema ao utilizar os indicadores do Pisa para evidenciar que a taxa de retenção escolar no Brasil é uma das mais elevadas do mundo.

As produções acadêmicas aqui expostas certamente contribuem para a compreensão do desempenho Pisa de cada participante; entretanto, outros aspectos relacionados ao programa carecem de investigação, como a evolução histórica dos indicadores de desempenho de um participante ou de um conjunto de participantes. Tome-se como exemplo a Organização dos Estados Ibero-americanos para Educação, Ciência e Cultura, comumente abreviada para Organização dos Estados Ibero-americanos (OEI), que representa número expressivo nas avaliações do programa. No Pisa 2012, eram dez países ibero-americanos, dentre um total de 65 participantes; em 2015, eram 11, dentre um total de 70; e em 2018, eram 12, dentre um total de 79. O cenário exposto possibilita a pergunta central: qual o desempenho dos países da OEI nas últimas edições do Pisa?

A elaboração de evidências científicas no tocante a essa questão se faz relevante para o campo educacional pela produção e disponibilização de informações que permitam cotejar o desempenho de um grupo representativo de participantes – em sua maioria países em desenvolvimento – por um período de tempo maior, lançando assim as bases para um estudo longitudinal. Também é relevante por apresentar um cenário ampliado da participação de nações que integram a mesma organização, comungam das mesmas finalidades, bem como por oferecer aos responsáveis pelas políticas educacionais casos de sucesso, estagnação e insucesso de seus pares no Pisa, haja vista a constatada dificuldade encontrada por muitos países no cumprimento de metas educacionais (Organização das Nações Unidas para a Educação, a Ciência e a Cultura, 2019b).

2. Método

Analisar os indicadores de desempenho Pisa dos países integrantes da OEI no período de 2012 a 2018, com vistas a disponibilizar informações que expressem a evolução histórica dos indicadores desses países e subsidiem os responsáveis por políticas educacionais nas tomadas de decisões implica em delimitações necessárias, entre elas o que é o Pisa, o que é a OEI e por que utilizar os dados das edições de 2012, 2015 e 2018.

Segundo o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), autarquia federal vinculada ao Ministério da Educação (MEC) e responsável pelas avaliações, exames e indicadores da educação básica e superior e pelas ações internacionais relacionadas à educação no Brasil, o Programa Internacional de Avaliação de Estudantes (Pisa) é

um estudo comparativo internacional, realizado a cada três anos pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE). O Pisa oferece informações sobre o desempenho dos estudantes na faixa etária dos 15 anos, vinculando dados sobre seus backgrounds e suas atitudes em relação à aprendizagem e também aos principais fatores que moldam sua aprendizagem, dentro e fora da escola. Desde sua primeira edição, em 2000, o número de países e economias participantes tem aumentado a cada ciclo (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020b).

Quanto à sua utilização na definição de políticas educacionais, o Inep destaca a centralidade das informações disponíveis nos relatórios Pisa:

os resultados do Pisa permitem que cada país avalie os conhecimentos e as habilidades dos seus estudantes em comparação com os de outros países, aprenda com as políticas e práticas aplicadas em outros lugares, bem como formule suas políticas e programas educacionais, visando melhorias na qualidade e na equidade dos resultados de aprendizagem (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020b).

A Organização de Estados Ibero-americanos para a Educação, a Ciência e a Cultura (OEI) inicia suas atividades em 1949 para a cooperação entre os países ibero-americanos no campo da educação, da ciência, da tecnologia e da cultura, e atualmente conta com 22 países: Argentina, Bolívia, Brasil, Colômbia, Costa Rica, Cuba, Chile, República Dominicana, Equador, El Salvador, Espanha, Guatemala, Guiné Equatorial, Honduras, México, Nicarágua, Panamá, Paraguai, Peru, Portugal, Uruguai e Venezuela (Organização dos Estados Ibero-americanos para a Educação, a Ciência e a Cultura, 2020). A OEI elenca nove “fins e objetivos” para sua atuação, um dos quais apresenta estreita relação com a participação, análise detalhada e importância da utilização do Pisa.

Colaborar permanentemente na transmissão e no intercâmbio das experiências de integração econômica, política e cultural produzidas nos países europeus e latino-americanos, que constituem as duas áreas de influência da Organização, assim como em qualquer outro aspecto susceptível de servir para o desenvolvimento dos países (Organização de Estados Ibero-americanos para a Educação, a Ciência e a Cultura, 2020).

De base documental, e dadas as limitações deste artigo, o campo empírico estabelecido são os indicadores de desempenho Pisa 2012, 2015 e 2018, nos três domínios – Leitura, Ciências e Matemática – avaliados dos países integrantes da OEI. As informações foram obtidas no endereço eletrônico oficial da OCDE (Organização para a Cooperação e Desenvolvimento Econômico, 2020a), registradas em planilhas por meio do programa Excel.

A organização dos registros se deu em dois momentos: no primeiro momento foram reunidos em três quadros – um para cada domínio avaliado nas três edições – o país líder geral do ranking Pisa, o último país cuja pontuação fosse maior ou igual a 500 pontos, seguido pelos países ibero-americanos ordenados da maior para a menor pontuação obtida. O estabelecimento de 500 pontos de desempenho como referência para as análises se deu por três justificativas: 1) a reconhecida importância da leitura com autonomia na atual sociedade; 2) a pontuação mínima de 480 pontos no domínio leitura estabelece que “o leitor localize e, em alguns casos, reconheça relações entre informações em diversos fragmentos que atendam a múltiplas condições” (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2013a, p.5); e 3) nenhum país ibero-americano avaliado nas três edições e nos três domínios avaliados, salvo Portugal em 2015, no domínio Ciências, alcançou essa pontuação.

No segundo momento foi reunido em quadro único – para os três domínios avaliados nas edições 2012, 2015 e 2018 – a pontuação do país líder geral do ranking, a pontuação dos países ibero-americanos avaliados e a diferença observada entre a pontuação do líder geral do ranking para o país europeu e integrante da OEI com maior pontuação, bem como a diferença observada entre a pontuação do líder geral do ranking para o país não europeu e integrante da OEI com maior pontuação.

72

As análises se basearam nas considerações de Gorur (2016), no sentido de oferecer, ainda que pautadas apenas em indicadores estatísticos, possibilidades de interpretações secundárias a partir de informações e considerações oficialmente disponibilizadas pela Organização das Nações Unidas para a Educação, a Ciência e a Cultura (2019a).

3. Resultados

A coleta e organização das informações possibilitou a confecção dos quadros de 1 a 5. É importante destacar que, após cada quadro, são apresentadas algumas análises comparativas quanto à pontuação e classificação de 2012 em relação a 2015; de 2015 em relação a 2018; e de 2012 em relação a 2018, seguidas de algumas considerações sobre alguns participantes cujos indicadores se destacaram. No campo “País” utilizou-se na primeira linha o termo “Líder”, e na segunda linha a expressão “Com 500 pontos ou mais” para não identificação do país em questão e pelo fato de que os países da OEI, objeto deste estudo, não ocupam nem a liderança nem apresentam, salvo Portugal em uma oportunidade, pontuação igual ou superior a 500.

3.1 Domínio leitura

“Atividades neste nível requerem que o leitor localize e, em alguns casos, reconheça relações entre informações em diversos fragmentos que atendam a múltiplas condições” (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2013a, p.5). A pontuação mínima que contemplava a realização das atividades descritas no excerto correspondia a 480 pontos. O Quadro 1 destaca o desempenho dos países ibero-americanos.

Quadro 1. Desempenho e posicionamento no domínio Leitura.

2012 (65 avaliados)			2015 (70 avaliados)			2018 (79 avaliados)		
País	Pontos	Posição	País	Pontos	Posição	País	Pontos	Posição
Líder	570	1	Líder	535	1	Líder	555	1
Com 500 pontos ou +	504	22	Com 500 pontos ou +	500	18	Com 500 pontos ou +	501	19
Espanha	488	30	Portugal	498	22	Portugal	492	25
Portugal	488	33	Espanha	496	25	Chile	452	44
Costa Rica	441	49	Argentina	475	38	Uruguai	427	49
Chile	441	47	Chile	459	42	Costa Rica	426	50
México	424	52	Uruguai	437	46	México	420	54
Uruguai	411	54	Costa Rica	427	53	Brasil	413	58
Brasil	410	55	Colômbia	425	54	Colômbia	412	59
Colômbia	403	57	México	423	55	Argentina	402	64
Argentina	396	60	Brasil	407	59	Peru	401	65
Peru	384	65	Peru	398	63	Panamá	377	72
Panamá	Não participou	Não participou	República Dominicana	358	66	República Dominicana	342	77
República Dominicana	Não participou	Não participou	Panamá	Não participou	Não participou	Espanha	Não publicado	Não publicado

Fonte: Elaborado pelo pesquisador a partir de Organização das Nações Unidas para a Educação, a Ciência e a Cultura (2020b, 2020c, 2020d).

Na edição de 2012, os países ibero-americanos apresentaram indicadores bem inferiores se comparados ao país líder, haja vista que o ibero-americano melhor posicionado ocupa apenas a 30ª posição entre 65 avaliados, tendo outro ibero-americano ocupando a última posição. Em 2015, os ibero-americanos apresentaram pontuação maior que em 2012, contrariando o desempenho do país líder do ranking, que apresentou uma pontuação menor. A edição de 2018, por sua vez, apresenta um cenário diferente, em que o país líder do ranking apresenta uma evolução em sua pontuação se comparado à edição de 2015, ao passo que os ibero-americanos um retrocesso. A pontuação de 2018 dos ibero-americanos, se comparada à pontuação de 2012, apresenta pouca evolução. Importante destacar os seguintes pontos de alguns países:

- Portugal e Espanha ocupam as primeiras posições entre os ibero-americanos. A Espanha não teve publicado seu desempenho no

domínio Leitura no Pisa 2018 (Organização das Nações Unidas para a Educação, a Ciência e a Cultura, 2020e);

- Chile e Uruguai apresentam certa constância e evolução em seus indicadores;
- Argentina tem grande oscilação em seus indicadores;
- Peru apresenta constância de indicadores próximos às últimas posições;
- República Dominicana e Panamá, mesmo iniciando suas participações em 2015 e 2018, ocupam as últimas posições do ranking.

3.2 Domínio Ciências

“Os estudantes são capazes de identificar questões científicas claramente definidas em uma série de contextos. Podem selecionar fatos e conhecimentos para explicar fenômenos e aplicar modelos simples e estratégias de pesquisa” (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2013b, p.5). A pontuação mínima que contemplava a realização das atividades descritas no excerto correspondia a 484,1 pontos. O Quadro 2 destaca o desempenho dos países ibero-americanos.

74

Quadro 2. Desempenho e posicionamento no domínio Ciências.

2012 (65 avaliados)			2015 (70 avaliados)			2018 (79 avaliados)		
País	Pontos	Posição	País	Pontos	Posição	País	Pontos	Posição
Líder	580	1	Líder	556	1	Líder	590	1
Com 500 pontos ou +	502	25	Com 500 pontos ou +	501	22	Com 500 pontos ou +	502	19
Espanha	496	29	Portugal	501	23	Portugal	492	27
Portugal	489	36	Espanha	493	29	Espanha	483	31
Chile	445	44	Argentina	475	38	Chile	444	46
Costa Rica	429	51	Chile	447	44	Uruguai	426	55
Uruguai	416	54	Uruguai	435	48	México	419	58
México	415	55	Costa Rica	420	55	Costa Rica	416	61
Argentina	406	58	México	416	57	Colômbia	413	63
Brasil	405	59	Colômbia	416	58	Argentina	404	66
Colômbia	399	60	Brasil	401	63	Brasil	404	65
Peru	373	65	Peru	397	64	Peru	404	67
Panamá	Não participou	Não participou	República Dominicana	332	70	Panamá	365	76
República Dominicana	Não participou	Não participou	Panamá	Não participou	Não participou	República Dominicana	336	79

Fonte: Elaborado pelo pesquisador a partir de Organização para a Cooperação e Desenvolvimento Econômico (2020b, 2020c, 2020d).

Na edição de 2012, os países ibero-americanos apresentaram indicadores bem inferiores se comparados ao país líder, haja vista que o ibero-americano melhor posicionado ocupa apenas a 29ª posição entre 70 avaliados, com outro ibero-americano ocupando a última posição. Na edição de 2015, os ibero-americanos, com exceção de Costa Rica e Brasil, apresentaram pontuação maior em relação a 2012, contrariando o desempenho do país líder do ranking, que apresentou uma pontuação menor. A edição de 2018, por sua vez, apresenta um cenário diferente, em que o país líder do ranking apresenta uma evolução em sua pontuação se comparado às edições anteriores, ao passo que os ibero-americanos um retrocesso. A pontuação de 2018, se comparada à pontuação de 2012, apresenta pouca evolução. Importante destacar os seguintes pontos de alguns países:

- Portugal e Espanha ocupam as primeiras posições entre os ibero-americanos;
- Chile e Uruguai apresentam certa constância e evolução em seus indicadores;
- Argentina tem grande oscilação em seus indicadores;
- Peru apresenta constância de indicadores ocupando as últimas posições ou próximos a elas;
- República Dominicana e Panamá, mesmo iniciando suas participações em 2015 e 2018, ocupam as últimas posições do *ranking*.

75

3.3 Domínio Matemática

“Os estudantes são capazes de executar procedimentos descritos com clareza, inclusive aqueles que exigem decisões sequenciais” (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2013c, p.3). A pontuação mínima que contemplava a realização das atividades descritas no excerto correspondia a 482,4 pontos. O Quadro 3 destaca o desempenho dos países ibero-americanos.

Quadro 3. Desempenho e posicionamento no domínio Matemática.

2012 (65 avaliados)			2015 (70 avaliados)			2018 (79 avaliados)		
País	Pontos	Posição	País	Pontos	Posição	País	Pontos	Posição
Líder	613	1	Líder	564	1	Líder	591	1
Com 500 pontos ou +	500	23	Com 500 pontos ou +	502	19	Com 500 pontos ou +	500	21
Portugal	487	31	Portugal	492	28	Portugal	492	29
Espanha	484	33	Espanha	486	32	Espanha	481	37
Chile	423	51	Argentina	456	42	Uruguai	418	59
México	413	53	Chile	423	48	Chile	417	60

2012 (65 avaliados)			2015 (70 avaliados)			2018 (79 avaliados)		
País	Pontos	Posição	País	Pontos	Posição	País	Pontos	Posição
Uruguai	409	55	Uruguai	418	51	México	409	62
Costa Rica	407	56	México	408	56	Costa Rica	402	64
Brasil	391	58	Costa Rica	400	59	Peru	400	66
Argentina	388	59	Colômbia	390	61	Colômbia	391	70
Colômbia	376	62	Peru	387	62	Brasil	384	71
Peru	368	65	Brasil	377	65	Argentina	379	72
Panamá	Não participou	Não participou	República Dominicana	328	70	Panamá	353	77
República Dominicana	Não participou	Não participou	Panamá	Não participou	Não participou	República Dominicana	325	79

Fonte: Elaborado pelo pesquisador a partir de Organização para a Cooperação e Desenvolvimento Econômico (2020b, 2020c, 2020d).

Na edição de 2012, os países ibero-americanos apresentaram indicadores bem inferiores se comparados ao país líder, haja vista que o ibero-americano melhor posicionado ocupa apenas a 31ª posição entre 79 avaliados, com outro ibero-americano ocupando a última posição. Na edição de 2015, os ibero-americanos, com exceção de Costa Rica, México e Brasil, apresentaram pontuação maior em relação a 2012, contrariando o desempenho do país líder do ranking, que apresentou uma pontuação menor. A edição de 2018, por sua vez, apresenta um cenário diferente, em que o país líder do ranking apresenta uma evolução em sua pontuação se comparado à edição de 2015, ao passo que os ibero-americanos um retrocesso. A pontuação de 2018, se comparada à pontuação de 2012, apresenta pouca evolução. Importante destacar os seguintes pontos de alguns países:

- Portugal e Espanha ocupam as primeiras posições entre os ibero-americanos;
- Chile e Uruguai apresentam certa constância e evolução em seus indicadores;
- Argentina tem grande oscilação em seus indicadores;
- Peru apresenta constância de indicadores ocupando as últimas posições ou próximos a elas;
- República Dominicana e Panamá, mesmo iniciando suas participações em 2015 e 2018, ocupam as últimas posições do *ranking*.

Os excertos utilizados para identificar o que se espera dos estudantes classificados com desempenho imediatamente inferior a 500 pontos especificam a necessidade de localizar informações (no domínio Leitura), identificar questões claramente definidas (no domínio Ciências) ou executar

procedimentos descritos com clareza (no domínio Matemática), sugerindo que tais atividades apresentam baixa complexidade. Entretanto, salvo Portugal e Espanha, os ibero-americanos encontram-se distantes desse desempenho.

O desempenho observado dos países ibero-americanos coloca em risco o cumprimento, em 2030, da meta 1 dos Objetivos de Desenvolvimento Sustentável (ODS) relacionados à aprendizagem, isto é, “garantir que todas as meninas e meninos completem o ensino primário e secundário livre, equitativo e de qualidade, que conduza a resultados de aprendizagem relevantes e eficazes” (Organização das Nações Unidas para a Educação, a Ciência e a Cultura, 2019, p.24).

Após a análise dos Quadros 1, 2 e 3, é possível inferir que, independente da edição, do domínio avaliado ou da pontuação obtida, alguns países se constituem em interessantes objetos de investigação mais adensada, uma vez que: ocupam e se alternam nas duas primeiras posições entre os ibero-americanos (Portugal e Espanha); apresentam constância e evolução de desempenho (Chile e Uruguai); apresentam grande oscilação na pontuação (Argentina); apresentam constância de indicadores ocupando as últimas posições ou próximos a elas (Peru); iniciam suas participações no Pisa em 2015 e 2018, contudo se mantêm nas últimas posições do ranking (República Dominicana e Panamá); flutuam entre as 15 últimas posições (demais países). O Quadro 4 ilustra o cenário obtido a partir dessas informações.

Quadro 4. Destaques dos países ibero-americanos no Pisa (2012, 2015 e 2018).

Domínio	PISA	Ocupam as duas primeiras posições entre todos os países da OEI	Apresentam constância e evolução no desempenho	Apresenta grandes oscilações no desempenho	Ocupam as últimas posições entre todos os avaliados pelo PISA
Leitura	2012	Portugal e Espanha	Chile e Uruguai	Argentina	Peru e Argentina
	2015	Portugal e Espanha	Chile e Uruguai	Argentina	Peru e República Dominicana
	2018	Portugal e Chile	Chile e Uruguai	Argentina	Panamá e República Dominicana
Ciências	2012	Portugal e Espanha	Chile e Uruguai	Argentina	Peru e Colômbia
	2015	Portugal e Espanha	Chile e Uruguai	Argentina	Peru e República Dominicana
	2018	Portugal e Chile	Chile e Uruguai	Argentina	Panamá e República Dominicana
Matemática	2012	Portugal e Espanha	Chile e Uruguai	Argentina	Peru e Colômbia
	2015	Portugal e Espanha	Chile e Uruguai	Argentina	Brasil e República Dominicana
	2018	Portugal e Chile	Chile e Uruguai	Argentina	Panamá e República Dominicana

Fonte: Elaborado pelo pesquisador.

No que se refere ao desempenho – expresso por meio da pontuação obtida – dos países ibero-americanos nos três domínios avaliados nas edições de 2012, 2015 e 2018, o Quadro 5 apresenta, com base em quatro agrupamentos, informações que possibilitam algumas inferências.

O agrupamento “A” destaca o desempenho dos países da OEI pertencentes à península Ibérica e a diferença desse desempenho para o país líder geral do Pisa; o agrupamento “B” apresenta o desempenho do país da OEI e não integrante da península Ibérica com melhor desempenho e a diferença desse desempenho para o país líder geral do Pisa; o agrupamento “C” destaca a diferença de desempenho do país com melhor desempenho integrante da península Ibérica e não integrante da península Ibérica em relação ao país líder geral do Pisa; e o agrupamento “D” apresenta a diferença de desempenho observada entre o país da OEI integrante da península Ibérica com melhor desempenho e o país da OEI não integrante da península Ibérica com melhor desempenho.

Quadro 5. Diferença de desempenho (pontuação) entre os participantes.

País		Leitura			Ciências			Matemática		
		2012	2015	2018	2012	2015	2018	2012	2015	2018
A	Líder	570	535	555	580	556	590	613	564	591
	1º da OEI (Europa)	488	498	492	496	501	492	487	492	492
	2º da OEI (Europa)	488	496	Não publicado	489	493	483	484	486	481
	Líder para o 1º da OEI (Europa)	82	37	62	84	55	98	126	72	99
B	1º da OEI (Demais)	441	475	452	445	475	444	423	456	418
	Líder para o 1º da OEI (Demais)	129	68	103	135	81	146	190	108	173
C	Líder para o 1º da OEI (Europa)	82	37	62	84	55	98	126	72	99
	Líder para o 1º da OEI (Demais)	129	68	103	135	81	146	190	108	173
D	1º da OEI (Europa) para o 1º da OEI (Demais)	47	23	40	51	36	48	64	36	74

Fonte: Elaborado pelo pesquisador a partir de Organização para a Cooperação e Desenvolvimento Econômico (2020b, 2020c, 2020d).

A considerar a pontuação obtida pelos países ibero-americanos, é possível observar, nos três domínios avaliados e nas três edições, um distanciamento grande entre o líder do ranking e os ibero-americanos da península Ibérica, ainda mais ampliado se comparado com os demais ibero-americanos.

Observa-se também uma redução desse distanciamento na edição de 2015 em comparação à de 2012, todavia esse distanciamento é ampliado na edição de 2018 se comparado à de 2015.

As informações do Quadro 5 permitem ainda afirmar que o desempenho dos países da OEI nos três domínios avaliados nas três edições configura dois grupos: o primeiro formado por Portugal e Espanha, que tem se mantido próximo a 500 pontos, e o segundo formado pelos demais países da OEI, cujo expoente, em cada domínio avaliado e em cada edição, busca alcançar 450 pontos, o que resulta em um cenário de poucos e mínimos avanços por parte dos países que compõem a OEI num período de seis anos.

4. Discussão

O objetivo central deste artigo é o de disponibilizar um conjunto de informações referentes ao desempenho dos países da OEI nas edições de 2012, 2015 e 2018 do Pisa, de modo a subsidiar discussões e debates dos responsáveis pelas políticas educacionais desses países. É possível afirmar que os resultados obtidos apontam que:

- o agrupamento de informações de mais de uma edição do Pisa possibilita a identificação de países a serem investigados como estudos de caso de sucesso, de estagnação ou de fracasso quanto ao desempenho. A construção de uma série histórica ampliada do desempenho do conjunto dos países ibero-americanos no Pisa pode contribuir ainda mais com as tomadas de decisão por parte dos responsáveis pelas políticas educacionais, uma vez que torna possível o estabelecimento de metas a partir de experiências de seus pares;
- os países ibero-americanos apresentam desempenho que os classifica muito abaixo de países que obtêm 500 pontos ou mais;
- o número de países integrantes da OEI no Pisa foi ampliado de 10 em 2012 para 12 em 2018, contudo ainda distante de seus 22 integrantes (OEI, 2020);
- Portugal, Espanha, Chile e Uruguai apresentam pouca oscilação quanto a seus desempenhos nos três domínios avaliados nas três edições, ao passo que os demais países ibero-americanos apresentam grandes oscilações e flutuam entre as 20 últimas posições do *ranking*;

República Dominicana e Panamá iniciaram suas participações em 2015 e 2018, contudo seus indicadores os posicionam nas últimas posições do *ranking*.

Os apontamentos aqui realizados possibilitam a identificação de um cenário no mínimo preocupante, uma vez que, salvo poucas exceções, independente do domínio avaliado ou da edição Pisa em análise, os países ibero-americanos destaques em cada categoria não se alteram, do mesmo modo que não se altera o distanciamento de desempenho observado entre os países ibero-americanos e os demais avaliados, assim como entre os ibero-americanos da península Ibérica e os demais ibero-americanos.

A comparação, a aproximação, o cotejamento ou qualquer outro termo utilizado que permita, a partir do simples ranqueamento do desempenho Pisa, a elaboração do cenário aqui exposto evidencia a urgente necessidade de os países ibero-americanos repensarem suas políticas educacionais, uma vez que seus indicadores apontam para pouca ou nenhuma efetividade.

A Organização das Nações Unidas para a Educação, a Ciência e a Cultura (2019) enfatiza quanto aos riscos da subutilização e/ou uso inadequado dos dados obtidos com as avaliações de aprendizagem, todavia ressalta a necessidade de constantes, diferentes e adensadas análises desses indicadores, uma vez que a base de dados disponível para essas análises está posta e dispensa a necessidade de se aguardar os resultados de uma próxima edição Pisa. Ressalta também que a análise adensada dos resultados das avaliações em larga escala possibilita a melhor compreensão da política educacional vigente, entretanto sua principal contribuição está na disponibilização de informações que subsidiaram os ajustes necessários às políticas educacionais em curso ou futuras.

5. Conclusões

O desempenho da maior parte dos países ibero-americanos nas edições 2012, 2015 e 2018 do Pisa evidencia avanços mínimos nos três domínios avaliados, tanto que alguns países apresentam em 2018 desempenho inferior ao apresentado em 2012 em mais de um domínio avaliado, o que corresponde a um alerta para investigações aprofundadas, uma vez que compreende um período de seis anos sem avanço ou com retrocesso no desempenho – o que, por sua vez, evidencia ações ou políticas educacionais ineficientes. Os casos de Panamá e República Dominicana, que iniciam suas

participações no Pisa anos após os demais países da OEI e apresentam desempenhos muito abaixo dos demais países da OEI avaliados, colocam em xeque a integração e cooperação entre os países que integram esta organização.

Os países da OEI avaliados pelo Pisa nas edições de 2012, 2015 e 2018 podem ser classificados como pertencentes à metade inferior do total de países avaliados nos três domínios, e cujos indicadores não sugerem alterações positivas nesse cenário. O que se observa é a ampliação do distanciamento entre os ibero-americanos e os países que obtêm desempenho igual ou superior a 500 pontos. Mesmo dentro do grupo, tal cenário guardado as proporções, se repete, pois os países da península Ibérica (Portugal e Espanha) apresentam desempenho significativamente superior em relação aos demais ibero-americanos avaliados.

Mais do que identificar informações importantes como o crescente distanciamento de desempenho entre outros países e os países da OEI avaliados ou o distanciamento de desempenho dentro da própria organização, torna-se fundamental a produção e disseminação de investigações que tenham como objeto de estudo não apenas os indicadores de desempenho desses países, mas sim – e com maior urgência – a investigação das condições objetivas em que esses indicadores são obtidos. Certamente o potencial econômico de alguns integrantes da OEI lhes permita a realização de alguns estudos, contudo é possível que outros não possuam essa condição, o que remete à centralidade da instituição OEI na inclusão dessa pauta como uma de suas prioridades.

Referencias

- Bacchetto, J. G. (2016). O Pisa e o custo da repetência no Fundeb. *Ensaio: Avaliação e Políticas Públicas em Educação*, 24(91), 424-444. Recuperado de <https://bit.ly/2F9p4Ds>.
- Bauer, A. (2019). Avaliação de redes de ensino e gestão educacional: aportes teóricos. *Educação e Realidade*, 44(1), 1-28. Recuperado de <https://bit.ly/2HWqZfr>.
- Brasil (2014). Lei nº 13.005, de 25 de junho de 2014. Aprova o Plano Nacional de Educação – PNE e dá outras providências. *Diário Oficial da União*. Recuperado de <https://bit.ly/3liYcQN>.
- Carvalho, L. M. (2016). Intensificação e sofisticação dos processos da regulação transacional da educação: o caso do Programa Internacional de Avaliação de Estudantes. *Educação & Sociedade*, 37(136), 669-683. Recuperado de <https://bit.ly/2SumzOS>.
- Cosso, D. C.M. (2013). *A organização do trabalho pedagógico em uma escola pública com alto IDESP*. [dissertação de mestrado. Universidade Estadual de Campinas – UNICAMP]. Recuperado de <https://bit.ly/3dfcdMJ>.

- Franco, F. M. (2018). *A qualidade da educação pública como um direito: do estabelecimento à sua efetivação*. [dissertação de mestrado, Universidade de Araraquara – UNIARA]. Recuperado de <https://bit.ly/3ligOvf>.
- Gorur, R. (2016). As “descrições finas” das análises secundárias do PISA. *Educação & Sociedade*, 37(163), 647-668. Recuperado de <https://bit.ly/3d7yjAr>.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2013a, março). *Programa Internacional de Avaliação de Estudantes (Pisa)*. Recuperado de <https://bit.ly/36ArhTx>.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2013b, março). *Programa Internacional de Avaliação de Estudantes (Pisa)*. Recuperado de <https://bit.ly/3lig521>.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2013c, março). *Programa Internacional de Avaliação de Estudantes (Pisa)*. Recuperado de <https://bit.ly/2HRM9vd>.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2020a, fevereiro). *Estudos Regionais Comparativos e Explicativos (Erce)*. Recuperado de <https://bit.ly/3ImN4Cf>.
- Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. (2020b, fevereiro). *Programa Internacional de Avaliação de Estudantes (Pisa)*. Recuperado de <https://bit.ly/3iHqCCs>.
- Klein, R. (2011). Uma re-análise dos resultados do PISA: problemas de comparabilidade. *Ensaio: Avaliação e Políticas Públicas em Educação*, 19(73), 717-768. Recuperado de <https://bit.ly/2GtGFGM>.
- Lingard, B. (2016). PISA: fundamentos para participar e acolhimento político. *Educação & Sociedade*, 37(136), 609-627. Recuperado de <https://bit.ly/3nkusVr>.
- Organização das Nações Unidas para a Educação, a Ciência e a Cultura. (2019a). *A promessa das avaliações de aprendizagem em larga escala: reconhecer os limites para desbloquear as oportunidades*. Recuperado de <https://bit.ly/36B3WBe>.
- Organização das Nações Unidas para a Educação, a Ciência e a Cultura (2019b). *UNESCO alerta que países estão fora do caminho para atingir metas de educação até 2030*. Recuperado de <https://bit.ly/2GCoUVs>.
- Organização de Estados Ibero-americanos para a Educação, a Ciência e a Cultura. (2020). *O que é a OEI?*. Recuperado de <https://bit.ly/3iCEyNTi>.
- Organização para a Cooperação e Desenvolvimento Econômico. (2020a, março). *Programa Internacional de Avaliação de Estudantes*. Recuperado de <https://www.oecd.org/pisa/>.
- Organização para a Cooperação e Desenvolvimento Econômico. (2020b, março). *Programa Internacional de Avaliação de Estudantes*. Recuperado de <https://bit.ly/36liglp>.
- Organização para a Cooperação e Desenvolvimento Econômico. (2020c, março). *Programa Internacional de Avaliação de Estudantes*. Recuperado de <https://bit.ly/3jAK05l>.
- Organização para a Cooperação e Desenvolvimento Econômico. (2020d, março). *Programa Internacional de Avaliação de Estudantes*. Recuperado de <https://bit.ly/3iA1fSV>.
- Organização para a Cooperação e Desenvolvimento Econômico. (2020e, março). *Programa Internacional de Avaliação de Estudantes*. Recuperado de <https://bit.ly/2GmbIKc>.

- Pinto, J.; Silva, J. C. e Bixirão Neto, T. (2016). Fatores influenciadores dos resultados de matemática de estudantes portugueses e brasileiros no PISA: revisão integrativa. *Ciência & Educação (Bauru)*, 22(4), 837-853. Recuperado de <https://bit.ly/3IjBIil>.
- Pizarro, M. V. e Lopes Junior, J. (2017). Os sistemas de avaliação em larga escala e seus resultados: o PISA e suas possíveis implicações para o ensino de Ciências. *Ensaio Pesquisa em Educação em Ciências (Belo Horizonte)*, 19, 1-24. Recuperado de <https://bit.ly/36FtU6C>.
- Salokangas, M e Kauko, J. (2015). Tomar de empréstimo o sucesso finlandês no PISA? Algumas reflexões críticas, da perspectiva de quem faz este empréstimo. *Educação e Pesquisa*, 41 (esp.), 1353-1366. Recuperado de <https://bit.ly/2GuITpl>.
- Vaillant, D e Zidan, E. R. (2016). Práticas de liderança para a aprendizagem na América Latina: uma análise do PISA 2012. *Ensaio: Avaliação e Políticas em Educação*, 24(91), 253-274. Recuperado de <https://bit.ly/3nllrLV>.
- Vicino, M. B. (2013). *Programa Bonificação por Resultados: opiniões de Professores Coordenadores da rede de ensino do Estado de São Paulo*. [Dissertação de Mestrado, Universidade Cidade de São Paulo – UNICID]. Recuperado de <https://bit.ly/3d2RIYM>.
- Villani, M. e Oliveira, D. A. (2018). Avaliação Nacional e internacional no Brasil: os vínculos entre o Pisa e o IDEB. *Educação & Realidade*, 43(4), 1343-1362. Recuperado de <https://bit.ly/3nm9ENc>.

Exploring trends in the relationship between child labour, gender and educational achievement in Latin America

Abigail Middel¹ ; Kalyan Kumar Kameshwara¹ ; Andrés Sandoval-Hernandez¹ 

¹ University of Bath, United Kingdom

Abstract. Participation in child labour, in both household and non-household activities, gender effects and low educational attainment remain challenges for countries in Latin America. Through hierarchical linear modelling of data from the OECD's Programme for International Student Assessment (PISA), this study seeks to explore the current cross-country trends in the relationship between educational attainment, child labour and gender. While non-household labour is found to have an effect, as per statistical significance and the magnitude, on educational achievement across all Latin American countries; participation in household labour is significant in only two countries (Peru and Uruguay). Girls are found to underperform compared to boys by a significant margin across Latin America. The later part of the study seeks to examine the interaction effects of gender and participation in labour activities. Results show that gender has no moderating effect, suggesting that the participation in work itself or workspace (household or non-household) does not influence or contribute to gender inequality in education outcomes. The explanatory factors for gender inequality in education outcomes is potentially rooted in a different sphere of influence which needs to be deciphered through deeper empirical investigation.

Keywords: child labour; gender; inequality; educational achievement; Latin America

Análisis de tendencias en la relación entre el trabajo infantil, el género y los logros académicos en Latinoamérica

Resumen. La participación de menores, tanto en tareas domésticas como no domésticas, los efectos del género y los bajos logros académicos siguen siendo un reto para los países de América Latina. A través del modelaje lineal jerárquico de datos del Programa Internacional de Evaluación de los Alumnos (PISA), este estudio busca explorar las tendencias entre los países en la relación entre los logros académicos, el trabajo infantil y el género. Si bien el trabajo fuera del hogar suele tener un efecto sobre los logros académicos en todos los países de Latinoamérica, tal como demuestran la importancia y la magnitud de las estadísticas; la participación en las tareas del hogar es relevante únicamente en dos (Perú y Uruguay). Se ha visto que las niñas obtienen peores resultados que los niños en un margen importante en toda Latinoamérica. La última parte del estudio busca analizar los efectos de interacción de género y participación en actividades laborales. Los resultados demuestran que el género no es un factor moderador, lo que sugiere que la participación en el trabajo o en el lugar de trabajo en sí mismo (en el hogar o fuera de él) no influye ni contribuye a la desigualdad de géneros en los resultados académicos. Los factores que explican la desigualdad en los resultados académicos se encuentran posiblemente en una esfera de influencia distinta que debe descifrarse mediante una investigación empírica más profunda.

Palabras clave. trabajo infantil; género; desigualdad; logros académicos; Latinoamérica.

Análise das tendências na relação entre o trabalho infantil, gênero e desempenho acadêmico na América Latina.

Resumo. A participação de menores, tanto em tarefas domésticas como não domésticas, os efeitos do gênero e o baixo rendimento escolar continuam sendo um desafio para os países da América Latina. Por meio de modelagem linear hierárquica de dados do Programa Internacional de Avaliação dos Alunos (PISA), este estudo busca explorar as tendências entre os países na relação entre desempenho acadêmico, trabalho infantil e gênero. Embora o trabalho fora de casa tenda a afetar o desempenho acadêmico em todos os países da América Latina, como mostra a importância e a magnitude das estatísticas, a participação nas tarefas domésticas é relevante apenas em dois (Peru e Uruguai). Viu-se que as meninas obtêm piores resultados que os meninos por uma margem significativa em toda a América Latina. A última parte do estudo busca analisar os efeitos da interação de gênero e a participação nas atividades de trabalho. Os resultados mostram que o gênero não é um fator moderador, sugerindo que a participação no trabalho ou no próprio local de trabalho (no lar ou fora dele) não influencia nem contribui para a desigualdade de gêneros nos resultados acadêmicos. Os fatores que explicam a desigualdade nos resultados acadêmicos estão possivelmente em uma esfera de influência diferente que deve ser decifrada por meio de uma pesquisa empírica mais profunda.

Palavra-chave: trabalho infantil, gênero; desigualdade; resultados acadêmicos; América Latina.

1. Introduction

Academic achievement is important because it is strongly linked to positive outcomes in many other areas. Adults who report high educational achievement during their school time are more likely to have stable employment, have more employment opportunities, and earn higher salaries than those with lower educational achievement (Card, 1999). They are less likely to engage in criminal activity (Machin et al., 2011; Lochner & Moretti, 2004), be more active as citizens (Lochner, 2011), and to be healthier (Bossuyt et al., 2004; Khang et al., 2004) and happier (Easterlin, 2003).

Factors negatively associated with poor educational outcomes are generally consistent across not just Latin America, but the globe. Lack of parental education is associated with lower attainment and higher dropout rates from high school (Barnard, 2004; Lee & Bowen, 2006) and parental involvement and expectations play an important role (Arends-Kuenning & Duryea, 2006; Driessen et al., 2005; Choi, 2008). Socioeconomic status is strongly associated with academic achievement (Nam & Juang, 2009; Reardon, 2011; Altschul, 2012; Pfeffer, 2018; Ziol-Guest & Lee, 2016) and belonging to an ethnic minority group is also strongly associated with poorer educational outcomes in many countries (Gillborn & Mirza, 2000; Kao & Thompson, 2003; Murillo, 2003; Archer & Francis, 2006). Other factors such as the level of urbanisation is found to be negatively associated with educational outcomes in developing countries for a variety of reasons including the higher cost per student of providing education in rural areas (Behrman et al., 1999; Gould, 2007).

It's been argued that in Latin America, unlike in many developing countries, gender is no longer associated with educational outcomes. Female students are reported to have not faced disadvantage in terms of enrolment across the region for over 30 years (Ahuja & Filmer, 1995) and have even begun to overtake male students, receiving equal or higher grades than males (Grant & Behrman, 2010). Along with factors mentioned above, child labour is found to be a significant predictor of poor educational outcomes including enrolment, attendance, grade repetition and attainment (Montmarquette et al., 2007; Heymann et al., 2013; Psacharopoulos, 1997; Assaad et al., 2010; Putnick & Bornstein, 2016; Eckstein & Wolpin, 1999; Parent, 2006; Gunnarsson et al., 2006). Latin American countries have large numbers of children working with varying legislation (Appendix 1) as well as social programmes in place aimed at reducing this practice.

In terms of enforcement of child labour laws, none of the Latin American countries in this study have a sufficient number of labour inspectors as per technical advice from the International Labour Organisation (ILO)

of a ratio of 1 inspector for every 15,000 workers. In Colombia, Dominican Republic and Peru there are less than half the recommended number of inspectors (ILAB, 2020). Not all countries have data available on the number of inspections conducted and even fewer report the number of child labour violations found and for which penalties were imposed and subsequently collected. In Chile and the Dominican Republic two of the countries for which we have some data, all child labour violations found had penalties imposed; however, in Chile less than half of these penalties were collected. In Colombia of 247 labour violations found in 2017, only 15 had penalties imposed (ILAB, 2020).

All countries have various social programmes aimed at reducing child labour. These include programmes which strengthen the employability of family members of at risk children (Walking Together for the Eradication of Child Labor, Chile), conditional cash transfer programs (More Families in Action, Colombia; Let's Get Ahead Program, Costa Rica; the Together Program, Peru), work and study programmes (I Study and Work, Uruguay), extension of the school day (Dominican Republic), educational and psychological help to at risk families (Carabayllo Project, Peru), targeting children in rural areas (Huánuco Project Peru; Houses of Joy Costa Rica) and education on children's rights (Present Against Child Labor, Colombia) (ILAB, 2020).

One of the challenges in studying child labour is that of defining child labour itself. The difficulty in arriving at a uniform definition is the fact that child labour intersects with local contextual and cultural factors. This complexity intensifies as we approach the adolescent years of the child, when his or her physical capacity to undertake work coincides with the most crucial secondary school years and, in many countries, legal working age. Minimum working age legislation often comes in the form of ratification of the Minimum Age Convention set forth by the ILO, who proffer their own definition of child labour.

The ILO notes that not all work undertaken by children can be classified as "child labour" and define it only as, "work that deprives children of their childhood, their potential and their dignity, and that is harmful to physical and mental development and/or interferes with their schooling." (ILO, 1996). Some kinds of work undertaken by children and adolescents are in fact considered positive. The ILO includes in this category activities such as household chores, helping in a family business or earning an allowance outside of school hours. They note that these activities "contribute to children's development and to the welfare of their families; they provide them with skills and experience and help to prepare them to be productive members of society during their adult life." (ILO, 1996)

Whether specific forms of work outside of the above can be considered 'child labour' is contingent on the age of the child, the nature and hours of the work undertaken and the working conditions and the cultural and legal contexts. This is then contingent on the individual country and the sector the work falls under (ILO, 1996). The ILO specifically considers work child labour if it has an adverse effect on a child's education by "depriving them of the opportunity to attend school; obliging them to leave school prematurely; or requiring them to attempt to combine school attendance with excessively long and heavy work." (ILO, 1996).

The definition of child labour is a complex one, and data available does not always allow operationalisation of it as per ILO guidelines. For the purpose of this analysis, child labour is understood as any labour undertaken by children (who in this study are aged 15) before or after school. Similar to previous studies on child labour in Latin America (Gunnarsson et al., 2006; Psacharopoulos, 1997), this study lacks information on the nature of work undertaken as well as the hours.

There is a potential argument made for a broader definition of child labour which is not tied to the specific nature of the work nor the number of hours. The image of a child working for long hours outside the home that informs legislation and social programmes aimed at its reduction, is not the profile of child labourers in many countries and particularly misrepresents the labour of girls (Assaad et al., 2010). Children are often engaged in work that is not captured by traditional definitions of work (i.e. market work) for example unpaid agricultural work in family enterprises. In doing so, research ignores the potential for responsibilities inside the household to effect educational attainment, in fact any work which interferes with human capital production that would benefit children and society should be considered (Putnick & Bornstein, 2016; Assaad et al., 2010). In addition, this kind of unpaid household work is said to be gendered and traditional definitions of work significantly misrepresent the work undertaken by girls (Putnick & Bornstein, 2016; Levison, 2000). This complexity can only be understood and tackled, if need be, by studying child labour in each context and examining the factors local forces that shape the process.

Therefore, it is important to empirically study how child labour interacts with influential factors such as gender. This paper sets out to achieve just that in the context of Latin America. It seeks to explore the trends and effects in/of participation of child in labour forms categorised broadly as household labour and non-household labour and its effect on education.

2. Literature review

There are various reasons posited for children's' participation in work. Using simultaneous equation models fitted to Indian data, Rosenzweig and Evenson (1977) analysed family decision making regarding fertility and the allocation of children's time to labour and education. They concluded that a high return on raw child labour as opposed to investment in skills acted as a motivation for the creation of large families. In analysis of data from Venezuela, Psacharopoulos (1997) found a similar trend - that the decision to work was associated with a larger family size. However, data from Botswana on the activities of youth led Chernikovsky et al. (1985) to conclude that there is in fact no trade-off between children's schooling and fertility.

Child's gender, familial wealth and composition and rural dwelling are important predictors of child labour. In an analysis of data from Latin America Psacharopoulos (1997) found working children were mostly male, indigenous, and from poorer female-headed households and their earnings contributed a significant amount of household income, amounting to 13% and 27% in Bolivia and Venezuela respectively. Analysis of Canadian data showed that student gender and education of parents most significantly predicted students' preference for labour market participation over schooling. As well as male students and students with parents with no post-secondary education, female students with children and students from single parent families were also overrepresented among those who drop out to pursue work (Montmarquette et al., 2007). In addition, several studies have found that residing in an urban area reduces the probability of children undertaking paid work (Psacharopoulos, 1996; Assaad et al., 2010).

Macro factors like minimum wage and levels of unemployment are strongly associated with child labour. The decision to drop out of school to pursue work is significantly affected by minimum wage (measured in real terms) - when students who are unsure as to whether to finish high school can earn a high minimum wage in the labour market, they tend to conclude there is little to gain from continuing their education (Montmarquette et al., 2007). As well as a high minimum wage, low unemployment rate can also cause students who may not be inclined to drop out, to do so under these particular conditions. Montmarquette et al. (2007) evidence this in their finding that the effect of these macroeconomic variables was stronger for those who express a preference for schooling (and therefore under normal conditions would not be inclined to drop out).

Studies of various Latin American countries in particular have shown that legislation has little effect on children's involvement in the labour force. Psacharopoulos' (1997) analysis revealed significant partici-

pation in the labour market among children who should be prevented from it by compulsory education or working age legislation. Similar evidence was found in Brazil by Bargain and Boutin (2017) and in data from 59 countries including Venezuela, the Dominican Republic and Bolivia by Edmonds and Shrestha (2012). Legislation often fails in eliminating child labour in its entirety as legislation does not cover the entire economy or only applies to specific activities or sectors (Boockmann, 2010). A ban that applies only to certain sectors also leads to a reduction in child wages due to the subsequent excess supply of child labour in sectors not enforcing the ban (Basu & Van, 1998). In addition if productivity and wages are higher in the sector where the legislation is enforced (e.g. industry) it pushes children into low paid low productivity employment such as agriculture. Households rely on this income need to maintain it. This means that more hours would be worked in order for household wealth not to be negatively affected meaning less time allocated to education.

The reduction of income also has a subsequent effect on other children within the family – although substantial working hours represent a clear detriment to a child's education, their earnings (which as previously discussed can be as high as 27% of household income) increase the probability that their siblings will attend school (Basu & Tzannatos, 2003). If wages are depressed, instead of more hours worked by one child, budget-constrained households may send more children to work, and legislation has had the exact opposite of the desired effect. A rise in adult wages is needed to offset that negative effects on household wealth to avoid the above consequences (Boockmann, 2010). This is also the case if the financial consequences of a ban (penalties or bribes) are shouldered by families (Basu & Van, 1998).

2.1 Household Labour

In addition the prohibition of child labour (through minimum legal working age laws) is rarely applied uniformly across all activities. Thus, it may in fact merely lead to the reallocation of child labour into unregulated sectors such as family businesses or work happening inside the household where these laws do not apply or are even more difficult to enforce (Bargain & Boutin, 2017). Using a two-sector model of employment in which legislation completely eliminates child labour, Basu (2005) found that a ban via minimum age legislation in this model instead pushed children into unregulated work. There are three conditions identified by Edmonds and Shrestha (2012) as necessary for sector reallocation to be neutral after a ban in the regulated sector. Firstly, that child and adult labour are exact substitutes subject to a productivity shifter (based on Basu and Van's (1998) 'substitution axiom'). The second condition is 'non-saturation' - that adult and child labour can be

easily substituted between productive tasks within the household. Thirdly is the need for ‘competitive adult labour markets’ – the free movement of adult labour between the household and the labour market, otherwise children’s work may merely be moved inside the home. Despite this, the paid work undertaken by children outside the home has received the most empirical attention (Putnick & Bornstein, 2016) and literature on child labour often ignores the potential for responsibilities inside the household to affect educational attainment.

Levison (2000) further argues that the traditional definition of work used to define child labour distinguishes arbitrarily between activities that are similar. She gives the example of food preparation, which if happening in a market stall or unpaid in a family enterprise is considered work, but the same activity is not considered work if undertaken for the purpose of household consumption. In the context of labour force statistics or national accounts the distinction between household / domestic work and market work is useful but causes biases when trying to understand the effect of child labour on schooling (Levison, 2000). Small jobs or chores may be beneficial for children, but the issue is with all work, whether included in the traditional definition or not, that interferes with children’s education or wellbeing (Assaad et al. 2010).

2.2 Gender

As previously discussed, the definition used when assessing the relationship between children’s work and educational attainment matters greatly, particularly for girls - gender differences in both the incidence and determinants of work are misrepresented when a traditional definition of work (i.e. market work) is used (Levison 2000). Overall, for middle- and low-income countries, there are higher rates of girls’ labour inside the home and higher rates of labour inside the home for boys, which often reflects “macrosystem-level gender inequality” (Putnick & Bornstein, 2016, p5). The high amount of household labour carried out by girls can be attributed to the cultural expectation that girls will be mothers and homemakers and thus early involvement in household work acts as preparation for these adult roles (Assaad et al., 2010).

As well as cultural, the reason is economic - in a number of the low- to middle- income countries studied by Putnick & Bornstein (2016) the levels of education and rates of employment are much lower for women than for men - in the face of these limited economic opportunities for adult women, parents may encourage girls to assume responsibilities within the household to prepare them for their likely adult role as a homemaker. It can

also be considered a rational decision by girls and adolescents themselves, who are aware of their incredibly limited access to the paid labour market (Assaad et al., 2010). In the same vain, boys' work outside the home allows them to develop skills that may apply to their adult work. In countries with better national gender equality, participation of boys and girls in excessive chores is similar (Putnick & Bornstein, 2016).

2.3 Educational outcomes

Studies have examined the effect of hours worked on educational outcomes measured by enrolment, attendance, drop out / graduation rates, grade repetition, years of schooling and in few cases, test results. Household and non-household labour has been found to negatively affect both enrolment and attendance (Beegle et al., 2009; Assaad et al., 2010). In a study of developing countries only, Guarcello et al. (2008) found that working children faced an attendance disadvantage of 10% or above in 56 out of 60 countries, and in 10 of those countries the disadvantage was as high as 30%. Working children also have a disadvantage in total years of schooling compared to non-working children. In Latin American countries, the difference in educational attainment between nonworking and working children begins as young as aged six and then increases rapidly - by the age of 14 (legal working age) Venezuelan working children have a deficiency of 2 years and Bolivian working children 1.4 years (rising to 2.5 by the age of 18) (Psacharopoulos, 1997).

However, measuring educational outcomes in this way has limitations. Attendance, enrolment and even years of schooling do not measure learning. Children might continue to attend school every day due to compulsory schooling legislation however working children may be too fatigued to study effectively in school nor have the time nor energy to study afterwards (Gunnarsson et al., 2006); the effects of which are not captured by these measures.

Various studies have found a strong link between work and high school drop-out rates (Parent, 2006; Montmarquette et al., 2007; Eckstein and Wolpin, 1999). While dropping out of high school to pursue work may be due to economic necessity of the household with child income often accounting for large percentage of a household's income (Psacharopoulos, 1997). Working during high school may cause a child to lag behind in their schoolwork to the point where dropping out in favour of entering the labour market full time is preferable (Eckstein and Wolpin, 1999). However, assessing educational outcomes as completion (graduation) or non-completion (drop-out) of high school does not take into account working children who stay in school and their possible lower achievement and the limit that pla-

ces on future pursuits. It has been noted in several studies including Marsh (1991) and Barone (1993) that young people working more than 20 hours per week while in school were much less likely to pursue higher education.

Educational attainment can be assessed through grade attainment. Although grade failure and repetition are associated with the same causal factors as children's participation in the labour market, working almost doubles the likelihood of failing a grade (Psacharopoulos, 1997). Beegle et al. (2009) found that child labour was significantly associated with a reduction of the highest grade attained. Although grade attainment may be a more nuanced measure than high school completion, the main limitation is the difficulty of cross-country comparison due to substantial difference in schooling across countries including the way in which grade attainment is assessed. In their study of Latin America, Gunnarsson et al. (2006) used results from mathematics and literacy tests and found that children who almost never work had a 22% and 27% advantage in mathematics and language respectively over working children. Although the use of test scores is arguably a better measure of educational outcomes as it measures learning, the use of pooled regression does not sufficiently address the issue of cross-country invariance.

3. Significance of the study

Previous studies examining the effect of child labour have largely focused on school enrolment, attendance, dropouts, grade repetition and aspirations. These can be considered proxies, and proxies for only schooling, not addressing the actual goal of schooling itself – learning. There is also a dearth in exploiting the large-scale cross-country surveys which offer a valuable snapshot on the latest trends in child labour, gender and education.

This study would help bridge the gap in literature by examining the association between child labour, and its interaction with gender, and educational attainment measured via test scores. In addition to that, it gives a more comprehensive picture of the trends in Latin America by giving a comparative analysis involving various countries. There are very few studies focused solely on Latin America, and the majority of those that have done so, or included Latin America in aggregate or pooled data, have used data from more than two decades ago. Latin America has undergone rapid change over the last 20 years, and although inequality persists, it has been declining in many countries (López-Calva & Listig, 2010). However, whilst access to primary and secondary education has been expanded, access to tertiary education, the next stage in Latin America's "educational upgrading" (Grynspan, 2010, p.vii), could be affected due to low educational attainment

in secondary schools, (López-Calva & Listig, 2010), precisely the outcome variable measure in this study, and a crucial factor that studies examining purely schooling would miss.

This study exploits the variance in labour participation levels, and student learning levels, both within and between countries to estimate its association. The study employs multilevel analysis, which accounts for clustering, to identify the contribution (positive or negative) and significance of child labour, and its association with gender, in shaping student test scores. This study aims to add robust empirical evidence to the body of knowledge on the status of child labour participation and its effect on child's education. This evidence is important through its uniqueness - it does not originate from a study focused exclusively on child labour or household survey but from a school survey focusing on education.

4. Research questions

94

This study uses large-scale datasets from OECD's PISA to explore the trends and effects of labour practices among 15-year-old children on their education outcomes in seven Latin American countries. The study exploits PISA 2015 survey data which captures if the student is engaged in household or non-household labour in all of the seven Latin American countries. It seeks to examine the effects of child labour participation (for household and non-household labour separately) on learning achievements in each of the seven countries, after controlling for various individual, household and school effects. Additionally, the study also examines if the participation and effects of participation in labour is moderated by individuals' gender.

1. What are the patterns in child labour participation (in household and non-household labour) across the sample of Latin American countries? How is it distributed along the lines of gender?
2. Does participation in household or/and non-household labour have any significant influence on student's learning levels? If so, what is the magnitude of effect in each country?
3. Does gender moderate the effect of labour participation on student attainment?

5. Data and methods

There were 72 countries in total which participated in PISA in 2015 and eight from Latin America. They include Brazil, Chile, Costa Rica, Colombia, Dominican Republic, Mexico, Peru and Uruguay. PISA follows a complex sampling design strategy. The schools are selected using systematic-random PPS (probability proportional to size) sampling. In the next stage, students who are 15 years old are randomly selected from each school. There is data collected at the student level, teacher level and school level from all the sampled schools and students (OECD, 2017).

Along with collecting information about student and home background, PISA also surveys if students participate in household and non-household labour activities. These are the main variables of interest in this study. They are captured for each student in the sample in all the PISA participating countries from Latin America. Each of the students is administered an assessment to capture his or her learning levels in mathematics, science, reading, financial literacy and collaborative problem solving. Due to the length of the assessment, the student undertakes only a part of the test in each subject and based on their performance, the final scores are imputed as plausible values. The mathematics scores are chosen as proxy for student learning and achievement in this study. With mathematics achievement as the outcome, two-level hierarchical regression models are constructed for each country with student at one level and school at another level.

95

Hierarchical regression analysis is chosen over OLS linear regression model to account for the clustering effects of students in a school (Goldstein 2011). This is due to the correlation between student performances from the same school. Hierarchical models would produce unbiased and robust estimates taking into consideration that student scores in the sample are not independent of each other. The inferences made from the analysis of PISA 2015 can be generalisable to a country level due to the representative sample and the sampling weights included as part of the analysis. In order to account for the imputation uncertainty derived from the PISA complex assessment design, the ten plausible values provided in the data set for the maths scores were used simultaneously in all the analyses (Rutkowski, et al., 2010). Furthermore, following the procedure suggested by Rutkowski and colleagues (2010), each level was weighted separately for all the models.

The first model (M_1) is constructed for each of the seven countries, to examine the effect of participation in labour activities on learning achievement in mathematics:

$$\begin{aligned} \text{Maths Scores } (y_{ij}) = & \beta_0 + \beta_1(\text{Household Labour})_{ij} + \beta_2(\text{Non-household Labour})_{ij} + \\ & \beta_3(\text{Female})_{ij} + \beta_4(\text{Private School})_j + \beta_5(\text{Grade})_{ij} + \beta_6(\text{ESCS})_{ij} \\ & + \beta_7(\text{Motivation levels})_{ij} + \beta_8(\text{Expected Occupational Status})_{ij} \\ & + \beta_9(\text{Math hours})_{ij} + \beta_{10}(\text{ST Ratio})_j + \beta_{11}(\text{School mean ESCS})_j \\ & + e_{ij} + u_j \\ & e_{ij} \sim N(0, \sigma_e^2); u_j \sim N(0, \sigma_u^2) \end{aligned}$$

While the variables of interest are mainly household, non-household labour and gender of the student, it is necessary to control for various student and school-level characteristics. The school-level variables accounted for in the modelling are the type of school (public or private), the Student-Teacher ratio of the school and most importantly the average of social, economic and cultural statuses (ESCS) of all sampled children in the school. The individual level characteristics controlled for in the above and below equations are the grade in which the student is studying, motivational levels, aspiration levels (desired occupation status) and the number of hours dedicated in total to study the subject of assessment (mathematics). These covariates are chosen from literature as they have shown to demonstrate a significant contribution in explaining the variance of the outcome variable (student achievement). The Model (M₁) assists in answering the second research question on the significance and effect of labour participation in student learning.

96

In order, to answer the third research problem of this paper, the following models (M₂) and (M₃) are constructed. In order to examine if gender moderates the effect of labour participation on student learning, an interaction term, between household labour and female dummy, between non-household labour and male dummy, is added to the above model.

$$\begin{aligned} \text{Maths Scores } (y_{ij}) = & \beta_0 + \gamma(\text{Household Labour*Female}) + \beta_1(\text{Household Labour})_{ij} \\ & + \beta_2(\text{Non-household Labour})_{ij} + \beta_3(\text{Female})_{ij} + \beta_4(\text{Private School})_j + \beta_5(\text{Grade})_{ij} + \beta_6(\text{ESCS})_{ij} + \beta_7(\text{Motivation levels})_{ij} \\ & + \beta_8(\text{Expected Occupational Status})_{ij} + \beta_9(\text{Math hours})_{ij} + \beta_{10}(\text{ST Ratio})_j + \beta_{11}(\text{School mean ESCS})_j + e_{ij} + u_j \\ & e_{ij} \sim N(0, \sigma_e^2); u_j \sim N(0, \sigma_u^2) \end{aligned}$$

$$\begin{aligned} \text{Maths Scores } (y_{ij}) = & \beta_0 + \gamma(\text{Non-household Labour*Male}) + \beta_1(\text{Household Labour})_{ij} \\ & + \beta_2(\text{Non-household Labour})_{ij} + \beta_3(\text{Female})_{ij} + \beta_4(\text{Private School})_j + \beta_5(\text{Grade})_{ij} + \beta_6(\text{ESCS})_{ij} + \beta_7(\text{Motivation levels})_{ij} \\ & + \beta_8(\text{Expected Occupational Status})_{ij} + \beta_9(\text{Math hours})_{ij} + \beta_{10}(\text{ST Ratio})_j + \beta_{11}(\text{School mean ESCS})_j + e_{ij} + u_j \\ & e_{ij} \sim N(0, \sigma_e^2); u_j \sim N(0, \sigma_u^2) \end{aligned}$$

Female is included as a dummy in model 2 (M_2) as it is often observed that girls engage more in household labour and it would be plausible to examine if female moderates the effect in the case of household labour participation. Likewise, following the similar reasoning, male is added as a dummy for model 3 (M_3) in the case of non-household labour. These models are similarly constructed for each country with the controls from first model (M_1) intact.

6. Results

The trends and distribution of participation in labour activities, in household and non-household domains, in seven^[1] Latin American countries are demonstrated by the following descriptive statistics in (Table 1) and (Table 2).

Table 1

Country	Household Labour (%)		Non Household Labour (%)	
	No	Yes	No	Yes
Chile	33.56	66.44	75.45	24.55
Colombia	22.98	77.02	55.54	44.46
Costa Rica	34.16	65.84	80.79	19.21
Dom Rep.	15.63	84.37	58.66	41.34
Mexico	17.46	82.54	70.5	29.5
Peru	10.3	89.7	68.24	31.76
Uruguay	20.66	79.34	67.59	32.41

Source: authors' descriptive statistics from PISA 2015.

All the Latin American countries studied here have at least 65% of the sample taking part in household labour and not more than 45% involved in non-household labour. The highest participation in household labour is found in Peru, followed closely by Dominican Republic and Mexico. In the case of non-household labour, Colombia has the highest percentage, immediately followed by Dominican Republic. Costa Rica has the lowest participation of the sample in both household and non-household labour. Not far below Costa Rica, Chile is found to have the second lowest participation levels in both domains. However, it cannot be said if having a higher or lower percentage of children involved in labour, has (or does not have) an effect on one's education. It remains to be seen if labour participation has any actual effect (positive or negative) on student's education and how this could be shaped by different contextual factors.

Table 2

Country	Gender	Household Labour		Non-Household Labour	
		No	Yes	No	Yes
Chile	Male	1,046	2,019	2,129	921
	Female	1,026	2,083	2,502	586
Columbia	Male	1,251	3,906	2,531	2,623
	Female	1,255	4,492	3,514	2,217
Costa Rica	Male	946	1,692	1,976	657
	Female	898	1,862	2,377	378
Dominican Rep	Male	311	1,341	800	836
	Female	234	1,600	1,222	589
Mexico	Male	634	2,634	2,022	1,237
	Female	521	2,827	2,630	710
Peru	Male	317	2,503	1,680	1,134
	Female	240	2,346	1,994	576
Uruguay	Male	474	1,781	1,296	938
	Female	523	2,047	1,922	605

Source: authors' descriptive statistics from PISA 2015.

98

Therefore, the first research question is answered through the descriptive statistics by depicting an overall picture of percentages and frequencies of participation rates in household and non-household labour activities and how they are split based on gender. The results of the hierarchical linear models for each country are represented below. The estimates and their standard errors of variables of interest and other student and school level covariates are listed in Table 3.

Table 3

Student achievement	Chile	Colombia	Costa Rica	Dom Rep.	Mexico	Peru	Uruguay
Household Labour	-1.31 (-3.11)	-4.51 (-2.75)	1.25 (-2.86)	3.18 (-4.32)	1.56 (-3.49)	-8.27** (-4.07)	-10.74*** (-3.69)
Non-Household Labour	-27.21*** (-3.47)	-23.29*** (-2.41)	-18.13*** (-3.14)	-25.59*** (-2.97)	-22.09*** (-2.84)	-25.20*** (-2.7)	-25.14*** (-3.54)
Female	-28.90*** (-3)	-26.34*** (-2.71)	-23.22*** (-3.17)	-12.68*** (-3.03)	-13.86*** (-2.33)	-20.53*** (-2.61)	-25.64*** (-3.88)
Private	-4.84 (-9.79)	0.68 (-7.36)	-2.59 (-10.21)	2.82 (-11.24)	-11.5 (-9.11)	-3.75 (-7.57)	-1.76 (-9.56)
Grade	30.82*** (-3.07)	23.70*** (-1.14)	21.63*** (-2.02)	17.97*** (-1.8)	18.48*** (-3.89)	22.74*** (-1.41)	27.89*** (-2.21)

Student achievement	Chile	Colombia	Costa Rica	Dom Rep.	Mexico	Peru	Uruguay
ESCS	9.50*** (-1.58)	5.47*** (-1.6)	6.64*** (-1.2)	5.78*** (-1.88)	4.15*** (-1.25)	6.96*** (-1.58)	8.55*** (-1.96)
Motivational Level	1.19 (-1.55)	8.69*** (-2.01)	2.29 (-2.1)	5.33** (-2.13)	6.39*** (-1.84)	11.03*** (-1.72)	9.21*** (-1.69)
Exp. Occupational Status	0.55*** (-0.1)	0.22*** (-0.07)	-0.08 (-0.06)	0.07 (-0.1)	0.28*** (-0.07)	0.58*** (-0.08)	0.36*** (-0.08)
Maths hours	-0.11 (-0.42)	0.49 (-0.53)	1.63 (-1.1)	-0.76 (-0.47)	0.79 (-0.69)	0.17 (-0.44)	2.33** (-1.12)
Student-Teacher Ratio	0.78 (-0.58)	-0.33 (-0.23)	0.05 (-0.19)	0.15 (-0.2)	-0.08 (-0.21)	0.36* (-0.21)	-0.12 (-0.24)
School mean (ESCS)	30.43*** (-5.22)	19.92*** (-5.6)	22.06*** (-4.37)	26.42*** (-7.82)	17.69*** (-4.46)	19.94*** (-4.12)	32.18*** (-6.35)
Constant	430.44*** (-13.69)	448.70*** (-11.24)	456.64*** (-8.56)	377.08*** (-16.12)	436.17*** (-12.01)	400.88*** (-11.33)	476.60*** (-11.86)
Ins1_1_1	3.31*** (-0.09)	3.23*** (-0.08)	3.29*** (-0.1)	3.09*** (-0.19)	3.27*** (-0.11)	3.16*** (-0.08)	3.27*** (-0.09)
Insig_e	4.09*** (-0.01)	4.03*** (-0.01)	3.96*** (-0.01)	3.91*** (-0.03)	4.10*** (-0.02)	4.08*** (-0.02)	4.12*** (-0.02)
Sample size (n=)	4,529	7,647	4,803	2,370	5,734	4,869	3,782

Standard errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Source: authors' work.

Household labour is not found to be significant (with 95% confidence levels) in shaping students learning except in the contexts of Peru and Uruguay. In these contexts, participating in household labour is shown to have a higher magnitude of effect than ones socio-economic and cultural status. Non-household labour participation, however, is found to be highly significant (with 99% confidence levels) and has a remarkable negative effect on student learning. This pattern is found in all the seven Latin American countries. These findings lead to the next question if the effects of non-household labour and household labour (in Peru and Uruguay) on student achievement are moderated by the individual's gender. In other words, the question can be rephrased as to understand if the gender effects (of being a female) on achievement, as seen in the above table, is influenced/moderated by male or female's participation in labour activities. The following tables (Table 4 and Table 5) aid in addressing the follow up question.

Table 4

Student achievement	Chile	Colombia	Costa Rica	Dom Rep.	Mexico	Peru	Uruguay
Female*Household Labour	-2.87 (-5.47)	-1.95 (-5.26)	-0.88 (-4.01)	0.2 (-7.35)	1.14 (-5.53)	-2.32 (-6.91)	8.87 (-6.78)
Female	-26.94*** (-4.87)	-24.82*** (-5.14)	-22.64*** (-3.8)	-12.84** (-6.3)	-14.80*** (-5.37)	-18.44*** (-6.98)	-32.58*** (-6.5)
Household Labour	0.15 (-3.96)	-3.47 (-4.04)	1.71 (-3.51)	3.09 (-6.01)	1.06 (-4.54)	-7.24 (-5.02)	-15.70*** (-5.74)
Non-Household Labour	-27.31*** (-3.45)	-23.34*** (-2.42)	-18.17*** (-3.15)	-25.58*** (-2.97)	-22.07*** (-2.83)	-25.23*** (-2.72)	-24.88*** (-3.58)
ESCS	9.51*** (-1.58)	5.46*** (-1.61)	6.63*** (-1.2)	5.78*** (-1.88)	4.15*** (-1.26)	6.95*** (-1.58)	8.49*** (-1.95)
Private	-4.87 (-9.76)	0.7 (-7.36)	-2.57 (-10.24)	2.82 (-11.24)	-11.5 (-9.11)	-3.77 (-7.56)	-1.77 (-9.62)
GRADE	30.82*** (-3.07)	23.70*** (-1.14)	21.62*** (-2.02)	17.97*** (-1.8)	18.47*** (-3.89)	22.75*** (-1.41)	27.93*** (-2.21)
Motivational Level	1.18 (-1.55)	8.67*** (-2)	2.29 (-2.1)	5.33** (-2.13)	6.40*** (-1.84)	11.01*** (-1.72)	9.20*** (-1.69)
Exp Occupational Status	0.55*** (-0.1)	0.22*** (-0.07)	-0.08 (-0.06)	0.07 (-0.1)	0.28*** (-0.07)	0.58*** (-0.08)	0.36*** (-0.08)
Maths hours	-0.11 (-0.42)	0.49 (-0.53)	1.62 (-1.1)	-0.76 (-0.47)	0.79 (-0.69)	0.17 (-0.44)	2.31** (-1.12)
Student-Teacher Ratio	0.78 (-0.58)	-0.33 (-0.23)	0.05 (-0.19)	0.15 (-0.2)	-0.08 (-0.21)	0.36* (-0.21)	-0.11 (-0.24)
School mean (ESCS)	30.40*** (-5.21)	19.88*** (-5.6)	22.04*** (-4.38)	26.42*** (-7.82)	17.68*** (-4.47)	19.95*** (-4.11)	32.18*** (-6.39)
Constant	429.61*** (-13.9)	447.79*** (-11.78)	456.35*** (-8.55)	377.15*** (-16.42)	436.56*** (-12.25)	399.98*** (-11.51)	480.28*** (-12.41)
Ins1_1_1	3.31*** (-0.09)	3.23*** (-0.08)	3.29*** (-0.1)	3.09*** (-0.19)	3.27*** (-0.11)	3.16*** (-0.08)	3.27*** (-0.09)
Insig_e	4.09*** (-0.01)	4.03*** (-0.01)	3.96*** (-0.01)	3.91*** (-0.03)	4.10*** (-0.02)	4.08*** (-0.02)	4.12*** (-0.02)
Sample size (n=)	4,529	7,647	4,803	2,370	5,734	4,869	3,782

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Source: authors' work.

Table 5

Student achievement	Chile	Colombia	Costa Rica	Dom Rep.	Mexico	Peru	Uruguay
Male*Non-Household Labour	0.57 (-6.76)	-2.4 (-3.62)	-14.00*** (-4.96)	-4.36 (-5.14)	-1.88 (-4.97)	-5.87 (-4.68)	-11.98** (-6.05)
Female	-28.75*** (-3.37)	-27.46*** (-3.03)	-25.72*** (-3.27)	-14.27*** (-3.17)	-14.37*** (-2.59)	-22.24*** (-2.83)	-29.39*** (-4.57)
Household Labour	-1.31 (-3.13)	-4.42 (-2.77)	1.39 (-2.86)	3.41 (-4.35)	1.59 (-3.5)	-8.12** (-4.09)	-10.49*** (-3.68)
Non-Household Labour	-27.53*** (-4.58)	-22.18*** (-2.67)	-9.66** (-4.46)	-23.37*** (-4.17)	-20.96*** (-4.01)	-21.73*** (-3.79)	-19.07*** (-4.83)
ESCS	9.50*** (-1.57)	5.47*** (-1.6)	6.63*** (-1.2)	5.80*** (-1.88)	4.16*** (-1.25)	6.99*** (-1.58)	8.48*** (-1.96)
Private	-4.83 (-9.79)	0.65 (-7.35)	-2.89 (-10.14)	2.92 (-11.25)	-11.48 (-9.11)	-3.7 (-7.57)	-2.07 (-9.67)
GRADE	30.82*** (-3.07)	23.71*** (-1.14)	21.56*** (-2.02)	17.97*** (-1.8)	18.48*** (-3.89)	22.72*** (-1.4)	27.84*** (-2.21)
Motivational Level	1.19 (-1.55)	8.68*** (-2)	2.25 (-2.1)	5.37** (-2.13)	6.37*** (-1.84)	11.07*** (-1.71)	9.24*** (-1.7)
Exp Occupational Status	0.55*** (-0.1)	0.22*** (-0.07)	-0.09 (-0.06)	0.06 (-0.1)	0.28*** (-0.07)	0.58*** (-0.08)	0.35*** (-0.08)
Maths hours	-0.11 (-0.43)	0.49 (-0.53)	1.7 (-1.1)	-0.73 (-0.47)	0.8 (-0.69)	0.18 (-0.44)	2.36** (-1.12)
Student-Teacher Ratio	0.78 (-0.58)	-0.34 (-0.23)	0.05 (-0.19)	0.15 (-0.2)	-0.08 (-0.21)	0.37* (-0.21)	-0.12 (-0.24)
School mean (ESCS)	30.42*** (-5.21)	19.89*** (-5.59)	21.99*** (-4.38)	26.24*** (-7.83)	17.66*** (-4.46)	19.82*** (-4.13)	32.14*** (-6.39)
Constant	430.39*** (-13.75)	449.43*** (-11.29)	457.80*** (-8.6)	377.82*** (-16.12)	436.43*** (-11.91)	401.57*** (-11.35)	478.92*** (-12.06)
Ins1_1_1	3.31*** (-0.09)	3.23*** (-0.08)	3.29*** (-0.1)	3.09*** (-0.19)	3.27*** (-0.11)	3.16*** (-0.08)	3.27*** (-0.09)
Insig_e	4.09*** (-0.01)	4.03*** (-0.01)	3.96*** (-0.01)	3.91*** (-0.03)	4.10*** (-0.02)	4.08*** (-0.02)	4.12*** (-0.02)
Sample size (n=)	4,529	7,647	4,803	2,370	5,734	4,869	3,782

Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Source: authors' work.

Table 4 shows that there is no significant difference in achievement between males and females, on average, engaging in household labour across all 7 countries. The magnitude of the difference, although non-significant, is negligible. There is also no difference in outcomes between males who participate in household labour compared to those who do not except in the context of Uruguay. Boys engaged in household labour in Uruguay scored 15.7 points lower, on average. In Peru, they scored 7.24 points lower than boys who did not engage in household labour although the difference is not statistically significant. The findings are consistent for household labour participation effects from table 3 (which showed the average of both males and females) where only Uruguay and Peru had significant differences.

The results from Table 5 show that gender does not moderate the effects of engagement in non-Household labour on achievement outcomes except in the context of Costa Rica and Uruguay. The penalty of participating in non-household labour is significantly higher for males over females in Costa Rica and Uruguay (-14 and -11.98 points higher respectively). However, on an average for both males and females, participating in non-household labour is associated with lower outcomes (from table 3 and table 4). Table 5 reiterates the same findings with respect to the effects of household labour participation from table 3 and also additionally demonstrates the significantly lower outcomes among females who engage in non-household labour compared to those who do not.

Some of the broad observations which speak to literature regarding student education status are that, unlike in many other developing country contexts, attending a private school has shown no significant effect on average in all the above Latin American countries. The results also point broadly to poor quality education offered in secondary schools, as number of hours invested in studying by the student does not show any significant gains in their learning except marginally in Uruguay.

Lastly, although gender does not moderate the effects of household labour participation on student learning, it still remains as an influential category. The magnitude of influence gender has in shaping learning is at least 2.5 to 5 times more than that of socio-economic and cultural status, depending on the country. Girls are found to consistently underperform in comparison to boys in every Latin American country without exception.

7. Conclusion

This study used international largescale assessment data (PISA) to highlight the trends and patterns in child labour participation in household and non-household activities in the seven Latin American countries. The results show a high level of participation, at least 65% in household activities and not less than 20% in non-household activities, in all countries in the sample. Furthermore, participation in household labour is not found to significantly affect students learning in majority of the countries except Peru and Uruguay.

Participation in non-household labour is shown to significantly hamper students' progress in learning across all of the seven countries of Chile, Colombia, Costa Rica, Dominican Republic, Mexico, Peru and Uruguay. This finding is in line with the findings of previous literature that non-household labour is a serious problem plaguing Latin American. Where previous studies have focused on the effect of child labour on school dropouts, repetitions of grade, attendance and years of schooling, this study contributes to knowledge by specifically demonstrating the magnitude of effect of participation in such activities on students learning levels, relative to other variables. It further adds how gender fails to moderate the household labour effects on learning and also in the case of non-household labour effects in majority of countries (except Costa Rica and Uruguay).

It does however point to gender as strongly associated with educational attainment, with girls underperforming boys across all countries despite literature positing that Latin America had in fact closed this gap, with female students even beginning to outperform males (Ahuja & Filmer, 1995; Grant & Behrman, 2010).

The strength of the study stems from the nature of survey data. The sample is representative for all the countries and measures of learning achievement are statistically sophisticated through inclusion of plausible values and appropriate sampling weights. This adds to the strength of generalisability of the results. The trends in participation levels, association and moderation between variables highlighted in the paper can be said to hold true for the entire population of 15-year olds in the above Latin American countries. Another strength of the study is the use of hierarchical modelling. In many of the empirical studies involving child labour, it is not common to account for the clustering of children's characteristics to estimate the effects of child labour on chosen outcomes.

The main weakness of the study is the measurement of participation in household and non-household activities using a binary scale. It collapses various forms of labour and duration of labour into a single category thereby

impeding the further deciphering of the nature of activities across various countries or socio-economic categories. However, this lack of detail can be justified as the focus of the study is exclusively on education and we exploited the opportunity of the dataset capturing participation in labour albeit in a rudimentary form.

This study does not establish any causal relationship between participation in child labour (in household or non-household activities), or gender, or their interaction on the student achievement. This paper gives a potentially valuable perspective into the reality of the nature of association and trends across countries. However, a more rigorous identification strategy is needed to address the endogeneity problem to infer causality.

At age 15 (the age of all students participating in PISA) participation in the labour market is legal in all Latin American countries (Appendix 1). This is often only with authorisation from the inspectorate or local authority, but given the previously discussed lack of inspectors, it is doubtful whether the work conditions or hour limits are being enforced. This study shows that even potentially legal child labour may play a significant role in the regions' low educational attainment and should not be neglected in future work in the area.

104

The empirical trends and analysis can potentially contradict some of the theoretical understanding on the gendered nature of household activities or non-household activities. It challenges claims that household and non-household activities have differing effects for male and female education outcomes (except for Costa Rica and Uruguay in the contexts of non-household labour). There is inequality between genders in terms of achievement levels and also between those who participate in child labour (more significantly in non-household activities) and those who do not; however, the harmful effects of participating in non-household labour is not connected to gender. This is because the participation in non-household activities is undesirable irrespective of one's gender and there are other spheres of discrimination or decision-making which are potentially driving the inequality in learning outcomes between males and females that demand attention. Therefore, the focus for further research or policy ought to include other spheres of influence, or socio-economic dimensions to address the gender inequality in achievement levels.

^[1] Brazil is omitted from the list due to large amount (approx. 45%) of missing data in the category of labour participation in household and non-household domains.

It is also important to note that this study doesn't use the latest PISA 2018 data as the survey no longer includes variables pertaining to any forms of labour.

Appendix 1

	Chile	Colom- bia	Costa Rica	Domi- nican Rep.	Mexico	Peru	Uruguay
School compulsory until	17	15	16	13	17	16	14
Minimum Age for Work	15	15	15	14	15	14	15
Minimum Age for Hazar- dous Work	18	18	18	18	18	18	18
Ratification of Interna- tional Conventions on Child Labour	✓	✓	✓	✓	✓	✓	✓

Source: ILAB, 2020.

References

Ahuja, V. & Filmer, D. (1995). Educational attainment in development countries: New estimates and projections disaggregated by gender. *Journal of Educational Planning and Administration* X(3): 229-254

Altschul, I. (2012). Linking socioeconomic status to the academic achievement of Mexican American youth through parent involvement in education. *Journal of the Society for Social Work and Research*, 3(1), 13-30.

Archer, L., & Francis, B. (2006). *Understanding minority ethnic achievement: Race, gender, class and 'success'*. Routledge.

Arends-Kuenning, M., & Duryea, S. (2006). The effect of parental presence, parents' education, and household headship on adolescents' schooling and work in Latin America. *Journal of Family and Economic Issues*, 27(2), 263-286.

Assaad, R., Levison, D., & Zibani, N. (2010). The effect of domestic work on girls' schooling: Evidence from Egypt. *Feminist Economics*, 16(1), 79-128.

Bargain, O., & Boutin, D. (2017). Minimum age regulation and child labor: New evidence from Brazil

Barnard, W. M. (2004). Parent involvement in elementary school and educational attainment. *Children and youth services review*, 26(1), 39-62.

Barone, F. J. (1993). The effects of part-time employment on academic performance. *NASSP Bulletin*, 77(549), 67-73.

Basu, K. (2005). Global labour standards and local freedoms. In *WIDER perspectives on global development* (pp. 175-200). Palgrave Macmillan, London.

Basu, K., & Tzannatos, Z. (2003). The Global Child Labor Problem: What do we know and what can we do?. *The world bank economic review*, 17(2), 147-173.

Basu, K., & Van, P. H. (1998). The economics of child labor. *American economic review*, 412-427.

- Beegle, K., Dehejia, R., & Gatti, R. (2009). Why should we care about child labor? The education, labor market, and health consequences of child labor. *Journal of Human Resources*, 44(4), 871-889.
- Behrman, J., Duryea, S., & Székely, M. (1999). Schooling investments and aggregate conditions: A household-survey-based approach for Latin America and the Caribbean. *IDB-OCE Working Paper*, (407).
- Boockmann, B. (2010). The effect of ilo minimum age conventions on child labor and school attendance: Evidence from aggregate and individual-level data. *World Development*, 38(5), 679-692.
- Bossuyt, N., Gadeyne, S., Deboosere, P., & Van Oyen, H. (2004). Socio-economic inequalities in health expectancy in Belgium. *Public health*, 118(1), 3-10.
- Card, D. (1999). The causal effect of education on earnings. In *Handbook of labor economics* (Vol. 3, pp. 1801-1863). Elsevier.
- Chernikovsky, D., Lucas, R. E., & Mueller, E. (1985). *The household economy of rural Botswana: An African case*. The World Bank.
- Choi, A. (2018). De padres a hijos: expectativas y rendimiento académico en España. *Presupuesto y gasto público*, 90, 13-32.
- Driessen, G., Smit, F., & Slegers, P. (2005). Parental involvement and educational achievement. *British educational research journal*, 31(4), 509-532.
- Easterlin, R. A. (2003). Explaining happiness. *Proceedings of the National Academy of Sciences*, 100(19), 11176-11183.
- Eckstein, Z., & Wolpin, K. I. (1999) 'Why youths drop out of high school: the impact of preferences, opportunities, and abilities,' *Econometrica* 67, 1295-339
- Edmonds, E. V., & Shrestha, M. (2012). The impact of minimum age of employment regulation on child labor and schooling. *IZA Journal of Labor Policy*, 1(1), 14.
- Gillborn, D., & Mirza, H. S. (2000). Educational Inequality: Mapping Race, Class and Gender. A Synthesis of Research Evidence.
- Gould, E. D. (2007). Cities, workers, and wages: A structural analysis of the urban wage premium. *The Review of Economic Studies*, 74(2), 477-506
- Grant, M. J., & Behrman, J. R. (2010). Gender gaps in educational attainment in less developed countries. *Population and development review*, 36(1), 71-89.
- Grynspan, R. (2010). Foreword. In *Declining inequality in Latin America: A decade of progress?* (pp. Vi-Vii). Brookings Institution Press.
- Guarcello, L., Lyon, S., & Rosati, F. C. (2008). Child labor and education for all: An issue paper. *The journal of the History of Childhood and Youth*, 1(2), 254-266
- Gunnarsson, V., Orazem, P. F., & Sánchez, M. A. (2006). Child labor and school achievement in Latin America. *The World Bank Economic Review*, 20(1), 31-54.
- Heymann, J., Raub, A., & Cassola, A. (2013). Does prohibiting child labor increase secondary school enrolment? Insights from a new global dataset. *International Journal of Educational Research*, 60, 38-45.

- ILAB (2020). *Findings on the Worst Forms of Child Labor*. U.S. Department of Labor, Bureau of International Labor Affairs. Available at <https://bit.ly/2Hwp3e2>.
- ILO (1996). *What is child labour*. International Programme on the Elimination of Child Labour (IPEC) Available at <https://bit.ly/3oudwfr>.
- Khang, Y. H., Lynch, J. W., & Kaplan, G. A. (2004). Health inequalities in Korea: age-and sex-specific educational differences in the 10 leading causes of death. *International Journal of Epidemiology*, 33(2), 299-308.
- Kao, G., & Thompson, J. S. (2003). Racial and ethnic stratification in educational achievement and attainment. *Annual review of sociology*, 29(1), 417-442.
- Lee, J.-S., & Bowen, N. K. (2006). Parent Involvement, Cultural Capital, and the Achievement Gap Among Elementary School Children. *American Educational Research Journal*, 43(2), 193-218. <https://doi.org/10.3102/00028312043002193>
- Levison, D. (2000). Children as economic agents. *Feminist Economics*, 6(1), 125-134.
- Lochner, L. (2011). *Non-production benefits of education: Crime, health, and good citizenship* (No. w16722). National Bureau of Economic Research.
- Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American economic review*, 94(1), 155-189.
- López-Calva, L. F., & Lustig, N. C. (Eds.). (2010). *Declining inequality in Latin America: A decade of progress?*. Brookings Institution Press.
- Machin, S., Marie, O., & Vujić, S. (2011). The crime reducing effect of education. *The Economic Journal*, 121(552), 463-484.
- Marsh, H. W. (1991). Employment during high school: Character building or a subversion of academic goals?. *Sociology of education*, 172-189.
- Montmarquette, C., Viennot-Briot, N., & Dagenais, M. (2007). Dropout, school performance, and working while in school. *The Review of Economics and Statistics*, 89(4), 752-760.
- Murillo, F. J. (2003). Una panorámica de la investigación iberoamericana sobre eficacia escolar. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 1(1).
- Nam, Y., & Huang, J. (2009). Equal opportunity for all? Parental economic resources and children's educational attainment. *Children and Youth Services Review*, 31(6), 625-634
- OECD (2017). *PISA 2015 assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving*. Organisation for Economic Co-operation and Development Publishing
- Parent, D. (2006). Work while in high school in Canada: its labour market and educational attainment effects. *Canadian Journal of Economics/Revue canadienne d'économique*, 39(4), 1125-1150.
- Pfeffer, F. T. (2018). Growing wealth gaps in education. *Demography*, 55(3), 1033-1068
- Psacharopoulos, G. (1997). Child labor versus educational attainment Some evidence from Latin America. *Journal of population economics*, 10(4), 377-386.

- Putnick, D. L., & Bornstein, M. H. (2016). Girls' and boys' labor and household chores in low- and middle-income countries. *Monographs of the Society for Research in Child Development*, 81(1), 104.
- Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor: New evidence and possible explanations. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity? Rising inequality, schools, and children's life chances* (pp. 91–115). New York, NY: Russell Sage Foundation
- Rosenzweig, M. R., & Evenson, R. (1977). Fertility, schooling, and the economic contribution of children of rural India: An econometric analysis. *Econometrica: journal of the Econometric Society*, 1065-10.
- Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142-151.
- Ziol-Guest, K. M., & Lee, K. T. (2016). Parent income-based gaps in schooling: Cross-cohort trends in the NLSYs and the PSID. *AERA Open*, 2(2). <https://doi.org/10.1177/2332858416645834>.

Should different countries participating in PISA interpret socioeconomic background in the same way? A measurement invariance approach

Nurullah Eryilmaz¹ ; Mauricio Rivera-Gutiérrez² ; Andrés Sandoval-Hernández¹ 

¹ University of Bath, United Kingdom; ² University of Brighton, United Kingdom

Abstract. It has been claimed that there is a lack of theory-driven constructs and a lack of cross-country comparability in International Large-Scale Assessment (ILSA)'s socio-economic background scales. To address these issues, a new socio-economic background scale was created based on Pierre Bourdieu's cultural reproduction theory, which distinguishes economic, cultural and social capital. Secondly, measurement invariance of this construct was tested across countries participating in the Programme for International Student Assessment (PISA). After dividing the countries which participated in PISA 2015 into three groups, i.e., Latin American, European, and Asian, a Multi-Group Confirmatory Factor Analysis was carried out in order to examine the measurement invariance of this new socio-economic scale. The results of this study revealed that this questionnaire, which measures the socio-economic background, was not found to be utterly invariant in the analysis involving all countries. However, when analysing more homogenous groups, measurement invariance was verified at the metric level, except for the group of Latin American countries. Further, implications for policymakers and recommendations for future studies are discussed.

Keywords: measurement invariance, multi-group confirmatory factor analysis, cultural reproduction theory, Pierre Bourdieu, socio-economic scales, PISA

¿Los países que participan en PISA deberían interpretar por igual el ambiente socioeconómico? Un enfoque de medición de invariancia

Resumen. Se ha argumentado que existe una falta de interpretaciones basadas en teorías, junto con una falta de comparabilidad entre países en las escalas de ambientes socioeconómicos de las evaluaciones internacionales a gran escala (ILSA, por sus siglas en inglés). A fin de dar respuesta a estos asuntos, se ha creado una nueva escala de ambiente socioeconómico basada en la teoría de reproducción cultural de Pierre Bourdieu, que distingue capital económico, cultural y social. En segundo lugar, la invariancia de medición de esta interpretación se ha probado en distintos países que participaron en PISA 2015 en tres grupos, es decir, se ha llevado a cabo un Análisis de Factor Confirmatorio Multigrupo de América Latina, Europa y Asia para examinar la medición de la variancia de esta nueva escala socio-económica. Los resultados han puesto de manifiesto que este cuestionario, que mide el ambiente socioeconómico, no es totalmente invariante en el análisis en relación con todos los países. No obstante, al analizar grupos más homogéneos, la invariancia de la medición se ha verificado a nivel métrico, salvo para el grupo de países de Latinoamérica. Además, se han debatido las implicaciones para el legislador junto con las recomendaciones para estudios futuros.

Palabras clave: invariancia de medición; análisis de factor confirmatorio multigrupo; teoría de reproducción cultural; Pierre Bourdieu; escalas socio-económicas; PISA.

Os países participantes do PISA deveriam interpretar o ambiente socioeconómico de maneira igual? Uma abordagem de medição de invariância

Resumo. Argumentou-se que há uma falta de interpretações baseadas em teorias, juntamente com uma falta de comparabilidade entre países nas escalas de ambientes socioeconómicos das avaliações internacionais em grande escala (ILSA, em sua sigla em inglês). Para responder a essas questões, uma nova escala de ambiente socioeconómico foi criada com base na teoria da reprodução cultural de Pierre Bourdieu, que distingue o capital económico, cultural e social. Em segundo lugar, a invariância de medição dessa interpretação foi testada em diferentes países que participaram do PISA 2015 em três grupos, ou seja, realizou-se uma Análise de Fator Confirmatório Multigrupo da América Latina, Europa e Ásia para examinar a medição da variância desta nova escala socioeconómica. Os resultados mostraram que este questionário, que mede o ambiente socioeconómico, não é totalmente invariável na análise em relação a todos os países. Porém, ao analisar grupos mais homogéneos, verifica-se a invariância da medição em nível métrico, exceto para o grupo de países América Latina. Além disso, as implicações para o legislador foram discutidas juntamente com as recomendações para estudos futuros.

Palavras-chave: invariância de medição, análise fatorial confirmatória multigrupo, teoria da reprodução cultural, Pierre Bourdieu, escalas socioeconómicas, PISA.

1. Introduction

International Large-Scale Assessments (ILSAs) have been given much attention due to the ever-increasing participation rate across countries in the world (Addey, Sellar, Steiner-Khamsi, Lingard & Verger, 2017). Retrospectively speaking, the International Association for the Evaluation of Educational Achievement (IEA) carried out the first ILSA in 1960, with the participation of twelve pilot countries (Addey & Sellar, 2018). By the end of the 1990s, the number of participating countries was approximately 40 (Tijana & Anna, 2015). Nowadays, nearly 70% of countries across the world participate in these evaluations (Lietz, Cresswell, Rust & Adams, 2017). Table 1 shows a selection of recent ILSAs and the respective number of participating countries for reference.

Table 1. Recent ILSA studies

ILSA studies	Participation
PISA 2018	79 countries, 37 OECD member countries
TIMSS 2015	57 countries and 7 benchmarking entities
PIRLS 2016	61 participants (50 countries and 11 benchmarking)
ICCS 2016	24 countries around the world
PASEC 2014	10 countries in Francophone West Africa
SACMEQ 2013	15 ministries of education
TERCE 2013	15 participants (14 countries and 1 Mexican state)
ERCE 2019	19 countries

110

The OECD’s Programme for International Student Assessment (PISA) shows the highest number of participating countries, compared to other ILSAs. Participation in PISA has also significantly increased over time. In 2000, 43 countries participated in this assessment, whereas 72 took part in 2015, and 80 in the latest round which took place in 2018 (Steiner-Khamsi, 2019). Since 2000, the proportion of countries participating in PISA has almost doubled worldwide.

There are two main objectives behind the application of ILSA studies: contributing comparatively to the functioning of educational systems, as well as illuminating the development of educational and training programmes in participating countries from many diverse regions (Torney-Purta & Amadeo, 2013). In this context, ILSAs introduce a major challenge relating to comparability, in that their underlying tools should enable sensible cross-country comparisons to comply with their aims (Goldstein, 2017; Segeritz & Pant, 2013).

The design process of ILSAs requires to adhere to rigorous standards to make comparisons possible across a wide range of participants, which are diverse in terms of culture, and economic and political contexts (Miranda & Castillo, 2018). Results of these assessments should be comparable because, as Mullis (2002, p. 2) states, they “provide an opportunity to examine the impact on achievement of different educational approaches and additional insight into ones’ educational system”. To meet these requirements, measurement instruments should ensure that participants who hold the same level of a certain characteristic obtain the same score in the test.

It is in this context that measurement invariance becomes a key condition that needs to be verified in these studies. The design and implementation of measurement instruments should allow all countries participating in ILSAs to be reflected in an equal manner. Measurement invariance should be taken into account as a significant matter in order to make group comparisons, that is to say, only if measurement invariance is ensured, then researchers can make comparisons between different cultures (Van de Vijver & Leung, 1997; Van de Vijver & Poortinga, 2002; Byrne & Van de Vijver, 2010). To put it differently, as long as a given scale’s measurement invariance is confirmed among relevant groups, scores obtained from it can be used to make a comparison across groups (Uysal & Arıkan, 2018). Conversely, if measurement invariance is not verified, both the validity of the scores and interpretations, and the fairness of the measurement process remain disputable (Gregorich, 2006). As a natural consequence of this, interpretations, and conclusions about group differences across countries may not be valid (Cheung and Rensvold, 2002).

The question on cross-cultural comparability of cognitive assessments in ILSAs has had considerable attention in the literature (e.g. Wu, 2010; Klieme, 2016 & Oliveri & Ercikan, 2011). Numerous studies have addressed the question of measurement invariance for PISA cognitive assessments, while less attention has been paid to PISA context questionnaires (Van de Vijver, 2018) (i.e. student questionnaires, e.g. He et al., 2018). Although Hopfenbeck et al. (2018) explicitly states that measurement invariance is just as important for background questionnaires, Rutkowski and Rutkowski (2010) highlight that for all participating countries, background questionnaires comparability has not been explored to the same extend as the cognitive assessments. Despite this, background questionnaire responses from participants are still utilized to make approximate estimations of the population and subpopulation achievements by using linear regression models (Rutkowski & Rutkowski, 2010). In light of this, ‘the degree to which a single measure of socioeconomic background is reliable and valid for all participating countries is not widely discussed’ (Rutkowski & Rutkowski, 2013, p. 260).

Socioeconomic status (SES) is one of the most frequently used predictive factors of academic achievement in the literature (Sirin, 2005; White, 1982). The socioeconomic background of students has increasingly become essential in educational research to determine whether there is segregation, differences, or inequalities between students in ILSAs, especially in PISA. In fact, this aspect is included in the fourth United Nation's Sustainable Development Goal (SDG4; UN, 2015), which aims to ensure inclusive and equitable quality education and promote opportunities for all students. After the study of Coleman et al. (1966), the link between socio-cultural and economic status and academic achievement has been demonstrated. To date, it has been clearly stated that SES is of great importance as an indicator. It has been integrated to studies on students' educational outcomes as a supplementary component (Bornstein & Bradley, 2003; White, 1982; Neff, 1938; Bradley & Corwyn, 2002; Sirin, 2005). For instance, Sirin (2005) review's findings highlighted that student's educational achievement is significantly affected by the socio-economic structure of families.

112 There are numerous studies addressing the association between SES and student academic achievement in the context of cross-cultural studies, particularly using PISA data (e.g., Park & Sandefur, 2016; Thein & Ong, 2015; Kalaycioglu, 2015; Pokropek et al. 2015). Nonetheless, there have been two fundamental criticisms regarding the use of SES in PISA, particularly when addressing questions on socio-economic unevenness: the lack-of-theory issue, and the problem of comparability. First, it is critical to note that, in general, decisions about what will be included in ILSA studies are made without taking into account existing theories, and analyses tend to only draw on statistical measures, such as correlations and regression models (Lauder et al. 1998; Coe & Fitz-Gibbon 1998). In that sense, the need to consolidate and understand the theoretical frame regarding socio-cultural and economic status as measured in ILSAs has emerged. Second, there is a fundamental debate as to whether SES has the same meaning across countries, particularly in terms of the indicators measuring this construct (Rutkowski & Rutkowski, 2013). This is a question on the validity of the interpretations made around SES and whether it can be measured across countries that have diverse contexts and conditions. Pokropek et al. (2017) gave an illustrative example in this point:

Having a car may not indicate socioeconomic status in the same way in the United States as it does in Japan. While in the United States car ownership is virtually universal (because distances between locations are large and the costs of maintaining a car low), in Japan car ownership is less common even in relatively wealthy families (as public transportation is widespread and efficient, and the cost of maintaining a car is high (p.244).

In order to address the first of the abovementioned criticisms, this paper aims to obtain and establish experimental confirmation for Pierre Bourdieu's cultural reproduction theory in order to theoretically support PISA's socio-economic status construct. This will be done by constructing one scale which does not originally exist in PISA in accordance with this theory. Cultural reproduction theory will be explained in detail in the literature review section.

Secondly, this paper aims to test the measurement invariance of the socioeconomic status construct across countries participating in PISA 2015 (OECD, 2018). When we look at the structure of PISA 2015, it can easily be stated that participating countries comprise of a wide range of populations, which includes different cultures, economic systems, and diverse spoken languages. The measurement invariance of PISA's SES structure has been tested across all countries but has not been properly confirmed (e.g., Rutkowski & Rutkowski, 2013; Pokropek et al., 2017)

To make cross-group comparisons more logical and reasonable, and based on the formation of more homogeneous groups, PISA 2015 participating countries will be split into three groups, i.e., Latin America, Asia, and Europe, considering the regions they belong to. While dividing the participating countries into three geographical groups, countries with similar cultural, historical and macroeconomic backgrounds were considered as a single group.

In summary, this paper intends to give theoretical support to the socio-economic status construct in PISA and, consequently, to verify whether this scale shows measurement equivalence across PISA participating countries. Therefore, this study aims to illustrate whether the questionnaire designed to measure the socioeconomic background of students who participated in PISA 2015 represents the same meaning across countries, particularly when grouped according to their region/continent. The results of this study will provide valuable information to improve those measures relating to concepts like socioeconomic status and the methods currently used to analyse its association with educational outcomes. National and local governments, as well as international organisations in charge of implementing this kind of assessments could be the main beneficiaries of the conclusions developed in this research.

2. Literature review

This section looks to address firstly the current lack of theory supporting ILSAs' SES constructs, particularly in works that use PISA data (Caro & Cortes, 2012) and, secondly, the lack of evidence supporting cross-cultural comparisons of these constructs. Hence, the literature will be organised around two main topics: (1) Pierre Bourdieu's cultural reproduction theory, (2) cross-cultural research works using SES indicators in ILSAs. Cultural reproduction theory will be discussed because a connection will be established between this theory and our proposed SES construct. Cross-cultural research using ILSA data will be reviewed in order to show the lack of empirical evidence to sustain the validity of comparison of SEs constructs across countries.

2.1 Cultural Reproduction Theory of Pierre Bourdieu

SES is described as a structure resulting from the combination of many components based on social, cultural, and economic factors such as individual's education level, household income, occupation, and home possessions. In the same way, the concept of capital pointed out by Bourdieu (1986) and Coleman (1988), expressed as three types of capital, i.e., economic, cultural and social, has been used in studies by most researchers to reveal the possible association between the family's socio-economic status and students' academic achievement. Capital is defined as a notion that "takes time to accumulate and which, as a potential capacity to produce profits and to reproduce itself in identical or expanded form, contains a tendency to persist in its being" (Bourdieu, 1986, p. 241).

Three forms of capital can be identified (Bourdieu 1986, p.242): economic capital, "which is immediately and directly convertible into money and might be institutionalized in the form of property rights"; cultural capital, "which is convertible, in certain conditions, into economic capital and might be institutionalized in the form of educational qualifications"; and social capital, "which is convertible, in certain conditions, into economic capital and might be institutionalized in the form of a title of nobility".

Bourdieu (1986) highlights that economic capital is the root of the other types of capital. In other words, cultural and social capital are a result of the modification of economic capital. Family income might lead to resources that allow them to participate in after-school activities as well as to reach high-quality instructional facilities and to build linkage with others (Lareau, 2011). There are three forms of cultural capital (Bourdieu, 1997): incorporated or embodied cultural capital, objectified cultural capital and institutionalized cultural capital. Embodied cultural capital includes linguistic

and cognitive competencies, cultural habits and tendencies. Objectified cultural capital contains possession and cultural goods, e.g., books, paintings. Institutionalized cultural capital comprises formal educational qualifications such as diplomas or certificates. It was revealed that cultural capital of students had significant effects on academic achievement (e.g. Yang, 2003; Barone, 2006). As DiMaggio (1982, p.190) points out: '[teachers] communicate more easily with students who participate in elite cultures, give them more attention and special assistance, and perceive them as more intelligent or gifted than students who lack cultural capital'. Social capital is expressed as belonging to a certain group based on the principle of recognizing and interacting with one another (Bourdieu, 1986). One reason for the differences in the educational level of students is the social capital produced as a result of the connections and interactions of the families at different levels (Rogosic & Baranovic, 2016).

There is a growing body of PISA-related research focusing on either largely cultural capital (Puzic et al., 2016; Puzic et al., 2018; Bodovski et al., 2017; Pitzalis & Porcu, 2016; Tan, 2015; Marteleto & Andrade, 2013) or social capital (Aloisi & Tymms, 2017). Furthermore, Garcia-Aracil et al. (2016) considered social and cultural capital whereas Caro et al. (2014) considered all three types, including economic capital. Education studies have traditionally conceptualised social inequality as a multidimensional phenomenon (Abel, 2008), however, most studies do not address the complex structure of cultural, economic and social capital. At least in quantitative studies, it is very rare to find studies where an integrated SES structure is considered. To address this gap, in this paper we designed a model considering economic, cultural, and social capital drawing on PISA 2015 socioeconomic background questionnaires.

Measurement invariance analysis has been frequently and widely used over the last decade and continues to attract interest. During the past years, much attention has been paid to testing measurement invariance of ILSAs' cognitive assessments. Wu, Li & Zumbo (2007) investigated the measurement invariance of the mathematic test using TIMSS 1999 data across seven countries but found that invariance was not supported. In the Italian context, Alivernini (2011) tested the measurement invariance of PIRLS 2006's reading literacy scale across students' gender and their immigration status and results showed that making such comparisons was not empirically supported.

Recently, studies have shifted their attention towards background questionnaires. For example, Segeritz and Pant (2013) examined the measurement invariance of the PISA 2003's Students' Approaches to Learning instrument across immigrant groups in Germany and did not achieve all levels

of invariance. In Turkey, Demir (2017) explored the measurement invariance of students' affective characteristics across gender categories and found that this scale was largely comparable between gender groups.

There is a limited number of studies addressing the measurement invariance of the socio-economic status indicator. In the United Kingdom, Hobbs and Vignoles (2007) stated that Free School Meal (FSM) Eligibility, which has been commonly used as a proxy for SES in UK educational research, has not enough supporting evidence to make comparison across families with dissimilar characteristics. Lenkeit et al. (2015) reveal that – using data from the Children of Immigrants Longitudinal Survey in Four European Countries (CILS4EU) in England – there are differences across immigrant groups in terms of the family background construct.

With regard to ILSA data, few studies relating to the measurement invariance of SES have been conducted. Hansson and Gustafsson (2013) found that invariance of SES was supported when comparing Swedish and non-Swedish populations, using TIMSS 2003 data. Rutkowski and Rutkowski (2013) found that the home possession indicator present in PISA 2009 SES index was not comparable across the 65 participant countries. Furthermore, Hernandez et al. (2019) explored the comparability of different socioeconomic scales of three ILSA studies: TERCE, PISA and TIMSS. None of the socioeconomic background scales was found to be fully invariant, which suggested that comparisons across countries should be made with caution.

Caro, Sandoval-Hernandez and Lüdtke (2014) highlight that, when using SES variables for making comparisons, recommendations or comments about participating countries, researchers should be extremely attentive and careful as comparisons are not fully supported by the evidence. Correspondingly, Hopfenbeck et al. (2018) emphasized in their systematic review that numerous articles suggest policymakers and researchers be careful and cautious when using PISA data as a valid benchmarking or informed policy-making tool.

3. Methodology

3.1 Sample

PISA is a triennial survey which was firstly launched by the Organisation for Economic Co-operation and Development (OECD) in 2000. The PISA 2015 study was administered in 35 OECD and 37 non-OECD (partner) countries. PISA implements a two-stage stratified sampling strategy. In the first stage, schools are sampled using a probability selection on the basis of the number of students enrolled in the school. In the second stage, a certain

sample of students is randomly selected within each school. 540.000 students took part in PISA 2015, representing about 29 million 15-year-olds in schools of the 72 participating countries (OECD, 2018). More detailed information of the sampling design, including weighting procedures can be found in the PISA 2015 Technical Report (OECD, 2017). To explore cross-cultural comparability across countries, the current study considered 35 OECD countries and 19 partner countries (a total of 54 countries). The rest of partner countries were removed from the analysis due to not having valuable information for some variables.

3.2 Measures

Nine subscales included in the PISA 2015's student questionnaire were selected to create a new SES scale based on Pierre Bourdieu's cultural reproduction theory. Indexes were used as indicators rather than each individual item, except for 'number of books' (ST013Q01TA). Table 2 indicates the items used to develop the new scale.

Table 2. PISA 2015 subscales used for the development of a new SES scale.

Code	Name	Description
Wealth	Family wealth possessions	Summary index consisting of a room of your own, internet, televisions, cars, rooms with a bath or shower, cell phones with internet access, computers, tablet computers, e-book readers and three country-specific items.
Pared	Parental education	Summary index of highest parental education schooling.
Hisei	Highest parental occupational status	Summary index of highest parental occupational status.
Cultposs	Cultural possessions	Summary index consisting of classic literature, books of poetry, works of art, books on art, music or design, musical instrument.
Hedres	Home education resources	Summary index consisting of a desk to study, a quiet place to study, a computer you can use for schoolwork, educational software, books to help with your schoolwork, technical reference book, a dictionary.
ST013Q01TA	Number of books	Single question asking about how many books are there in your home? 0-10 books (1), 11-25 books (2), 26-100 books (3), 101-200 books (4), 201-500 books (5), more than 500 books (6).
Cooperate	Enjoy Co-operation	Summary index consisting of 'I am a good listener', 'I enjoy seeing my classmates be successful', 'I take into account what others are interested in', 'I enjoy considering different perspectives'.

Code	Name	Description
Cpsvalue	Value Co-operation	Summary index consisting of 'I prefer working as a part of team to working alone', 'I find that teams make better decisions than individuals', 'I find that teamwork raises my own efficiency', 'I enjoy cooperating with peers'.
Emosups	Parents Emotional Support	Summary index consisting of 'My parents are interested in my school activities', 'My parents support my educational efforts and achievements', 'My parents support me when I am facing difficulties at school', 'My parents encourage me to be confident'.

Table 3 shows the respective descriptive statistics. Items were grouped into three groups indicating whether they measure economic capital, cultural capital or social capital.

Table 3. Descriptive statistics for the variables used in this study

	Minimum (min)	Maximum (max)	Mean	Standard deviation (SD)
<i>Economic Capital</i>				
WEALTH (index)	-7.635	4.715	-0.321	1.26
PARED (index)	3	18	13.34	3.25
HISEI (index)	11	89	50.44	22.36
<i>Cultural Capital</i>				
CULTPOSS (index)	-1.84	2.63	-0.05	0.95
HEDRES (index)	-4.412	1.177	-0.178	1.07
ST013Q01TA	1.00	6.00	2.963	1.46
<i>Social Capital</i>				
COOPERATE (index)	-3.33	2.29	0.05	1.01
CPSVALUE (index)	-2.83	2.14	0.10	1.00
EMOSUPS (index)	-3.08	1.10	-0.03	0.99

3.3 Analytical Strategy

The psychometric characteristics of the created scale were evaluated following a number of procedures. First, reliability (internal consistency) was evaluated using Cronbach's alpha coefficient (Cronbach, 1951). This coefficient ranges from 0 to 1, with values close to 1 indicating high levels of reliability. Second, a confirmatory factor analysis was implemented to evaluate the model fit for each country (see more information in the results section). We then applied a multi-group confirmatory factor analysis (MG-CFA) to examine the model fit and cross-cultural comparability of this scale across all education

systems. Lastly, countries were split into three different sub-groups (Latin American countries, Asian countries, and European countries) to examine the cross-cultural comparability of this scale within more homogeneous groups.

3.3.1 Confirmatory Factor Analysis

Models were estimated using maximum likelihood (ML). Model fit was tested using the Comparative Fit Index (CFI) and the Tucker-Lewis index (TLI) as goodness of fit statistics, and the root-mean squared error of approximation (RMSEA) and the standardized root mean-squared residual (SRMR) as residual fit statistics. It is important to highlight that the closer the CFI and TLI values are to 1, and the closer the RMSEA and SRMR values are to 0, the better model fit. Acceptable model fit was given by CFI > .90; TLI > .90; RMSEA < .10; and SRMR < 0.08 as proposed by Hu and Bentler, (1999) and Rutkowski and Svetina (2014).

3.3.2 Cross-cultural Comparability

MG-CFA is a method widely used to test measurement invariance (Widaman & Rice, 1997; Vandenberg & Lance, 2000; Hair et al. 2010; Kline, 2011; Milfont & Fischer, 2015). MG-CFA is a continuation of classic CFA, and it is based on multi-group comparison. It divides the data into groups and determines the model fit for each one of them (Kline, 2011; Bialosiewicz, Murphy & Berry, 2013). MG-CFA is also widely used to test measurement invariance, where different levels of comparability must be explored, i.e., configural invariance, metric invariance, and scalar invariance (Kline, 2011; Vandenberg & Lance, 2000).

Configural invariance constitutes the first step when testing measurement invariance. It is associated with a model where the latent structure is equivalent across groups (Kline, 2011), i.e., the common factors and items measuring these factors are the same (Vandenberg & Lance, 2000). Although achieving this level of invariance does not mean that the groups are comparable, it is a prerequisite for testing other invariance levels (Kline, 2011).

Metric invariance implies that each group has equal factor loadings (Kline, 2011). If this level of invariance is verified, latent variances and covariances between latent variables can be compared (Millsap & Olivera-Aguilar, 2012). When metric invariance conditions are not met, that implies items/indicators do not have the same meaning across groups (Gregorich, 2006).

Scalar invariance must be verified after metric invariance has been tested. This level implies that item constants/intercepts are equivalent among groups (Millsap & Olivera-Aguilar, 2012) and that latent and observed variable

means are comparable (Kline, 2011; Gregorich, 2006). In other words, if scalar invariance conditions are met, this will allow us to compare the level of the latent variable among different education systems.

Finally, strict invariance is the last level of invariance that can be tested and implies that residual covariances are equivalent across groups (Brown, 2015). However, this last step was not taken into account in this study as the scalar level was considered sufficient to make meaningful comparisons of latent factors across education systems (Meredith, 1993).

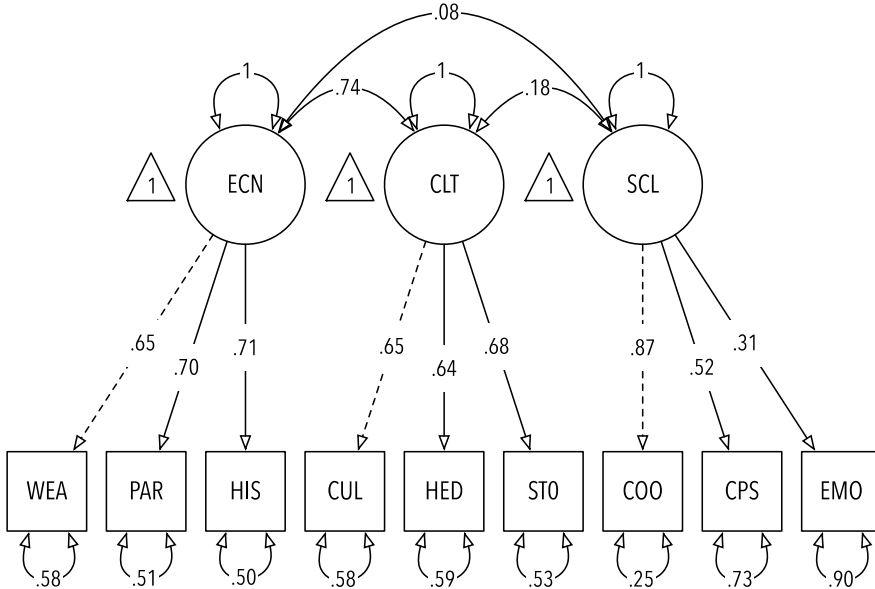
Two approaches to test measurement invariance are generally accepted in the literature: the chi-square (χ^2) test and changes in CFI and RMSEA statistics (Byrne & Stewart, 2006; Cheung & Rensvold, 2002). In this study, χ^2 , ΔCFI , $\Delta RMSEA$ were calculated and assessed. Using the chi-square test to decide on the overall model fit is said not to be reasonable in this context due to the large sample sizes (Rutkowski and Svetina, 2014). Therefore, ΔCFI and $\Delta RMSEA$ values were assessed in order to determine metric and scalar invariance, drawing on the criteria suggested by Rutkowski and Svetina (2014) when analysing large and variable sample sizes and a large number of groups. To determine metric invariance, these authors provide a slightly more liberal criterion of around -0.020 for ΔCFI and 0.030 for $\Delta RMSEA$. To determine scalar invariance, the traditional cut-off values were taken into consideration, i.e., -0.010 for ΔCFI and a $\Delta RMSEA$ of 0.010.

All analyses were executed in the R statistical software (R Core Team, 2019), using *lavaan* (Rosseel, 2012), and *lavaan.survey* (Oberski, 2014) packages.

4. Findings

First, an overall reliability estimate and CFA results are provided, as well as country-level reliability estimates and CFA results for a model that consists of economic (ECN), cultural (CLT) and social capital (SCL). Next, measurement invariance analysis results are presented considering the three abovementioned groups: Latin American countries, Asian countries, and European countries. Figure 1 shows overall CFA results and Table 4 shows country-level reliability estimates as well as country-level CFA models.

Figure 1 Measurement model including parameter estimates



The overall reliability was good (Cronbach's alpha = 0.7). Factor loadings ranged from 0.31 to 0.87, error variances ranged from 0.25 to 0.90 as shown in Figure 1. Results indicate that this model including all countries shows a weak fit to the data ($\chi^2 = 77538.351$; $DF = 24$; $CFI = 0.893$; $TLI = 0.84$; $RMSEA = 0.094$; $SRMR = 0.059$). However, this model was further considered in the analysis as it is a theory-based model.

With regard to country-level results, reliability estimates ranged from 0.77 (BSJG China) to 0.59 (The Netherlands). Whereas in OECD countries, the average reliability estimate was 0.66, ranging from 0.74 to 0.59, in partner countries, the average reliability was 0.69, ranging from 0.77 to 0.60.

Country-level CFA models are shown in Table 4. As can be seen, no country met the minimum TLI cut-off value of 0.90. For this reason, countries that satisfy the minimum criteria in three of the four fit measures are shown in bold. A total of 19 nations reached three fit measure cut-off values, of which 12 were OECD countries, and 7 were partner countries. Among the partner countries, there were three Asian countries (Chinese Taipei, Hong Kong and BSJG China) and three Latin American countries (Colombia, Costa Rica and the Dominican Republic), while there was only one European country (Russian Federation). All OECD countries were European countries, except for Korea.

Although there are education systems with relatively adequate fit both in OECD countries and in partner countries, there are still some education systems that do not show a good fit to the data. This evidence does not support cultural reproduction theory as an accurate model for some educational systems in this study. Particularly, the model poorly fitted in Canada (CFI=0.835 and TLI=0.752) and in New Zealand (CFI=0.826 and TLI=0.739).

Table 4. Reliability estimates and CFA results by country

Educational System	OECD	Reliability	CFI	TLI	RMSEA	SRMR	df	Chi-Square	n
Australia (36)	Yes	0.67	0.855	0.782	0.087	0.065	24	2290.451	12395
Austria (40)	Yes	0.63	0.889	0.833	0.081	0.061	24	1019.086	6364
Belgium (56)	Yes	0.62	0.896	0.843	0.076	0.055	24	1165.708	8273
Brazil (76)	No	0.7	0.899	0.849	0.088	0.051	24	3023.911	16237
Bulgaria (100)	No	0.69	0.882	0.823	0.09	0.061	24	944.037	4777
Canada (124)	Yes	0.66	0.835	0.752	0.094	0.068	24	3698.557	17372
Chile (152)	Yes	0.74	0.892	0.838	0.091	0.054	24	1272.31	6240
Chinese Taipei (158)	No	0.72	0.915	0.873	0.081	0.063	24	1091.014	6726
Colombia (170)	No	0.7	0.905	0.858	0.090	0.043	24	2038.137	10321
Costa Rica (188)	No	0.66	0.913	0.87	0.086	0.048	24	1007.429	5604
Croatia (191)	No	0.65	0.897	0.845	0.081	0.056	24	850.098	5190
Czech Republic (203)	Yes	0.66	0.89	0.834	0.077	0.055	24	896.191	6067
Denmark (208)	Yes	0.63	0.918	0.877	0.065	0.053	24	608.497	5697
Dominican Rep. (214)	No	0.68	0.926	0.889	0.078	0.041	24	550.861	3636
Estonia (233)	Yes	0.66	0.882	0.823	0.081	0.058	24	848.484	5251
Finland (246)	Yes	0.66	0.886	0.829	0.079	0.058	24	844.623	5496
France (250)	Yes	0.64	0.91	0.865	0.075	0.058	24	734.364	5279
Germany (276)	Yes	0.67	0.905	0.858	0.075	0.052	24	677.473	4853
Greece (300)	Yes	0.67	0.899	0.848	0.08	0.061	24	779.538	4933
Hong Kong (344)	No	0.73	0.917	0.876	0.081	0.060	24	752.816	4677
Hungary (348)	Yes	0.71	0.910	0.865	0.084	0.060	24	881.835	5027
Iceland (352)	Yes	0.64	0.895	0.843	0.071	0.055	24	402.752	3095
Ireland (372)	Yes	0.62	0.879	0.818	0.084	0.059	24	901.061	5201
Italy (380)	Yes	0.65	0.891	0.836	0.081	0.057	24	1705.972	10644
Japan (392)	Yes	0.63	0.885	0.827	0.08	0.052	24	902.483	5778
Korea (410)	Yes	0.73	0.913	0.87	0.081	0.06	24	858.347	5313
Latvia (428)	Yes	0.66	0.907	0.861	0.075	0.055	24	610.259	4390
Lithuania (440)	No	0.7	0.88	0.819	0.092	0.064	24	1133.432	5494
Luxemburg (442)	Yes	0.69	0.914	0.871	0.080	0.064	24	721.717	4546
Macao (446)	No	0.67	0.88	0.82	0.083	0.064	24	737.167	4268
Mexico (484)	Yes	0.71	0.888	0.831	0.103	0.051	24	1800.803	6977
Montenegro (499)	No	0.67	0.892	0.837	0.082	0.053	24	729.035	4379
Netherlands (528)	Yes	0.59	0.915	0.873	0.064	0.048	24	504.057	4913

Educational System	OECD	Reliability	CFI	TLI	RMSEA	SRMR	df	Chi-Square	n
New Zealand (554)	Yes	0.65	0.826	0.739	0.094	0.069	24	824.449	3773
Norway (578)	Yes	0.63	0.875	0.813	0.079	0.059	24	751.533	4886
Peru (604)	No	0.73	0.895	0.842	0.098	0.045	24	1528.132	6491
Poland (616)	Yes	0.64	0.86	0.789	0.091	0.067	24	859.031	4166
Qatar (634)	No	0.62	0.874	0.811	0.078	0.053	24	1393.175	9368
Russian Federat. (643)	No	0.67	0.911	0.866	0.072	0.052	24	668.919	5183
Singapore (702)	No	0.71	0.883	0.825	0.093	0.074	24	1217.337	5728
Slovak Republic (703)	Yes	0.71	0.893	0.84	0.08	0.055	24	845.719	5380
Slovenia (705)	Yes	0.67	0.907	0.86	0.076	0.055	24	833.518	5781
Spain (724)	Yes	0.69	0.907	0.86	0.079	0.057	24	945.932	6084
Sweden (752)	Yes	0.65	0.89	0.836	0.078	0.059	24	717.383	4706
Switzerland (756)	Yes	0.61	0.894	0.84	0.077	0.058	24	787.448	5319
United Arab Emirates (784)	No	0.6	0.847	0.77	0.083	0.058	24	1967.732	11679
Tunisia (788)	No	0.67	0.89	0.835	0.098	0.065	24	948.659	3988
Turkey (792)	Yes	0.73	0.88	0.819	0.105	0.065	24	1327.982	4959
United Kingdom (826)	Yes	0.66	0.877	0.816	0.085	0.061	24	1968.007	11157
United States (840)	Yes	0.72	0.863	0.795	0.102	0.068	24	1297.074	5060
Uruguay (858)	No	0.69	0.894	0.841	0.095	0.064	24	1066.445	4863
BSJG China (970)	No	0.77	0.913	0.87	0.094	0.057	24	1877.519	8760
Spain(regions) (971)	Yes	0.68	0.907	0.861	0.080	0.058	24	4554.095	29836
Portugal (620)	Yes	0.73	0.906	0.859	0.089	0.064	24	1326.639	6785

It is important to point out that a well-fitted CFA model is essential before examining measurement invariance. Although not all education systems showed a good fit, invariance analyses were carried out because most of the education systems did. Table 5 shows the baseline, configural, metric and scalar invariance models and their respective fit measures considering the 54 education systems. As can be observed, the baseline model showed fit indices slightly within acceptable levels. When moving from the baseline model to the configural model, the fit indices did not show much visible variation. Moving from the configural model to the metric model, the variation in fit indices was a minor. The change in RMSEA was of an acceptable level, while the change in CFI was not within the expected value. Moving to the scalar invariance model, fit indices worsened notoriously and changes in CFI and RMSEA were not acceptable. These results clearly indicate that neither metric nor scalar levels of invariance were reached, and thus it is not possible to compare latent variances, covariances and means across all participating countries.

Table 5. Measurement invariance results for all countries.

Level of invariance	Chi-Square	df	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA
Baseline	77538.35	24	0.893	0.840	0.094	0.059		
Configural	64501.252	1296	0.892	0.839	0.084	0.058		
Metric	94423.134	1614	0.842	0.810	0.092	0.076	-0.05	0.007
Scalar	276736.061	1932	0.532	0.53	0.144	0.118	-0.31	0.052

In the following stage, countries were grouped into three regions, namely, Latin America, Asia and Europe. Next, a MG-CFA was carried out within each group in order to explore whether measurement equivalence was supported.

Table 6 shows results for Latin American countries. The baseline model showed satisfactory fit indices. Moving to the configural model, there was a slight improvement in terms of fit indices. When moving from the configural to the metric level, it can be seen that the CFI value decreased from 0.89 to 0.86 and the RMSEA value increased from 0.092 to 0.095. These differences are just over those proposed by Rutkowski and Svetina (2014). Moving to the scalar model the CFI changed from 0.86 to 0.78 and RMSEA from 0.096 to 0.111. These results demonstrate that factor loadings and intercepts are not equivalent across Latin American countries, and thus no comparisons between latent variances, covariances and means can be made.

Table 6. Measurement invariance results for Latin American countries

Level of invariance	Chi-Square	df	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA
Baseline	12617.355	24	0.894	0.840	0.093	0.047		
Configural	12492.02947	192	0.899	0.850	0.092	0.049		
Metric	16450.02153	234	0.868	0.837	0.096	0.065	-0.032	0.004
Scalar	26032.8548	276	0.790	0.781	0.111	0.077	-0.078	0.015

In Asian countries (see Table 7) the baseline model fit indices were CFI = 0.88, TLI = 0.82 and RMSEA = 0.096. Moving from the baseline model to the configural model, there was an increase in fit indices from 0.88 to 0.89 for CFI, from 0.82 to 0.84 for TLI and from 0.096 to 0.084 for RMSEA. When moving from the configural to the metric model, the changes in CFI and in RMSEA were within acceptable levels. Using Rutkowski and Svetina (2014)'s criteria, results indicate that factor loadings are equivalent across Asian countries. CFI reduced from 0.87 to 0.66 and RMSEA increased from 0.083 to 0.124 when moving from the metric invariance model to the scalar

invariance model, which is higher than the expected values. Again, these results indicate that intercepts are not equivalent across Asian countries, and thus no comparison between latent means can be made.

Table 7. Measurement invariance results for Asian countries

Level of invariance	Chi-Square	df	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA
Baseline	14052.11418	24	0.885	0.828	0.096	0.061		
Configural	10900.12349	216	0.894	0.841	0.084	0.059		
Metric	12992.2109	264	0.873	0.845	0.0834	0.067	-0.020	-0.001
Scalar	33652.24698	312	0.669	0.656	0.124	0.099	-0.204	0.041

In European countries (see Table 8), the baseline model's fit indices were acceptable (CFI = 0.90, TLI = 0.86 and RMSEA = 0.080). Moving from the baseline model to the configural model, fit indices marginally worsened (CFI from 0.90 to 0.89 and TLI from 0.86 to 0.84). Similarly, when moving from the configural model to metric model, no considerable change in fit indices was observed. CFI reduced from 0.89 to 0.87 and RMSEA remained unchanged. These values are below those proposed by Rutkowski and Svetina (2014), which suggests that factor loadings are equivalent across countries. There was an extreme deterioration of model fit indices when switching from the metric model to the scalar model as changes in CFI and RMSEA were not of an acceptable level.

Table 8. Measurement invariance results for European countries

Level of invariance	Chi-Square	df	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA
Baseline	31526.0374	24	0.907	0.861	0.080	0.056		
Configural	31886.0435	768	0.897	0.845	0.079	0.057		
Metric	38967.8366	954	0.874	0.848	0.079	0.065	-0.022	-0.0006
Scalar	110087.353	1140	0.64	0.636	0.122	0.1	-0.234	0.043

5. Discussion

Identifying the differences in student academic achievement across countries is one of the main challenges facing education designers and practitioners who especially dedicate themselves to eliminate disparities among students across the world. Although the scale that measures the socio-economic background explains this difference to a great extent, the adequacy of this scale in explaining this difference remains to be discussed as there is a wide variety of groups in PISA. There occur still two main criticisms to studies based on secondary analyses of PISA in education, which are

this scale's lack of theoretical background to formulate the hypotheses that they test, and the alleged lack of comparability of this construct. Therefore, theoretically supporting the underlying mechanisms of SES and making valid comparisons of this measure across countries are essential requirements.

The primary purpose of this paper was to address these criticisms by using items included in the PISA 2015 student background questionnaire to create a SES scale based on Bourdieu's reproduction theory (i.e. latent variables measuring students' economic, cultural and social capital) and to test the measurement invariance of these constructs across PISA participating countries. In other words, this study aimed to develop a reasonable theory-based structure from PISA existing indicators and to examine the comparability of this theory-supported structure across countries.

Regarding the first criticism, we have formalized a model that consists of economic, cultural and social capitals considering cultural reproduction theory of Pierre Bourdieu. On the one hand, economic and cultural capital measures were selected based on this theory and a wide range of related previous studies. WEALTH, PARED, and HISEI indicators for economic capital and CULTPOSS, HEDRES and number of books for cultural capital showed higher factor loadings, which is similar to Caro et al. (2014)'s findings. On the other hand, COOPERATE, CPSVALUE and EMOSUPS indicators showed acceptable factor loadings for social capital. This factor, however, is a multi-dimensional concept that cannot be easily measured using the available data.

Our results showed that a construct derived from PISA's indicators did not support cross-cultural comparability across all countries, but just at the configural level. However, after countries were split into more homogeneous groups (Latin America, Asia, Europe), cross-cultural partial comparability was supported. Rutkowski and Rutkowski (2018) have pointed out that ILSAs include linguistically, geographically, economically, and culturally diverse participating countries. Therefore, they suggest that well-structured country-specific indicators should be produced rather than single indicators for all participating countries. This way, it would be possible for each participating country to incorporate their territorial conditions into comparable international scales (Rutkowski and Rutkowski, 2018; Sandoval-Hernandez et al., 2019).

The results of this study provide solutions and recommendations that should be considered and implemented. The analyses including all countries, do not support comparisons across education systems when using this socio-economic status scale, as neither the metric level of invariance nor the scalar level were reached. This may be partly related to regional and socio-cultural factors, as well as language as stated by Lee (2019) in her work focusing on the home possessions scale. Although the Latin American

group did not achieve the metric invariance level, results were close to acceptable values. In both Asian and European countries, the metric level was achieved but not at scalar invariance level. Sandoval-Hernandez et al. (2019) have highlighted that in TERCE - a much more regional assessment – the socio-economic background scale reached the metric level of invariance. Our suggestion goes in line with what Rutkowski and Rutkowski (2018) state, in that ILSAs would benefit from the “the active involvement of countries or regions to develop and include more country or regional options into the background questionnaire” (p.365).

6. Limitations

This study has limitations that should not be ignored. Firstly, it is worth mentioning that the variable ‘number of books’ is categorical. In order to carry out the analysis in a way that takes into account the survey design, variables must be continuous. However, Liu et al. (2017) state that if there are more than five response categories in ordered-categorical data, it may be acceptable to analyse them as continuous data. Since this variable has more than five response categories, it is reasonable to assume that there was no significant bias in parameter estimation.

Another limitation is that social capital is an indicator of socio-economic background, however, there are not enough items that capture and measure social capital in PISA. Therefore, this study encourages policy-makers and educational research designers to consider this and act towards this direction. Moreover, social capital is an extensive and multidimensional notion that comprises different dimensions such as structural, cognitive and relational factors. We have mostly conceptualised social capital using variables relating to interpersonal relationships and parental responsibility in education. Since indicators of social capital are limited, we could not focus on all aspects of this construct.

It is worth noting that this study was carried out to determine whether a SES scale was invariant and did not focus on the reasons triggering invariance. In this context, if measurement invariance is detected in a given step, successive analyses should be carried out to determine the reasons for this invariance before proceeding to the next stage.

7. Conclusion

This paper has supplied evidence that PISA indicators of socioeconomic background have serious psychometric deficiencies when used to elucidate differences in educational achievement across different educational systems. Further investigation on the comparability of other scales included in PISA's background questionnaires, such as teaching practices, could be carried out given the diversity of participating countries. Such studies are necessary because PISA's report provides information on such scales, and many researchers around the world use these variables to explain academic achievement. Making an evidence-based comparison among countries is undoubtedly a need for educators in each country.

As revealed in this study, when dividing countries into groups according to region/continent, comparability across education systems of some background scales could be supported by evidence. In that sense, two alternatives can be considered. On the one hand, ILSAs could use continent-specific or country-specific items for its background questionnaires. On the other hand, the process of developing background questionnaires could be benefitted from more heterogeneous groups of experts that represent different countries and languages. By adopting these suggestions, the necessity of designing assessments with a focus on specific regions will have been addressed.

References

- Abel, T. (2008). Cultural capital and social inequality in health. *Journal of Epidemiology & Community Health*, 62(7), e13-e13.
- Addey, C. & Sellar, S. (2018). Why Do Countries Participate in PISA? Understanding the Role of International Large-Scale Assessments in Global Education Policy. En Verger, A., Altinyelken, H.K. and Novelli, M. (Eds.), *Global education policy and international development: New agendas, issues and policies*(pp. 98-117). Bloomsbury Publishing.
- Addey, C., Sellar, S., Steiner-Khamisi, G., Lingard, B., & Verger, A. (2017). The rise of International large-scale assessments and rationales for participation. *Compare: A journal of Comparative and International Education*, 47(3), 434-452.
- Alivernini, F. (2011). Measurement invariance of a reading literacy scale in the Italian Context: a psychometric analysis. *Procedia-Social and Behavioral Sciences*, 15, 436-441.
- Aloisi, C., & Tymms, P. (2017). PISA trends, social changes, and education reforms. *Educational Research and Evaluation*, 23(5-6), 180-220.
- Barone, C. (2006). Cultural capital, ambition and the explanation of inequalities in learning outcomes: A comparative analysis. *Sociology*, 40(6), 1039-1058.

- Bialosiewicz, S., Murphy, K. & Berry, T. (2013). *An introduction to measurement invariance testing: Resource packet for participants*. Retrieved from <https://bit.ly/3iQoCl5>
- Bodovski, K., Jeon, H., & Byun, S. Y. (2017). Cultural capital and academic achievement in post-socialist Eastern Europe. *British Journal of sociology of Education*, 38(6), 887-907.
- Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education*. New York: Greenwood.
- Bourdieu, P. (1997). The Forms of Capital. In *Education: Culture, Economy, and Society*, edited by A. H. Halsey, H. Lauder, P. Brown, and A. Stuart Wells, 47-58. Oxford: Oxford University Press.
- Bornstein, M. C., & Bradley, R. H. (Eds.). (2003). *Socioeconomic status, parenting, and child development*. Mahwah, NJ: Lawrence Erlbaum.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53, 371-399.
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research*. New York: Guildford Press.
- Byrne, B. M., & Stewart, S. M. (2006). Teacher's corner: The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural equation modeling*, 13(2), 287-321.
- Byrne, B. M., & Van de Vijver, F. J. R. (2010). Testing for measurement and structural equivalence in large-scale cross-cultural studies: Addressing the issue of nonequivalence. *International Journal of Testing*, 10, 107-132.
- Caro, D. H., & Cortés, D. (2012). Measuring family socioeconomic status: An illustration using data from PIRLS 2006. *IERI Monograph Series Issues and Methodologies in Large-Scale Assessments*, 5, 9-33.
- Caro, D. H., Sandoval-Hernandez, A., & Lüdtke, O. (2014). Cultural, social, and economic capital constructs in international assessments: An evaluation using exploratory structural equation modelling. *School Effectiveness and School Improvement*, 25(3), 433-450.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modelling*, 9(2), 233-255.
- Coe, R., & Fitz-Gibbon, C. T. (1998). School Effectiveness Research: Criticisms and Recommendations. *Oxford Review of Education*, 24(4), 421-438.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A. M., Weinfeld, F. D., & York, R. L. (1966). *Equality of Educational Opportunity*. Washington, DC: US Government Printing Office.
- Coleman, J. S. (1988). Social capital in the creation of human capital. *The American Journal of Sociology*, 94(1) Supplement. Organizations and institutions: Sociological and economic approaches to the analysis of social structure, pp. 95-120.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Demir, E. (2017). Testing Measurement Invariance of the Students' Affective Characteristics Model across Gender Sub-Groups. *Educational Sciences: Theory and Practice*, 17(1), 47-62.

- DiMaggio, P. (1982). Cultural Capital and School Success: The Impact of Status Culture Participation on the Grades of U.S. High School Students. *American Sociological Review*, 47 (2), 189-201.
- Garcia Aracil, A. D. E. L. A., Neira, I., & Albert, C. (2016). Social and Cultural Capital Predictors of Adolescents' Financial Literacy: Family and School Influences. *Revista de Educación*, (374), 91-115.
- Gordon W. C. & Roger B., R. (2002). Evaluating Goodness-of-Fit Indexes for Testing Measurement Invariance, *Structural Equation Modeling*, 9(2), 233-255.
- Goldstein, H. (2017). Measurement and evaluation issues with PISA. En Louis Volante (Eds.), *In The PISA effect on global educational governance*. New York and London, Routledge.
- Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, 44, 78-94.
- Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2010). *Multivariate Data Analysis*. Seventh Edition. New Jersey: Prentice Hall, Upper Saddle River.
- Hansson, Å., & Gustafsson, J. E. (2013). Measurement invariance of socioeconomic status across migrational background. *Scandinavian Journal of Educational Research*, 57(2), 148-166.
- He, J., Barrera-Pedemonte, F., & Buchholz, J. (2018). Cross-cultural comparability of noncognitive construction in TIMSS and PISA. *Assessment in Education: Principles, Policy & Practice*, 1-17.
- Hobbs, G., & Vignoles, A. (2007). *Is free school meal status a valid proxy for socio-economic status (in schools research)?* (No. 84). Centre for the Economics of Education, London School of Economics and Political Science.
- Hopfenbeck, T. N., Lenkeit, J., El Masri, Y., Cantrell, K., Ryan, J., & Baird, J. A. (2018). Lessons learned from PISA: A systematic review of peer-reviewed articles on the programme for international student assessment. *Scandinavian Journal of Educational Research*, 62(3), 333-353.
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, 6(1), 1-55.
- Kalaycioglu, D. B. (2015). The Influence of Socioeconomic Status, Self-Efficacy, and Anxiety on Mathematics Achievement in England, Greece, Hong Kong, the Netherlands, Turkey, and the USA. *Educational Sciences: Theory and Practice*, 15(5), 1391-1401.
- Klieme, E. (2016). TIMSS 2015 and PISA 2015. How are they related on the country level. *German Institute for International Educational Research (DIPF)*. Retrieved from http://www.dipf.de/de/publikationen/pdf-publikationen/Klieme_TIMSS2015andPISA2015.pdf.
- Kline, R.B., (2011). *Principles and Practices of Structural Equation Modelling*. 3rd ed. New York: The Guilford Press.
- Lareau (2011). *Unequal childhoods: Class, race, and family life* (2nd ed.). Berkeley, CA: University of California Press.
- Lauder, H., Jamieson, I., & Wikeley, F. (1998). Models of Effective Schools: Limits and Capabilities. In R. Slee, G. Weiner, & S. Tomlinson (Eds.), *School Effectiveness for*

whom?: Challenges to the School Effectiveness and School Improvement Movements. London: Falmer Press.

- Lee, S. S. (2019). Longitudinal and Cross-Country Measurement Invariance of The Pisa Home Possessions Scale. *Publicly Accessible Penn Dissertations*. 3310. <https://repository.upenn.edu/edissertations/3310>
- Lenkeit, J., Caro, D. H., & Strand, S. (2015). Tackling the remaining attainment gap between students with and without immigrant background: An investigation into the equivalence of SES constructs. *Educational Research and Evaluation*, 21(1), 60-83.
- Lietz, P., Cresswell, J. C., Rust, K. F., & Adams, R. J. (2017). Implementation of Large-Scale Education Assessments. En Petra Lietz, John C. Cresswell, Keith F. Rust and Raymond J. Adams (Eds.) *Implementation of Large-Scale Education Assessments*. John Wiley & Sons.
- Liu, Y., Millsap, R. E., West, S. G., Tein, J. Y., Tanaka, R., & Grimm, K. J. (2017). Testing measurement invariance in longitudinal data with ordered-categorical measures. *Psychological methods*, 22(3), 486.
- Marteletto, L., & Andrade, F. (2014). The educational achievement of Brazilian adolescents: Cultural capital and the interaction between families and schools. *Sociology of Education*, 87(1), 16-35.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525-543.
- Milfont, T.L., & Fischer, R. (2015). Testing measurement invariance across groups: Applications in cross-cultural research. *International Journal of Psychological Research*, 3, 111-130.
- Millsap, R. E. & Olivera-Aguilar, M. (2012). Investigating measurement invariance using confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modelling* (pp. 380-392). New York, NY: Guilford Press.
- Miranda, D. and Castillo, J. C. (2018). Measurement model and invariance testing of scales measuring egalitarian values in ICCS 2009. En A. Sandoval-Hernandez, M. M. Isac and D. Miranda (Eds.) *Teaching Tolerance in a Globalized World*. Cham: Springer International Publishing
- Mullis, I. V. (2002). Background questions in TIMMS and PIRLS: An overview. *Paper commissioned by the National Assessment Governing Board*, <http://www.nagb.org/release/Mullis.doc>.
- Neff, W. S. (1938). Socioeconomic status and intelligence: A critical survey. *Psychological Bulletin*, 35(10), 727.
- Oberski, D. (2014). lavaan. survey: An R package for complex survey analysis of structural equation models. *Journal of Statistical Software*, 57(1), 1-27.
- OECD (2017). *PISA 2015 Technical Report*. Paris, France: OECD Publishing.
- OECD (2018). PISA 2015 Results in Focus. Available at: <https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf> Accessed 9 December 2019.
- Oliveri, M. E., & Ercikan, K. (2011). Do different approaches to examining construct comparability in multilanguage assessments lead to similar conclusions?. *Applied Measurement in Education*, 24(4), 349-366.

- Park, H., & Sandefur, G. D. (2006). Families, schools, and reading in Asia and Latin America. In *Children's Lives and Schooling across Societies* (pp. 133-162). Emerald Group Publishing Limited.
- Pitzalis, M., & Porcu, M. (2017). Cultural capital and educational strategies. Shaping boundaries between groups of students with homologous cultural behaviours. *British Journal of Sociology of Education*, 38(7), 956-974.
- Pokropek, A., Borgonovi, F., & Jakubowski, M. (2015). Socio-economic disparities in academic achievement: A comparative analysis of mechanisms and pathways. *Learning and Individual Differences*, 42, 10-18.
- Pokropek, A., Borgonovi, F., & McCormick, C. (2017). On the Cross-Country Comparability of Indicators of Socioeconomic Resources in PISA. *Applied Measurement in Education*, 30(4), 243-258.
- Puzic, S., Gregurović, M., & Košutić, I. (2016). Cultural capital—a shift in perspective: An analysis of PISA 2009 data for Croatia. *British journal of sociology of education*, 37(7), 1056-1076.
- Puzic, S., Gregurović, M., & Košutić, I. (2018). Cultural Capital and Educational Inequality in Croatia, Germany and Denmark: A Comparative Analysis of the PISA 2009 Data. *Revija za socijalnu politiku*, 25(2), 133-156.
- R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>.
- Rogošić, S., & Baranović, B. (2016). Social capital and educational achievements: Coleman vs. Bourdieu. *Center for Educational Policy Studies Journal*, 6(2), 81-100.
- Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of statistical software* 48(2), 1-36.
- Rutkowski, L., & Rutkowski, D. (2010). Getting it 'better': the importance of improving background questionnaires in international large-scale assessment. *Journal of Curriculum Studies*, 42(3), 411-430.
- Rutkowski, D., & Rutkowski, L. (2013). Measuring socioeconomic background in PISA: One size might not fit all. *Research in Comparative and International Education*, 8(3), 259-278.
- Rutkowski, L., & Svetina, D. (2014). Assessing the hypothesis of measurement invariance in the context of large-scale international surveys. *Educational and Psychological Measurement*, 74(1), 31-57.
- Rutkowski, L., & Rutkowski, D. (2018). Improving the comparability and local usefulness of international assessments: A look back and a way forward. *Scandinavian Journal of Educational Research*, 62(3), 354-367.
- Sandoval-Hernandez, A., Rutkowski, D., Matta, T., & Mirnda, D. (2019). Back to the drawing board: Can we compare socioeconomic background scales?. *REVISTA DE EDUCACION*, (383), 37-61.
- Segeritz, M., & Pant, H. A. (2013). Do they feel the same way about math? Testing measurement invariance of the PISA "students' approaches to learning" instrument across immigrant groups within Germany. *Educational and Psychological Measurement*, 73(4), 601-630.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research*, 75(3), 417-453.

- Steiner-Khamsi, G. (2019). Conclusions: What Policy-Makers Do with PISA. En Florian Waldow and Gita Steiner-Khamsi (Eds.) *Understanding PISA's Attractiveness Critical Analyses in Comparative Policy Studies*. Bloomsbury Publishing.
- Tan, C. Y. (2015). The contribution of cultural capital to students' mathematics achievement in medium and high socioeconomic gradient economies. *British Educational Research Journal*, 41(6), 1050-1067.
- Thein, L. M., & Ong, M. Y. (2015). Malaysian and Singaporean students' affective characteristics and mathematics performance: evidence from PISA 2012. *SpringerPlus*, 4(1), 563.
- Tijana, P. B., & Anna, S. (2015). *PISA The Experience of Middle-Income Countries Participating in PISA 200-2015*. OECD Publishing.
- Torney-Purta, J., & Amadeo, J. A. (2013). International large-scale assessments: Challenges in reporting and potentials for secondary analysis. *Research in Comparative and International Education*, 8(3), 248-258.
- UN (2015). *Transforming Our World: The 2030 Agenda for Sustainable Development*, 2nd August 2015. New York: United Nations.
- Uysal, N. K., & Arikan, Ç. A. (2018). Measurement Invariance of Science Self-Efficacy Scale in PISA. *International Journal of Assessment Tools in Education*, 5(2), 325-338.
- Vanderberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions Practices, and Recommendations for Organizational Research. *Organizational Research Methods*, 3(1), 4-70.
- Van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis for cross-cultural research*. Thousand Oaks, CA: Sage.
- Van de Vijver, F. J. R., & Poortinga, Y. H. (2002). Structural equivalence in multilevel research. *Journal of Cross-Cultural Psychology*, 33, 141-156.
- Van de Vijver, F. J. (2018). Towards an Integrated Framework of Bias in Noncognitive Assessment in International Large-Scale Studies: Challenges and Prospects. *Educational Measurement: Issues and Practice*, 37(4), 49-56.
- White, K. R. (1982). The relationship between socioeconomic status and academic achievement. *Psychological Bulletin*, 91, 461-481.
- Widaman, K. F., Reise, S. P., Bryant, K. J., & Windle, M. (1997). Exploring the measurement invariance of psychological instruments: Applications in substance use domain. *Ariel*, 165, 220-14.
- Wu, A. D., Li, Z., & Zumbo, B. D. (2007). Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data. *Practical Assessment, Research & Evaluation*, 12(3), 1-26.
- Wu, M. (2010). Comparing the Similarities and Differences of PISA 2003 and TIMSS. *OECD Education Working Papers*, 32, Paris: OECD Publishing. <https://doi.org/10.1787/5km4psnm13nx-en>
- Yang, Y. (2003). Dimensions of socio-economic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian Journal of Educational Research*, 47(1), 21-41.

Uso de TERCE-UNESCO para informar la práctica educativa: factores que influyen en el aprendizaje en Ciencias en Chile y Paraguay

Soledad Miranda ¹  ; Ernesto Treviño ¹ 

¹ Pontificia Universidad Católica de Chile (UC), Chile.

Resumen. Las evaluaciones internacionales proveen información valiosa más allá de la atención que acaparan los rankings del logro promedio por país que aparecen en la prensa. Los estudios de factores asociados al logro aportan evidencia para la toma de decisiones en educación. Diversas mediciones internacionales levantan información de este tipo, incluso para Latinoamérica. Pese a esto, pocos estudios han abordado comparativamente dentro de la región los factores que influyen el aprendizaje en ciencias. El presente estudio examinó los efectos de factores personales y prácticas docentes que afectan el logro en ciencias en estudiantes de Chile y Paraguay, usando los datos del TERCE. Mediante un conjunto de modelos multinivel, controlando por variables socioeconómicas, se identificó que las variables repitencia y asistencia a preescolar son predictores del logro en ambos países. Por otro lado, las prácticas docentes presentaron resultados diversos, pues para Chile sólo la inasistencia del docente resultó ser un predictor, mientras que en Paraguay todas las prácticas medidas mostraron una influencia significativa. De esta forma, se identificaron similitudes entre los factores que impactan en distintos países, y consecuentemente, generar sugerencias de política educativa que apunten hacia una mejora de la educación científica en Latinoamérica.

Palabras clave: educación científica; factores personales; prácticas docentes efectivas; modelo multinivel.

Uso do TERCE-UNESCO para informar a prática educacional: fatores que influenciam a aprendizagem de Ciências no Chile e no Paraguai

Resumo. As avaliações internacionais fornecem informações valiosas, além da atenção que recebem os rankings de desempenho médio por país que aparecem na imprensa. Estudos de fatores associados ao desempenho fornecem evidências para a tomada de decisões na educação. Diversas medições internacionais levantam informações deste tipo, mesmo para a América Latina. Apesar disso, poucos estudos abordaram comparativamente dentro da região os fatores que influenciam o aprendizado de ciências. O presente estudo examinou os efeitos de fatores pessoais e práticas de ensino que afetam o desempenho em ciências em estudantes do Chile e do Paraguai, usando dados do TERCE. Por meio de um conjunto de modelos multiníveis, controlando por variáveis socioeconômicas, identificou-se que as variáveis repetência e frequência à pré-escola são preditores do aproveitamento em ambos os países. Por outro lado, as práticas de ensino apresentaram resultados diferentes, uma vez que para o Chile apenas a ausência do professor acabou por ser um preditor, enquanto no Paraguai todas as práticas medidas mostraram uma influência significativa. Dessa forma, foram identificadas semelhanças entre os fatores que impactam os diferentes países e, consequentemente, gerar sugestões de políticas educacionais que apontem para uma melhoria na educação científica na América Latina.

Palavras-chave: educação científica; fatores pessoais; práticas de ensino eficazes; modelo multinível.

Use of TERCE-UNESCO to report the education practice: factors that affect learning in Science in Chile and Paraguay

Abstract. International large-scale assessments offer valuable information beyond the rankings of countries according to their average achievement that is regularly picked up by the media. Associated factor studies allow producing important evidence for decision-making in education. Several international assessments continuously collect data on educational outcomes and its explanatory factors, even for Latin America. Despite that, there are limited studies within the region that research on the factors that explain student learning in science. This study analyzes the influence of different teaching practices on learning achievement in 6th grade for Chile and Paraguay using data from TERCE. A set of multilevel models are fitted, controlling for the socioeconomic status of students and schools; the study found that grade retention and attending preschool were predictors of science achievement in both countries. In Chile, only teacher's absence from class was a significant predictor, while in Paraguay different types of classroom practices showed an important influence on achievement. In this way, we found similarities and differences in school factors related to achievement and, consequently, propose policy recommendations to improve science education in Latin America.

Keywords: science education; personal factors; effective teaching practices; multilevel models.

1. Introducción

La educación científica se ha convertido en un tema central en las últimas décadas. Esta busca promover la formación de personas que sean capaces comprender el mundo que los rodea, usar el conocimiento y la información de manera integrada, combinándola con sus habilidades para resolver una diversidad de problemas (OECD, 2017). Así, la educación en ciencias se vuelve una herramienta fundamental para comprender y tomar decisiones informadas relativas a los grandes desafíos actuales, tales como: el cambio climático, el desarrollo sostenible, obesidad y pandemias.

Mejorar la educación científica requiere de evidencia sólida que permita distinguir los factores que, en distintos niveles, y particularmente en aula, tienen una asociación robusta con el logro. Es por ello que los países invierten importantes recursos en educación científica, en la formación de docentes del área, en el diseño de nuevos programas, en la incorporación de infraestructura, etc. con el objetivo de que los estudiantes adquieran los aprendizajes y competencias necesarias (Rivas, 2015). Pese a lo anterior, algunos países de Latinoamérica y el Caribe no han logrado avanzar sustantivamente en lo referido al logro en ciencias. Esto se ha evidenciado en diversas evaluaciones internacionales donde estos países tienden a presentar bajos niveles de rendimiento.

136

Para conocer en profundidad la realidad latinoamericana y obtener información contextualizada acerca de los factores que influyen en el aprendizaje de los estudiantes de cada país, es importante trabajar con estudios que pongan foco en la región. La medición más reciente para países de América Latina y el Caribe que combina evaluación del logro, factores personales del estudiante y prácticas docentes, es el Tercer Estudio Regional Comparativo y Explicativo (TERCE). Los resultados de este estudio aportan antecedentes relevantes respecto de los aprendizajes de los estudiantes de la región, sin embargo, estos no se han utilizado de manera suficiente, menos desde el área de ciencias, que tal como se ha planteado anteriormente es un área educativa de gran importancia.

Los resultados de este estudio muestran que existe alta variabilidad en los resultados de los países de la región y que no todos adquieren los mismos conocimientos. En este sentido, se puede mencionar a países como Chile que es uno de los países latinoamericanos que en esta medición ha presentado un mejor desempeño (UNESCO, 2015). Aunque no es menos cierto que en términos internacionales está muy por debajo del promedio de los países desarrollados (Agencia de Calidad de la Educación, 2017). En el otro extremo podemos encontrar a países como Paraguay, que presenta una

realidad muy distinta. A pesar de tener una economía estable y ser el único país bilingüe de América Latina, con una diversidad y riqueza cultural, es uno de los países de la región con más bajos logros (UNESCO, 2015).

Considerando esta variabilidad entre países pertenecientes a un contexto común, en este artículo se explota la riqueza de información contenida en esta medición regional, a través del estudio de los factores que influyen en el desempeño en estudiantes de Chile y Paraguay. En particular, se propone un método multinivel que busca comprender con mayor profundidad las variables que inciden en los resultados de aprendizaje en el área de ciencias en ambos países e identificar aquellos elementos que prometen mayor eficacia para mejorar la calidad de la educación de los países estudiados. Así, el uso de evaluaciones a gran escala como TERCE permite responder las siguientes preguntas: ¿En qué medida factores personales y de prácticas docentes se relacionan con el aprendizaje en ciencias? ¿Cuáles son las similitudes y diferencias entre Chile y Paraguay?

Las siguientes secciones de este documento se ordenan de la siguiente forma. En la sección 2 se presenta la revisión de la literatura, donde se exponen factores del estudiante y prácticas docentes vinculadas con el aprendizaje escolar. La metodología es expuesta en la sección 3. Posteriormente, en la sección 4, se presentan los resultados para cada país y sus respectivas comparaciones. Finalmente, conclusiones y futuros trabajos se exhiben en la sección 5.

2. Revisión de literatura

Variadas investigaciones del área educativa convergen en la idea de que la educación científica es fundamental para mejorar las oportunidades de aprendizaje. La educación en ciencias brinda a los estudiantes la oportunidad de comprender el mundo y posibilita el desarrollo del razonamiento y de actitudes que les permiten una mejor integración y respuesta a las demandas de la sociedad actual (OREALC-UNESCO, 2016). Para poder enfrentar exitosamente los desafíos actuales y futuros, es necesario que los estudiantes puedan desarrollar múltiples habilidades que les permitirán hacerlo. En esta línea, Bybee (2010) indica que la educación científica permite el desarrollo: de la capacidad de adaptarse a distintas situaciones; de la capacidad de resolución de problemas no rutinarios, de la autogestión y autodesarrollo, además de la habilidad de pensar en forma sistémica o integrada. En suma, la comprensión de la ciencia es fundamental para la preparación para la vida de los jóvenes (OECD, 2017).

En las últimas décadas ha cambiado el enfoque de educación científica, iniciándose una transición hacia un paradigma constructivista donde los docentes deben practicar la enseñanza de la ciencia más centrada en el alumno, considerando que el aprendizaje es activo y socialmente constructivo (Areepattamannil, Freeman y Klinger, 2011). Este cambio de enfoque se vincula también con una serie de prácticas que favorecen el pensamiento crítico y la adquisición de habilidades para resolver problemas. Dentro de estas prácticas se cuenta la participación de estudiantes en investigación científica; en instancias de interacción dialógica y argumentativa; y la inclusión de actividades prácticas en la enseñanza de las ciencias (Blanchard *et al.* 2009; Hofstein *et al.* 2008; McNeill y Pimentel 2009).

Pese a esta evidencia, no existen mediciones observacionales que permitan evaluar prácticas específicas de la enseñanza de las ciencias en Latinoamérica. En este sentido, evaluaciones como el TERCE, permiten realizar acercamientos a los procesos de las salas de clases, brindando información relativa a la calidad de estos procesos. Dicha información, junto a datos de contexto, permiten estudiar la relación entre prácticas docentes generales y el logro educativo en ciencias. A continuación, se presentan las variables que se deben tener en consideración para realizar un análisis apropiado de la asociación entre prácticas y logro.

2.1 Antecedentes del estudiante y su relación con el aprendizaje

En términos generales, la condición socioeconómica y cultural es uno de los factores que más evidencia acumula como un importante predictor sobre el aprendizaje de los estudiantes (UNESCO, 2015). Sin embargo, es claro que existe una gama amplia de otros factores que pueden afectar el aprendizaje. Esto es de especial importancia en el área de ciencias, donde la literatura ha mostrado que el estatus socioeconómico predice en menor grado el rendimiento, y que por lo tanto, en dicha área la escuela tiende a ser un elemento más poderoso para predecir el rendimiento (Heyneman, 2004). En adelante, se analizan los principales antecedentes del estudiante que se han asociado con el logro educativo, además se aborda la importancia del conocimiento científico y de las de prácticas docentes efectivas que empíricamente se han vinculado con el aprendizaje de los estudiantes.

Durante décadas la evidencia empírica ha mostrado que el género es un elemento vinculado con rendimiento escolar. Distintas evaluaciones de aprendizaje en ciencias muestran que las mujeres tienden a obtener menores puntajes de que los hombres (OCDE, 2016; UNESCO, 2015). Esta situación, aunque no es generalizada para todos los países, debe mejorarse. Para esto es necesario un trabajo explícito y deliberado desde todos los niveles de la

política y la práctica educativa para cerrar las disparidades de aprendizaje asociadas al sexo, puesto que estas se relacionan con prácticas de socialización que las escuelas suelen transmitir inadvertidamente (Aikman y Rao, 2012).

Otro antecedente determinante, es la asistencia a educación preescolar. De acuerdo con la literatura, la educación inicial es un gran igualador de oportunidades que no sólo trae beneficios de corto plazo a nivel cognitivo o socioemocional, sino que también trae grandes beneficios y oportunidades posteriores (Vegas y Santibañez, 2010). En América Latina y el Caribe la participación en educación preescolar presenta una alta heterogeneidad, tal como se evidencia en los casos de los países considerados para este estudio, donde en Chile, la tasa neta de matrícula en educación preescolar es superior al 80% y en Paraguay está tasa no supera el 40% (UNESCO, 2015). Poner énfasis en la educación preescolar es fundamental, dado que durante esta etapa se construyen las oportunidades que abrirán las puertas a beneficios de largo plazo, particularmente para niños provenientes de hogares con bajo capital cultural (Treviño, Aguirre y Varela, 2018).

La repitencia es otra variable que resulta tener un impacto negativo sobre el aprendizaje, mostrando que los estudiantes que enfrentan esta situación obtienen puntuaciones menores (UNESCO, 2015). El fenómeno de la repetición de curso es una práctica común en la región, llevada a cabo bajo el supuesto de que repetir un año escolar puede mejorar el aprendizaje de un estudiante. Contrario a esto, la evidencia muestra que la repetición aparece como el primer paso hacia, y el mejor predictor de la deserción escolar (UNICEF, 2007). Como contramedida de este efecto, es necesario que los sistemas escolares pongan en marcha estrategias que atiendan a la particularidad del alumno, para promover las competencias sociales y cognitivas y facilitar el éxito académico de los estudiantes (Méndez y Cerezo, 2018).

Un último factor individual considerado para este estudio, es el ausentismo de los estudiantes. Esta problemática resulta importante pues tiene relación con el tiempo efectivo de aprendizaje en la escuela. La inasistencia escolar tiende a ser prevalente entre estudiantes de contextos desaventajados y puede contribuir a la brecha socioeconómica en resultados de pruebas (Cook *et al.*, 2017). Los procesos de aprendizaje constituyen elementos básicos de transformación y superación de la exclusión social, en consecuencia, los alumnos que no asisten regularmente a la escuela sufrirán una merma en este proceso de formación (González, 2006; González, 2016).

Considerando la literatura relativa a los factores personales, a través de este estudio se busca validar las siguientes hipótesis: las niñas obtienen en ciencias menos puntaje que los niños; la Educación preescolar afecta positivamente en el logro; la repitencia y el ausentismo son factores que afectan de manera negativa el aprendizaje en ciencias.

2.2 Calidad y prácticas docentes efectivas

La literatura ha demostrado que el aporte que las escuelas pueden hacer en relación con el aprendizaje de las ciencias en particular, es mayor que el que pueden aportar en otras áreas como lenguaje (Heyneman, 2004). En consecuencia, temas como la efectividad docente cobran especial relevancia por el impacto que pueden provocar en el desempeño de los estudiantes.

El estudio de la efectividad docente es un factor con abundante evidencia teórica, y empírica a su favor. En las últimas décadas, múltiples investigaciones han abordado este tema, en especial desde que Barber y Mourshed (2008) plantearon la idea de que ningún sistema educativo puede lograr resultados de aprendizaje sin preocuparse de la calidad de sus docentes. Avances significativos han surgido respecto de este tema, en especial en lo referido a prácticas docentes efectivas. Estas son entendidas como aquellas interacciones que ocurren en la clase, esenciales para el aprendizaje en la medida que éstas configuran un ambiente grato y promueven la estimulación cognitiva y lingüística (Hunt, 2009; Pianta y Hamre, 2009). En la medida que estas ocurren se transforman en una fuente importante de variación en el aprendizaje de los estudiantes (Creemers y Kyriakides, 2008).

Las prácticas docentes efectivas en enseñanza se han tratado principalmente como la estructura y naturaleza de las interacciones profesor-alumno en las aulas, organizadas en tres dominios: apoyo emocional, organización del aula y apoyos instructivos (Downer *et al.*, 2015; Hamre y Pianta, 2007).

Con apoyos emocionales se enfatiza en la habilidad de los maestros para apoyar el funcionamiento social y emocional en el aula. Los aspectos centrales de este ámbito son la retroalimentación positiva y constructiva por parte de los docentes; un enfoque positivo de los errores de los estudiantes; y el comportamiento cuidadoso del maestro (Klieme *et al.*, 2009; Pianta *et al.*, 2012). La organización del aula considera los procesos relacionados con la organización y gestión del comportamiento, el tiempo y la atención de los estudiantes en el aula (Emmer y Stough, 2001). Esta es una característica crítica del medio ambiente, con vínculos directos con una gama de resultados sociales y académicos que aumenta los niveles de compromiso de los estudiantes y disminuye los comportamientos de oposición (Pianta y Hamre, 2009).

Finalmente, los apoyos instructivos refieren a la forma en que docentes implementan el currículo y las actividades de aprendizaje, orientando al desarrollo cognitivo y académico. Se espera que estos apoyos fomenten el compromiso cognitivo y a su vez conduzcan a un conocimiento elaborado (Klieme *et al.*, 2009). Los profesores que usan estrategias de apoyo instruccional tienden a tener estudiantes que obtienen mayores logros, pues enfocan a sus estudiantes en habilidades de pensamiento de orden superior a través de la retroalimentación oportuna y desarrollo las habilidades lingüísticas (Hamre y Pianta, 2007).

Adicionalmente, el uso del tiempo para la enseñanza, es otro aspecto importante vinculado al logro educativo. Este refiere al tiempo que los estudiantes están “expuestos” a actividades de aprendizaje, más allá de los tiempos de organización y orden (Scheerens, 2013). La proporción de tiempo de clase que dedican los docentes a la instrucción, es uno de los factores que más incide en el aprendizaje (Bruns y Luque, 2014). En el plano internacional es variada la literatura que respalda dicha relación (Bruns y Luque, 2014; Cueto *et al.*, 2008; Scheerens, 2013). De la misma manera, este vínculo se ha visto evidenciado en el último estudio TERCE realizado en la región, donde se muestra una asociación positiva entre el uso del tiempo y el logro académico de los alumnos. Esta variable además se encuentra íntimamente relacionada con la asistencia y puntualidad docente (UNESCO, 2015). Esta evidencia releva la importancia del manejo del tiempo y la adecuada planificación por su incidencia sobre el aprendizaje de los alumnos.

141

En línea con la literatura presentada, este estudio incorporará como dimensiones de prácticas docentes: apoyo emocional, organización del aula, apoyos instructivos y asistencia docente. Teniendo como hipótesis que los estudiantes obtienen mejores logros cuando los docentes implementen estas prácticas. Esto permitirá, además, generar un contraste entre las prácticas docentes ejercidas en cada país.

3. Métodos

3.1 Muestra e instrumento

Los datos utilizados en esta investigación se obtuvieron desde el Tercer Estudio Regional Comparativo y Explicativo (TERCE), cuyas bases de datos son de libre disposición. La investigación se enfoca en los resultados de aprendizaje en 6° básico, debido a que la disciplina de interés de este estudio sólo fue evaluada en este nivel. En forma adicional a las pruebas de evaluación de aprendizaje, el TERCE utiliza cuestionarios de contexto que son

respondidos por estudiantes, familias, profesores y directores. La información que ofrecen estos instrumentos permite realizar análisis de factores asociados como el que se pretende realizar en el presente estudio.

El TERCE usa un muestreo complejo, bi-etápico y estratificado. En él primero se seleccionaron aleatoriamente escuelas en los respectivos estratos identificados en cada país. En una segunda etapa se elige aleatoriamente un aula intacta con todos los estudiantes dentro de cada escuela (UNESCO - OREALC, 2016). Por este motivo, en los modelos multinivel se utilizó el peso final de los estudiantes como factor de expansión para proyectar las estimaciones a nivel poblacional.

Finalmente, en esta investigación se utilizaron las muestras de dos países, la primera compuesta por todos los estudiantes de Chile que participaron de la medición y la segunda conformada por los estudiantes de Paraguay. La muestra final de estudiantes de sexto grado para Chile es de un total de 5.029 estudiantes provenientes de 197 escuelas. Mientras que en el caso de Paraguay la muestra fue de 3.231 provenientes de 204 escuelas (UNESCO - OREALC, 2016).

3.2 Variable dependiente

La variable dependiente corresponde al puntaje de logro en ciencias. Para evaluar dicho logro se implementó un diseño de cuadernillos rotados en que no todos los estudiantes responden todas las preguntas. Para estimar el puntaje se utilizaron modelos Rasch que generan cinco “valores plausibles” para calcular el puntaje obtenido en la prueba. Para analizar dicha variable, se utilizan procedimientos de estimación con diseño complejo lo que permite obtener el promedio internacional y su respectiva varianza (Promedio= 700, Error Estándar=1.07) (UNESCO, 2015).

3.3 Variables independientes

Las variables independientes se componen de variables individuales del estudiante como índice socioeconómico, género, asistencia a nivel preescolar e inasistencia a la escuela. Y también de variables en relación con las prácticas docentes declaradas por los estudiantes. La descripción de las variables la podemos ver en la Tabla 1.

Tabla 1. Variables independientes

Ítem/Pregunta	Niveles de la variable
Variables individuales del estudiante	
Género (nina)	0 'Niño' 1 'Niña'
Repite (repite)	0 'No'
Pregunta: ¿Cuántas veces has repetido de curso?	1 'Si'
Escala de respuesta:	
1 'Nunca he repetido'	
2 'Una vez'	
3 'Dos veces o más'	
4 'No sé, no recuerdo'	
Asistencia a la educación inicial entre los 4 y 6 años (prekfor6)	0 'No asistió entre los 4 y 6 años' 1 'Asistió a un centro formal entre los 4 y 6 años'
Inasistencia a clases (inasclas)	0 'Menos de una vez al mes'
¿Con qué frecuencia el estudiante ha faltado al colegio en los últimos seis meses?	1 'Más de una vez al mes'
Escala de respuesta:	
1 'Nunca'	
2 'Un par de veces al semestre'	
3 'Un par de veces al mes'	
4 'Una vez por semana'	
5 'Varios días por semana'	
6 'No sé'	
Variables referidas a prácticas docentes	
Índice socioeconómico del estudiante centrado al promedio a de la escuela (Isecf est-país)	
Índice de Apoyo emocional (apoyo-emo)	Min -1.53
Pregunta: ¿Con qué frecuencia ocurren estas cosas en tu clase?	Max 0.33
DQA6IT17_07 Los profesores están contentos de hacernos clase	
DQA6IT17_08 Los profesores nos felicitan cuando hacemos algo bien	
DQA6IT17_09 Los profesores nos motivan para que sigamos estudiando	
DQA6IT17_10 Los profesores me animan cuando encuentro difícil la materia	
DQA6IT17_11 Los profesores son simpáticos conmigo	
DQA6IT17_12 Los profesores escuchan con atención cuando hago algún...	
DQA6IT17_13 Los profesores nos explican con paciencia	
DQA6IT17_14 Los profesores llegan con las clases bien preparadas	

Ítem/Pregunta	Niveles de la variable
Variables referidas a prácticas docentes	
Índice socioeconómico del estudiante centrado al promedio a de la escuela (lsecf est-país)	
DQA6IT17_16 Los profesores se preocupan de que aprovechemos el tiempo...	
DQA6IT17_20 Los profesores me preguntan qué entendí y qué no	
DQA6IT17_21 Si no entendemos algo, los profesores buscan otras formas...	
DQA6IT17_22 Si me equivoco, los profesores me ayudan a ver mis errores	
Escala de respuesta:	
1 Nunca o casi nunca	
2 A veces	
3 Siempre o casi siempre	
Apoyo pedagógico (apoyo-ped)	Min -1.05
Pregunta: ¿Con qué frecuencia ocurren estas cosas en tu clase?	Max 0.35
DQA6IT17_26 Los profesores me piden que explique mis respuestas	
DQA6IT17_27 Los profesores usan nuestros propios ejercicios o...	
DQA6IT17_28 Cuando los profesores nos hacen preguntas esperan que les...	
DQA6IT17_29 Nuestros profesores esperan que hagamos los ejercicios de la...	
Escala de respuesta:	
1 Nunca o casi nunca	
2 A veces	
3 Siempre o casi siempre	
Asistencia del docente (asist-prof)	Min -0.33
Pregunta: ¿Con qué frecuencia ocurren estas cosas en tu clase?	Max 1.13
DQA6IT17_04 Los profesores faltan a clases	
DQA6IT17_05 Los profesores llegan tarde a clases	
DQA6IT17_06 Los profesores se van temprano	
Escala de respuesta:	
1 Nunca o casi nunca	
2 A veces	
3 Siempre o casi siempre	
Organización de las clases (organización)	Min -0.53
Pregunta: ¿Con qué frecuencia ocurren estas cosas en tu clase?	Max 0.89
DQA6IT17_01 Hay ruido y desorden	
DQA6IT17_02 Se pone atención cuando los profesores hablan	
DQA6IT17_17 Cuando pasamos de una actividad a otra, se arma desorden...	
Escala de respuesta:	
1 Nunca o casi nunca	
2 A veces	
3 Siempre o casi siempre	

Ítem/Pregunta	Niveles de la variable
Variables de dependencia del establecimiento	
Índice socioeconómico de la escuela centrado al promedio del país (Isecf esc país)	
Tipo de escuela (factor(tipoesc)2/factor(tipoesc)3)	1 'Rural' 2 'Urbana-Pri- vada' 3 'Urbana-Pu- blica'

Fuente: Elaboración propia.

4. Proceso de estimación

Para este estudio se utilizaron regresiones lineales multinivel, donde se considera nivel estudiante y nivel escuela. Este tipo de modelo es pertinente para esta investigación debido a que nos permite separar la varianza explicada por factores del estudiante de la varianza explicada por los colegios (ejemplo: prácticas y nivel socioeconómico). Se diseñaron 6 modelos de regresión multinivel, que incluyen variables socioeconómicas, variables individuales del estudiante, variables de la escuela y variables relacionadas con las prácticas ejercidas por las docentes declaradas por los estudiantes. Se aplicó un modelo por cada valor plausible disponible en TERCE. La estimación final de los parámetros se realizó mediante la regla de Rubin, que pondera los valores de los parámetros obtenidos por cada valor plausible (Rubin y Schenker, 1986).

Se generaron índices para generar las siguientes variables independientes: apoyo pedagógico, apoyo emocional, asistencia y rganización. Para esta construcción se utilizó un análisis factorial exploratorio. Las preguntas incluidas en cada factor son presentadas en la tabla 1.

Para realizar este estudio se utilizó R-statistics, en particular se utilizaron las librerías “lme4” (Bates *et al.*, 2008).

5. Resultados

A continuación, se presentan las tablas de los modelos construidos para cada país y sus respectivos resultados. La tabla 2 presenta los resultados de Chile y la tabla 3 presenta los resultados de Paraguay.

Tabla 2. Resultados modelos multinivel para Chile

	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
(intercepto)	768,18*** (1,24)	782,78*** (1,13)	776,68*** (1,40)	786,30*** (1,24)	779,23*** (1,55)	781,13*** (2,00)	776,43*** (7,68)
Isecf. Est- escuela.		0.060*** (0.009)	0.062*** (0.009)	0.059*** (0.010)	0.048*** (0.012)	0.046*** (0.011)	0.046*** (0.011)
		12,12*** (1,87)	12,37*** (1,85)	11,89*** (2,13)	9,75*** (2,42)	9,25*** (2,37)	9,34*** (2,33)
Niña (vs niño)			0.051*** (0.004)	0.034*** (0.004)	0.039*** (0.006)	0.032*** (0.007)	0.033*** (0.007)
			12,10*** (1,06)	8,12*** (1,06)	9,24*** (1,56)	7,69*** (1,66)	7,98*** (1,66)
Repite (vs No repite)				-0.142*** (0.004)	-0.137*** (0.004)	-0.136*** (0.004)	-0.131*** (0.004)
				-48,71*** (1,68)	-46,84*** (1,59)	-46,56*** (1,66)	-44,84*** (1,65)
Inasistencia a clases					-0.022 (0.012)	-0.024 (0.013)	-0.024 (0.013)
					-6,99 (3,82)	-7,42 (4,01)	-7,60 (4,02)
Educación preescolar					0.056*** (0.014)	0.058*** (0.013)	0.058*** (0.013)
					13,72*** (3,41)	14,24*** (3,36)	14,15*** (3,26)
Prácticas de prácticas docentes							
Apoyo emocional						0.028 (0.024)	0.024 (0.024)
						9,45 (8,00)	8,28 (8,04)
Apoyo pedagógico						0.012 (0.020)	0.008 (0.021)
						5,79 (9,49)	3,79 (9,76)
Inasistencia profesor						-0.132*** (0.013)	-0.130*** (0.012)
						-61,45*** (6,13)	-60,37*** (6,02)
Organización clases						-0.007 (0.017)	-0.008 (0.017)
						-3,88 (8,43)	-4,30 (8,56)
Variables de la escuela							
Isecf. Esc-país.		0.430*** (0.011)	0.429*** (0.011)	0.402*** (0.010)	0.382*** (0.011)	0.372*** (0.010)	0.322*** (0.025)
		66,74*** (1,80)	66,53 (1,78)	62,42*** (1,56)	59,24*** (1,82)	57,80*** (1,66)	50,00*** (3,96)
Particular urbano							0.060 (0.047)
							10,75 (8,43)

	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
Público urbano							-0,054 (0,022) -9,71 (3,92)
Efectos aleatorios							
s2	9592,58	9654,24	9646,73	9373,90	9304,6	9080,50	9075,08
too	4418,34	3459,05	4424,30	4070,31	3791,3	3662,38	1512,55
ICC	0,31	0,32	0,31	0,30	0,30	0,29	0,14
N	197	195	195	195	194	194	194
Obs	5029	4347	4347	4182	3959	3950	3950

Nota: Error Estándar en paréntesis. Coeficientes estandarizados.

Error Estándar en paréntesis. Coeficientes no estandarizados.

***= p<0.001; **= p<0.01; *=p<0.05

En la Tabla 2 que muestra los modelos estimados para Chile se presenta, en primer lugar, el modelo nulo donde se considera un modelo sin predictores que permite descomponer la varianza en los dos niveles de análisis (estudiantes al interior de las escuelas y entre escuelas). Esto permite evaluar la proporción de varianza asociada a la variabilidad entre escuelas. Como se observa, un 31% de la varianza del logro en ciencias puede ser atribuida a las diferencias entre escuelas (ICC= 0.31) y un 69% de la variabilidad puede ser atribuida a la variabilidad de los estudiantes al interior de las escuelas. Lo anterior indica que una porción importante de la variabilidad del logro en ciencias puede estar asociado a características del nivel escuelas.

El segundo modelo presentado incluye como predictor un indicador de nivel socioeconómico. A nivel del estudiante este fue incluido como una variable centrada al promedio de la escuela. A nivel de escuela el indicador corresponde al promedio para cada escuela, centrado al promedio del país. Estas condiciones de centrado permiten interpretar el efecto de las características socioeconómicas como independientes (ortogonales). Es decir, dado este centrado se puede interpretar el coeficiente del nivel individual como el efecto del nivel socioeconómico del estudiante, independiente del nivel socioeconómico promedio de la escuela a la que asiste. Así mismo, permite interpretar el coeficiente del nivel de la escuela como el efecto del nivel socioeconómico de la escuela, independiente de las condiciones individuales. En ese sentido, el modelo 1 indica que un estudiante que se desvía un punto del nivel socioeconómico promedio de su escuela obtiene 12.12 (p<0.001) puntos más en la prueba de ciencias. También indica que, a nivel de las escuelas, una escuela que se desvía 1 punto del promedio socioeconómico nacional obtiene 66.74 (p<0.001) puntos más de logro en ciencias. Este resultado muestra que tanto las condiciones socioeconómicas individuales, como las características socioeconómicas de la escuela resultan predictores relevantes del puntaje obtenido en la prueba.

Los modelos del 2 al 4 ingresan características de los estudiantes como: el género, la repitencia, la inasistencia a clases y asistencia a educación preescolar. Como se aprecia, todos son predictores estadísticamente significativos del logro en ciencias. Estos resultados permiten decir que: una niña obtiene 7.69 ($p<0.001$) puntos más que un niño, que un estudiante que repite obtiene 46.56 ($p<0.001$) puntos menos que uno que no repite, que un estudiante que no asiste a clases “más de una vez al mes” obtiene 7.42 ($p<0.001$) puntos menos que un niño que asiste regularmente, y que un estudiante que asistió a educación preescolar obtiene 14.24 ($p<0.001$) puntos más que uno que no asistió. Todos estos efectos se observan aún controlando por las características socioeconómicas individuales y de las escuelas.

En el modelo 5 se agregan como predictores las prácticas docentes de apoyo emocional, apoyo pedagógico, organización de la clase y de inasistencia del docente, controlando también por las variables socioeconómicas. Sólo la inasistencia del docente presenta un efecto estadísticamente significativo. Esto indica que en la medida que los profesores asisten menos a clases, los estudiantes obtienen un menor puntaje en la prueba de ciencias (-61.45; $p<0.001$). Respecto de las otras prácticas evaluadas, las hipótesis planteadas no se cumplen.

148

Finalmente, el modelo 6 incluye dos características adicionales del nivel escuela: si la escuela es particular-urbana y si la escuela es pública-urbana, en referencia a aquellas escuelas rurales-públicas. Como se aprecia, estas características no presentan efectos estadísticamente significativos. Sin embargo, todos los otros efectos reportados anteriormente se mantienen, lo que indica que los efectos descritos de las características individuales y de inasistencia del docente, se pueden afirmar independientemente del nivel socioeconómico de los estudiantes, del nivel socioeconómico de la escuela, de la dependencia del establecimiento y la zona en que está emplazada la escuela.

Tabla 3. Resultados modelo multinivel para Paraguay

	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
(intercepto)	646,39*** (1,81)	660,17*** (1,77)	655,09*** (2,68)	662,91*** (2,06)	668,35*** (2,33)	666,89*** (2,28)	676,41*** (2,90)
Isecf. Est- escuela.		0.106*** (0,005) 21,12*** (1,05)	0.106*** (0,005) 21,10*** (1,05)	0.095*** (0,005) 19,04*** (1,12)	0.088*** (0,010) 17,53*** (2,08)	0.076*** (0,009) 15,17*** (1,88)	0.078*** (0,009) 15,55*** (1,95)
Niña (vs niño)			0.043*** (0,008) 10,28*** (1,96)	0.038*** (0,006) 8,99*** (1,52)	0.036*** (0,008) 8,50*** (1,91)	0.030*** (0,008) 7,19*** (2,02)	0.027*** (0,008) 6,50*** (1,96)

	Modelo Nulo	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5	Modelo 6
Repite (vs No repite)				-0.084*** (0.008) -28,90*** (3,03)	-0.079*** (0.007) -27,31*** (2,41)	-0.069*** (0.006) -23,64*** (2,29)	-0.065*** (0.006) -22,50*** (2,18)
Inasistencia					-0.034*** (0.005) -10,67*** (1,55)	-0.028*** (0.003) -8,82*** (1,17)	-0.027*** (0.004) -8,47*** (1,25)
Educación preescolar					0.057*** (0.021) 13,81*** (5,17)	0.060*** (0.018) 14,55*** (4,41)	0.054*** (0.018) 13,29*** (4,43)
Prácticas docentes							
Apoyo emocional						0.178*** (0.023) 59,12*** (7,81)	0.180*** (0.023) 59,71*** (7,77)
Apoyo pedagógico						-0.068*** (0.019) -31,53*** (9,11)	-0.068*** (0.019) -31,52*** (9,10)
Inasistencia profesor						-0.039*** (0.007) -18,37*** (3,62)	-0.039*** (0.007) -18,46*** (3,68)
Organización clases						-0.047*** (0.008) -22,89*** (4,23)	-0.041*** (0.008) -20,34*** (4,04)
Variables de la escuela							
Isecf. Esc-país.		0.421*** (0.012) 65,36*** (1,89)	0.421*** (0.012) 65,29*** (1,88)	0.394*** (0.012) 61,11*** (1,93)	0.379 *** (0.019) 58,90*** (3,01)	0.374 *** (0.017) 58,11*** (2,64)	0.397 *** (0.019) 61,70*** (3,00)
Particular urbano							-0.040 (0.030) -7,23 (5,32)
Público urbano							-0.102*** (0.015) -18,12*** (2,66)
Efectos aleatorios							
s2	7035,68	7465,428	7434,60	7212,64	6979,51	6783,74	6794,81
too	6638,92	4562,40	10554,77	4248,78	4269,06	3844,81	1915,12
ICC	0,48	0,37	0,38	0,37	0,38	0,36	0,21
N	204	193	193	192	190	189	189
Obs	3442	2927	2927	2801	2174	2136	2136

Nota: Error Estándar en paréntesis. Coeficientes estandarizados.

Error Estándar en paréntesis. Coeficientes no estandarizados.

***= p<0.001; **= p<0.01; *=p<0.05

Para Paraguay se utilizó la misma lógica de agregación de predictores para cada modelo y de control por características socioeconómicas individuales y de las escuelas. Los principales resultados son los siguientes.

El modelo nulo muestra que en este país un 48% de la varianza del logro en ciencias puede ser atribuida a las diferencias entre escuelas ($ICC=0.48$) y un 52% de la variabilidad puede ser atribuida a la variabilidad de los estudiantes al interior de las escuelas. Esto indica que al igual que en el caso chileno, una parte importante de variabilidad del logro en ciencias puede estar asociado a características del nivel escuela.

El modelo 1, que incorpora los indicadores de dimensión socioeconómica, indica que un estudiante que se desvía un punto del nivel socioeconómico promedio de su escuela obtiene 21.12 ($p<0.001$) puntos más en la prueba de ciencias. También indica que una escuela que se desvía 1 punto del promedio socioeconómico nacional obtiene 65.36 ($p<0.001$) puntos más de logro en ciencias. Este resultado muestra que tanto las condiciones socioeconómicas individuales, como las condiciones socioeconómicas de la escuela son predictores relevantes del logro.

150

Los modelos del 2, 3 y 4, indican que todos los factores personales son predictores significativos del logro en ciencias. A partir de estos resultados se puede decir que: una niña obtiene 8.50 ($p<0.001$) puntos más que un niño, que un estudiante que repite obtiene 27.31 ($p<0.001$) puntos menos que uno que no repite, que un estudiante que no asiste a clases “más de una vez al mes” obtiene 10.67 ($p<0.001$) puntos menos que aquellos niños que asisten regularmente, y que un estudiante que asistió a educación preescolar obtiene 13.81 ($p<0.001$) puntos más que uno que no asistió.

En el modelo 5, se muestra que en el caso de Paraguay todas las variables relativas a las prácticas docentes presentan un efecto estadísticamente significativo. Para este país la variable más importante es el apoyo emocional, seguido de inasistencia de los profesores. Esto indica que en la medida que los profesores realizan prácticas de apoyo emocional, los estudiantes obtienen 59.12 ($p<0.001$) puntos más en la prueba de ciencias. También indica que en la medida que los profesores asisten menos a clases, los estudiantes obtienen menor puntaje en ciencias (-16.37, $p<0.001$). Las otras prácticas evaluadas, apoyo pedagógico y organización, presentan un efecto estadísticamente significativo. Sin embargo, la asociación observada contradice las hipótesis planteadas. Es decir, a partir de la literatura se esperaba un efecto positivo sobre el logro en ciencias, pero en Paraguay este efecto es negativo.

Finalmente, en el modelo 6 se aprecia que los predictores del nivel escuela no presentan efectos estadísticamente significativos sobre el logro en ciencias. Pese a esto, todos los otros efectos reportados anteriormente se mantienen, lo que indica que los efectos descritos de las características individuales y de las prácticas docentes, tienen un efecto independiente del nivel socioeconómico de los estudiantes, del nivel socioeconómico de la escuela, de la dependencia del establecimiento y la zona en que está emplazada la escuela.

5.1 Comparación entre Chile y Paraguay

Los análisis presentados para Chile y Paraguay permiten observar que existen similitudes y diferencias en los factores asociados al logro en ciencias en ambos países.

Primero, una importante diferencia es la amplitud de la proporción de la varianza asociada a la variabilidad entre escuelas. Para Chile esta representa un tercio de la varianza total ($ICC=0.31$), mientras que para Paraguay representa prácticamente la mitad de esta varianza ($ICC=0.48$). Pese a estas diferencias de proporción, en ambos países una importante porción de la variabilidad del logro en el área de ciencias puede asociarse a las diferencias entre escuelas.

Segundo, en ambos países son relevantes las características socioeconómicas a nivel de estudiantes y a nivel de escuela. Por un lado, a nivel de escuela los efectos son muy similares según lo observado en el modelo 1. Considerando los coeficientes estandarizados, en Chile el efecto es de 0.430 ($p<0.001$) y en Paraguay es de 0.421 ($p<0.001$). Por otro lado, es posible afirmar que el efecto a nivel individual es mayor en Paraguay, en donde el efecto del nivel socioeconómico llega a ser casi el doble. En Chile el efecto es de 0.060 ($p<0.001$) y en Paraguay es de 0.106 ($p<0.001$).

Tercero, los factores personales de los estudiantes presentan patrones similares de relación, exceptuando la influencia de la variable de inasistencia para Chile. En ambos países ser niña y asistir a educación preescolar mejoran el logro en ciencias, mientras que la repitencia y la inasistencia disminuyen el puntaje obtenido. Sin embargo, algunas diferencias en las intensidades son relevantes de mencionar. Observando los coeficientes estandarizados del modelo 4 (Tabla 2 y 3), se puede afirmar que: el género y la asistencia a educación preescolar tienen efectos similares en ambos países. Por un lado, en el caso del género, el efecto es de 0.039 ($p<0.001$) para Chile y de 0.036 ($p<0.001$) para Paraguay. Mientras que la asistencia a educación preescolar muestra un coeficiente de 0.058 ($p<0.001$) en Chile y un coeficiente de

0.057 ($p < 0.001$) en Paraguay. Por el contrario, la repitencia presenta una diferencia en la amplitud de su efecto en el logro. En Paraguay presenta una incidencia negativa de -0.079 ($p < 0.001$), mientras que en Chile la incidencia es casi el doble -0.137 ($p < 0.001$). Por último, la inasistencia sólo es un factor relevante en Paraguay -0.034 ($p < 0.001$), pues en Chile no es factor de importancia (-0.024; $p > 0.05$).

Finalmente, en las prácticas docentes se observan mayores diferencias en los predictores. Mientras que en Chile el apoyo emocional no es un factor relevante (0.024; $p > 0.05$), en Paraguay es la práctica pedagógica más con mayor importancia (0.180; $p < 0.001$). Por otro lado, en Chile sólo la inasistencia docente es un factor relevante (-0.130; $p < 0.001$). Mientras que en Paraguay la inasistencia (-0.039; $p < 0.001$), el apoyo pedagógico (-0.068; $p < 0.001$) y la organización (-0.041; $p < 0.001$) disminuyen el logro de los estudiantes. Este último efecto negativo del apoyo pedagógico y de la organización llaman fuertemente la atención, ya que la literatura indica que estos factores influyen en la dirección opuesta.

6. Conclusión y discusión

El objetivo principal de este estudio fue evaluar los factores que influyen en el aprendizaje de los estudiantes en el área de ciencias, en forma independiente de su nivel socioeconómico, argumentando que el aprendizaje en ciencias esta influido por distintos factores tanto individuales, como también por prácticas ejercidas por los docentes dentro de la sala de clases.

En términos generales, es posible concluir que existe una serie de variables que son grandes predictores del aprendizaje de los estudiantes en el área de ciencias. Variables individuales como el género, la asistencia a educación preescolar, la repitencia y la inasistencia de los estudiantes son importantes. También algunas prácticas docentes influyen significativamente en los resultados que obtienen los estudiantes en mediciones en ciencias. Por tanto, conocer estos factores y su efecto sobre el aprendizaje parece una decisión correcta.

Respecto de las variables individuales del estudiante evaluadas en este estudio, los resultados muestran que en general las hipótesis planteadas se cumplieron, pues una parte importante de los factores evaluados, resultaron ser predictores del logro en ciencias. De esta manera, factores individuales como la asistencia a educación preescolar son importantes para tener en cuenta por su influencia positiva en el aprendizaje. Por el contrario, aspectos como la repitencia y la inasistencia regular de los estudiantes a la escuela,

influyen de manera negativa en el aprendizaje. Teniendo en cuenta esta información, se hace importante el énfasis en políticas educativas que por un lado fomenten la asistencia a educación preescolar, sobre todo en países como Paraguay en donde el porcentaje de niños que participan de este nivel es aún muy bajo. También, es importante que se promuevan políticas que permitan sustituir el mecanismo de repitencia, por un mecanismo realmente efectivo; y que promuevan la asistencia de aquellos estudiantes que faltan en forma regular a la escuela.

Por otra parte, en relación con las prácticas docentes evaluadas, es posible indicar que la evidencia encontrada solo se alinea a la literatura en relación con algunos factores. Distinta evidencia plantea que las prácticas docentes aquí medidas afectan significativamente el aprendizaje de los estudiantes (Fauth *et al.*, 2014; Havik y Westergård, 2020; Hu *et al.*, 2020; Reyes *et al.*, 2012). Sin embargo, los resultados encontrados en este estudio, muestran que en Chile solo la inasistencia por parte del docente es un factor que influye en el aprendizaje, dejando de lado las otras prácticas incluidas. Una influencia distinta se observa en el caso de Paraguay. En este país se releva la importancia del apoyo emocional, lo que está alineado con la literatura, y al igual que en Chile la inasistencia del profesor es un factor importante. Por el contrario, el efecto del apoyo pedagógico y la organización son negativos, aspecto que va en línea contraria de lo esperado.

Llaman la atención diferentes aspectos de los resultados encontrados. Por un lado, tal como se evidencia en múltiples estudios, el aprendizaje presenta un sesgo socioeconómico en ambos países. Sin embargo, en Paraguay este efecto es mucho más relevante. Esto va en la línea de la evidencia denominada *double jeopardy* que muestra la doble ventaja a la que acceden estudiantes con mayores recursos que van a colegios con mayores recursos (Caro y Lenkeit, 2012). Por otro lado, lo que ocurre con los resultados en relación con el género de los estudiantes es un antecedente interesante de mirar. Estos resultados no se alinean a la literatura, y contrario a lo que ocurre frecuentemente, las niñas en ambos países obtienen mejores resultados en el área científica.

Esta investigación avanza en el conocimiento de los factores que influyen en el aprendizaje de los estudiantes en ciencias, sin embargo, es importante mencionar que una de sus limitaciones, tiene relación con el tipo de prácticas docentes evaluadas. El estudio utilizado provee información sólo de prácticas generales y no de prácticas pedagógicas específicas de la enseñanza- aprendizaje de las ciencias como son la investigación científica, interacción dialógica o el desarrollo de actividades prácticas.

Como pudo apreciarse, los estudios de gran escala dan lineamientos claros de las principales variables procesuales que se relacionan con el aprendizaje y sirven para identificar fortalezas y áreas de trabajo para mejorar. Asimismo, estos hallazgos permiten estudiar las desigualdades de los sistemas educativos. Por este motivo, es necesario que la comunidad académica y educativa haga un mayor y mejor uso de los estudios nacionales e internacionales de gran escala. Solo de esa forma podremos indagar en las variables clave para promover el aprendizaje.

Este estudio abre la posibilidad a futuras investigaciones, en primer lugar, se podrían estudiar prácticas docentes específicas para la enseñanza científica, incluyendo una mayor cobertura de prácticas propias de esta área de aprendizaje, por ejemplo: la investigación científica y el desarrollo de actividades prácticas. Por otra parte, con los nuevos datos de ERCE para los países de Latinoamérica, que están próximos a ser publicados, se podrían generar nuevamente modelos multinivel, para estudiar la evolución de los factores personales y de prácticas docentes efectivas aquí estudiadas.

7. Agradecimientos

Soledad Miranda agradece el financiamiento de la investigación realizado por la Agencia Nacional de Investigación y Desarrollo (ANID) a través del programa CONICYT-PFCHA/MagísterNacional/2018–50180099. Ernesto Treviño agradece el apoyo de la ANID Proyecto PIA-CIE160007.

Referencias

- Agencia de Calidad de la Educación. (2017). *Informe de Resultados PISA 2015 Competencia científica, lectora y matemática en estudiantes de quince años en Chile*. Santiago de Chile.
- Aikman, S., y Rao, N. (2012). Gender equality and girls' education: Investigating frameworks, disjunctures and meanings of quality education. *Theory and Research in Education*, 211-228.
- Aarepattamannil, S., Freeman, J. G., y Klinger, D. A. (2011). Influence of motivation, self-beliefs, and instructional practices on science achievement of adolescents in Canada. *Social Psychology of Education*, 233-259.
- Bates, D., Maechler, M., Bolker, B., Walker, S., Christensen, R. H. B., Singmann, H., y Green, P. (2018). Package 'lme4'. *Version*, 1, 17.
- Barber, M., y Mourshed, M. (2008). *Como hicieron los sistemas educativos con mejor desempeño del mundo para alcanzar sus objetivos*. Santiago de Chile: PREAL.

- Blanchard, M. R., Southerland, S. A., y Granger, E. M. (2009). No silver bullet for inquiry: Making sense of teacher change following an inquiry-based research experience for teachers. *Science Education*, (93), 322–360.
- Bruns, B., Luque, J., De Gregorio, S., Evans, D., Fernández, M., Moreno, M., Rodríguez, J., Toral, G., y Yarrow, N. (2014). *Profesores excelentes: cómo mejorar el aprendizaje en América Latina y el Caribe*. Banco Mundial.
- Bybee, R. W. (2010). *The teaching of science: 21st century perspectives*. NSTA press.
- Caro, D. H., y Lenkeit, J. (2012). An analytical approach to study educational inequalities: 10 hypothesis tests in PIRLS 2006. *International Journal of Research & Method in Education*, 35(1), 3-30.
- Cook, Philip, Kenneth Dodge, Elizabeth Gifford y Amy Schulting (2017). A New Program to Prevent Primary School Absenteeism: Results of a pilot study in five schools, *Children and Youth Services Review*, 82, 262-270.
- Creemers, B., y Kyriakides, L. (2008). *The dynamics of educational effectiveness. London: A contribution to policy, practice and theory in contemporary schools*. Routledge.
- Cueto, S., Torero, M., León, J., y Deustua, J. (2008). *Asistencia docente y rendimiento escolar: el caso del Programa META*. (Documento de Trabajo, 53). Lima: GRADE Group for the Analysis of Development.
- Downer, J. T., Stuhlman, M., Schweig, J., Martínez, J. F., y Ruzek, E. (2015). Measuring Effective Teacher-Student Interactions from a Student Perspective: A Multi-Level Analysis. *Journal of Early Adolescence*, 723-724.
- Emmer, E. T., y Stough, L. M. (2001). Classroom Management: A Critical Part of Educational Psychology, with Implications for Teacher Education. *Educational psychologist*, 36(2), 103-112.
- Fauth, B., Decristan, J., Rieser, S., Klieme, E., y Büttner, G. (2014). Student ratings of teaching quality in primary school: Dimensions and prediction of student outcomes. *Learning and Instruction*, 29, 1-9.
- González., M. (2006). Absentismo y abandono escolar: una situación singular de la exclusión educativa. *REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 4(1), 1-15.
- González González, M. T. (2016). Absentismo Escolar: Posibles Respuestas desde el Centro Educativo. *REICE. Revista Iberoamericana Sobre Calidad, Eficacia Y Cambio En Educación*, 12 (2), 5-27.
- Hamre, B. K., y Pianta, R. C. (2007). Learning opportunities in preschool and early elementary classrooms. *Paul H Brookes Publishing*, 49-83.
- Havik, T., y Elsa Westergård, E. (2020). Do Teachers Matter? Students' Perceptions of Classroom Interactions and Student Engagement. *Scandinavian Journal of Educational Research*, 64(4), 488-507.
- Heyneman, S. P. (2004). International education quality. *Economics of Education Review*, 23(4), 441-452.
- Hofstein, A., Kipnis, M., y Kind, P. (2008). Learning in and from science laboratories: Enhancing students' metacognition and argumentation skills. *Science education issues and developments*, 59-94.
- Hunt, B. C. (2009). *Efectividad del desempeño docente. Una reseña de la literatura internacional y su relevancia para mejorar la educación en América Latina*. PREAL.

- Hu, B. Y., Fan, X., Wu, Y., LoCasale-Crouch, J., y Song, Z. (2020). Teacher-child interaction quality and Chinese children's academic and cognitive development: New perspectives from piecewise growth modeling. *Early Childhood Research Quarterly*, 51, 242-255.
- Klieme, E., Pauli, C., y Reusser, K. (2009). The Pythagoras study: Investigating effects of teaching and learning in Swiss and German mathematics classrooms. *The Power of Video Studies in Investigating Teaching and Learning in the Classroom*, 137-160.
- McNeill K. L., Pimentel D. S. (2009) Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education*, 94(2), 203-229.
- Méndez, I. y Cerezo, F. (2018). La repetición escolar en educación secundaria y factores de riesgo asociados [Repetition rate in secondary education and associated risk factors]. *Educación XXI*, 21(1), 41-62.
- OCDE. (2016). *PISA 2015 Resultados clave*. Recuperado de <https://bit.ly/3jRP44x>
- OECD. (2017). *PISA 2015 assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving*. OECD Publishing.
- OREALC-UNESCO. (2016). Aportes para la enseñanza de las ciencias naturales. Santiago, Chile.
- Pianta, R. C., y Hamre, B. K. (2009). Conceptualization, Measurement, and Improvement of Classroom Processes: Standardized Observation Can Leverage Capacity. *Educational Researcher*, 38(2), 109-119.
- Pianta, R. C., Hamre, B. K., y Allen, J. P. (2012). Teacher-Student Relationships and Engagement: Conceptualizing, Measuring, and Improving the Capacity of Classroom Interactions. *Handbook of Research on Student Engagement*, 365-386.
- Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., y Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, 104(3), 700-712.
- Rivas, A. (2015). *América Latina después de PISA: Lecciones aprendidas de la Educación en siete países 2000-2015*. Fundación CIPPEC.
- Rubin, D. B., y Schenker, N. (1986). Multiple Imputation for Interval Estimation from Simple Random Samples with Ignorable Nonresponse. *Journal of the American Statistical Association*, 81(394), 366-374.
- Scheerens, J. (2013). *Effectiveness of Time Investments in Education: Insights from a review and meta-analysis*. Springer Science y Business Media.
- Treviño, E., Aguirre, E., y Varela, C. (2018). *Un buen comienzo para los niños de Chile*. Ediciones Universidad Diego Portales.
- UNESCO-OREALC (2016). *Reporte técnico. Tercer Estudio Regional Comparativo y Explicativo*. TERCE. Santiago Chile.
- UNESCO (2015). *Informe de Resultados TERCE: Factores asociados*. UNESCO Publishing.
- UNICEF (2007). *Todos pueden aprender. Propuestas para superar el fracaso escolar*. Buenos Aires, Argentina.
- Vegas, E., y Santibañez, L. (2010). *The Promise of Early Childhood Development in Latin America and the Caribbean*. World Bank Publications.

Determinante del desempeño en lectura en Uruguay: un análisis multinivel a partir de TERCE

Adrián Silveira Aberastury 

Universidad de la República (UdelaR), Uruguay

Resumen. El objetivo de este trabajo es identificar los factores que influyen en el desempeño de los estudiantes al final de la escuela primaria en Uruguay. Los datos del Tercer Estudio Regional Comparativo y Explicativo (TERCE) de UNESCO se utilizaron para identificar los factores asociados con el rendimiento escolar en lectura. Los efectos sobre el desempeño de los estudiantes se analizaron a través de modelos multinivel que consideran las características de las escuelas, los maestros, los estudiantes y sus familias en el dominio de la lectura. El análisis confirma la relevancia de los factores individuales y composicionales basados en el origen socioeconómico y cultural y las expectativas educativas de los padres. Los factores institucionales representaron una parte menor de la varianza total explicada, lo cual indica la limitada capacidad de las escuelas uruguayas de influir sobre los aprendizajes de sus estudiantes.

Palabras clave: modelos multinivel; factores asociados a los desempeños; evaluaciones internacionales a gran escala; TERCE.

Determinante do desempenho na leitura no Uruguai: uma análise multinível baseada no TERCE

Resumo. O objetivo deste trabalho é identificar os fatores que influenciam o desempenho dos estudantes ao final do ensino fundamental no Uruguai. Os dados do Terceiro Estudo Regional Comparativo e Explicativo da UNESCO (TERCE) foram usados para identificar os fatores associados ao desempenho escolar em leitura. Os efeitos no desempenho dos alunos foram analisados através de modelos multiníveis que consideram as características das escolas, professores, alunos e suas famílias no domínio da leitura. A análise confirma a relevância de fatores individuais e composicionais com base na origem socioeconômica e cultural e nas expectativas educacionais dos pais. Os fatores institucionais representaram uma parte menor da variância total explicada, o que indica a capacidade limitada das escolas uruguaias de influenciar a aprendizagem de seus estudantes.

Palavras-chave: modelos multiníveis; fatores associados ao desempenho; avaliações internacionais em grande escala. TERCE.

Performance indicator in reading in Uruguay; a multi-level approach from TERCE

Abstract. The aim of this study is to identify the factors that influence students' achievement at the end of elementary school in Uruguay. Data from UNESCO's Third Regional Comparative and Explanatory Study (TERCE) were used to identify the factors associated with school performance in reading. The effects on student performance were analyzed through multilevel models that consider the characteristics of schools, teachers, students and their families in the reading performance. The analysis confirms the importance of individual and compositional effects based on socioeconomic and cultural background and educational expectations of parents. Institutional factors represented a smaller part of the total explained variance, which indicates the limited capacity of Uruguayan schools to influence in academic achievement.

Keywords: multilevel models; factors associated with academic achievement; international large-scale assessments; TERCE.

157

1. Introducción

Entre los numerosos desafíos que enfrenta cualquier sistema educativo, se impone el doble desafío de mejorar los aprendizajes de los estudiantes y abordar el problema de la desigualdad en sus logros académicos. En efecto, la experiencia internacional ha mostrado (Willms, 2006; 2010)

que los sistemas educativos de calidad son aquellos que logran mejorar sus desempeños y, simultáneamente, reducir las desigualdades. Ambos asuntos son una cuestión recurrente de la agenda educativa uruguaya.

Pese a los notorios avances en el acceso y cobertura del sistema educativo nacional (ANEP, 2015b), la cuestión de la calidad de los aprendizajes es asunto de álgidos debates a la luz de los resultados de las evaluaciones internacionales a gran escala. En efecto, aunque el país cuenta con destacados desempeños en el contexto latinoamericano (UNESCO, 2016b) su performance se encuentra a una distancia considerable de países desarrollados (ANEP, 2017; INEED, 2015; OECD, 2016).

Asimismo, el desempeño uruguayo no es ajeno a un rasgo distintivo de los sistemas educativos latinoamericanos: altos niveles de desigualdad de aprendizaje con una marcada incidencia del origen socioeconómico de los estudiantes (ANEP, 2015a; Cardozo, 2016; Fernández, 2004; INEED, 2014, 2017; UNESCO, 2016a).

Las preocupaciones por la desigualdad educativa suelen conceptualizarse desde dos nociones distintas pero complementarias. En primer lugar, las concepciones basadas en las nociones de justicia asociadas a la equidistribución de aquellos bienes socialmente valiosos (Atkinson, 2015; Cowell, 1998; Roemer y Trannoy, 2015). Bajo este enfoque la desigualdad distribución de los desempeños educativos se observa en la dispersión de los resultados. En segundo término, otros enfoques sobre la desigualdad se concentran en analizar el grado en que la dispersión de los desempeños se asocia a otros factores. Esta perspectiva suele referirse como desigualdad por estratificación o desigualdad de oportunidades educativas (Cortés y Rubalcava, 1984; Fernández y Cardozo, 2009).

Por otra parte, en la medida que crece el interés en los procesos educativos y su impacto en el futuro de las vidas de las personas, las evaluaciones internacionales de aprendizaje a gran escala (ILSA, por su sigla en inglés) se han expandido considerablemente (Hanushek, 2011, 2013; Kirsch y Braun; 2020). Estas evaluaciones proporcionan información para una variedad de actores, decisores de política, investigadores y la comunidad educativa en general (Blossfeld *et al.*, 2009; Ritzen, 2013; Von Davier *et al.*, 2013). En efecto, además de la medición de los desempeños, las ILSAs recogen información relevante para el estudio de la desigualdad, a partir de los cuestionarios aplicados, así como del propio diseño muestral.

Los estudios sobre la desigualdad en los desempeños se basan en el desarrollo de análisis multivariados que buscan aislar los efectos específicos de cada una de las variables de interés sobre los resultados. En

educación, es frecuente que estos modelos multivariados distingan, además, entre efectos asociados a características individuales de los estudiantes y efectos agregados, de tipo composicional, institucional o contextual, lo que representa más adecuadamente el hecho de que los estudiantes que asisten a una misma escuela están expuestos a un conjunto de situaciones comunes que contribuyen a moldear sus oportunidades y riesgos educativos. Técnicamente, esto implica desarrollar modelos estadísticos de tipo multinivel (Raudenbush y Bryk, 2002).

Este trabajo consiste en focalizar un conjunto de dimensiones individuales e institucionales, con el propósito de analizar la estructura básica de la desigualdad de los logros educativo. En concreto, el objetivo de este artículo es identificar los principales factores asociados al desempeño en el dominio de la lectura de los estudiantes al final de la escuela primaria a través de modelos multinivel, que considera los efectos de las características de las escuelas y su entorno, los estudiantes y sus familias.

2. Fundamentación teórica

Hasta hace poco tiempo, al calor de los debates sobre el acceso y progresión dentro de los sistemas educativos, se consideraba el resultado del proceso educativo a través indicadores como la expansión de la matrícula escolar. A su vez, la noción de logro académico se observaba en la aprobación de los cursos y/o las calificaciones escolares y, su contracara, como la repetición y desvinculación. Con el tiempo, al influjo del interés por la evaluación internacional a gran escala, los estudios se han reorientado hacia el análisis de los efectos de los factores familiares y escolares en el rendimiento de los estudiantes, lo que ha habilitado una discusión más fructífera sobre la calidad de la educación (Albernaz *et al.*, 2002; Carnoy, 2015; Soares, 2004; Wagemaker, 2014).

El estudio de los factores asociados consiste en identificar las condiciones que influyen en el aprendizaje y dimensionar sus efectos. Su abordaje es complejo, en parte, por la variedad de conceptos y variables a considerarse, tanto a nivel individual como institucional. Sin embargo, en el análisis de los determinantes es usual que la literatura distinga, *grosso modo*, dos grupos principales: los relacionados con los estudiantes y sus familias, y por otra parte, los vinculados a la estructura y condiciones escolar (Andrade y Soares, 2008; Blanco, 2011; Fernández, 2004; Willms, 2006).

2.1 Factores individuales y familiares

El desempeño de los estudiantes en cada etapa de la vida escolar es el resultado de un conjunto de factores que a menudo escapan a los límites de las escuelas. Entre la multiplicidad de factores intervinientes (personales, familiares, sociales e institucionales), existe relativo consenso en la identificación de los principales determinantes del rendimiento escolar los que se refieren a las características socioeconómicas de las familias de los estudiantes. Hace más de 50 años, los informes “Coleman” (Coleman *et al.*, 1966) y “Plowden” replantearon la cuestión de la influencia de las escuelas en el rendimiento y la relevancia de los aspectos familiares, jerarquizando los factores psicológicos (actitudes y valores de los padres en relación con los estudios y el futuro escolar) y los factores materiales (ingresos y condiciones materiales). Estos resultados, debido a su interés y controversia, ejercieron una gran influencia en la investigación posterior, demostrando sistemáticamente que el nivel educativo de los padres de los estudiantes, las condiciones culturales del hogar y la situación económica de la familia son los aspectos que más afectan el aprendizaje de los estudiantes (Soares, 2004).

Por otra parte, la revisión de la literatura indica sistemáticamente que eventos como el rezago (por repetición) y la desvinculación en las trayectorias escolares, las prácticas educativas en el hogar y los atributos socioeconómicos, demográficos y culturales de las familias están asociados con el rendimiento académico (Albernaz *et al.*, 2002; Cardozo, 2016; Castro, 2009; Soares, 2003, 2004; Willms, 2006).

Mientras en la trayectoria educativa, la repetición y el ausentismo mantienen una sólida asociación negativa con el rendimiento escolar; por el contrario, la asistencia a la educación preescolar es un factor que tiene una relación positiva con el aprendizaje. Además, los antecedentes muestran que las expectativas educativas, el monitoreo y retroalimentación de los docentes del desempeño académico de los estudiantes, los hábitos de lectura y el estudio en el hogar, son factores positivamente asociados con el aprendizaje (ANEP, 2015a; Blanco, 2008; Cervini, 2002; Fernández, 2004; INEED, 2017; UNESCO, 2016a).

2.2 Los efectos escolares

Si las diferencias sociales entre los estudiantes que reciben las escuelas son inmutables en función de las capacidades del sistema educativo, existe una serie de variables que la institución escolar está en condiciones de manipular. Ahora bien, la escuela, y específicamente el aula, se considera el núcleo fundamental de los procesos de interacción, a través de los cuales se produce el aprendizaje escolar.

La magnitud de los “efectos escolares” depende, en teoría, de dos tipos diferentes de aspectos. Por un lado, aquellos factores conocidos como “composicionales” y que refieren directamente al accionar escolar. Son la expresión agregada de diversos atributos individuales, como la proporción de varones o el nivel socioeconómico promedio de los estudiantes que concurren a la misma escuela (Raudenbush y Willms, 1995). Se asume que estos factores operan sobre el desempeño individual, con independencia de la condición individual de cada estudiante. Diversos estudios muestran que, por ejemplo, estudiantes que asisten a escuelas donde sus compañeros provienen de hogares con alto capital cultural, logran, en promedio, mejores desempeños independientemente de su contexto familiar de origen. No existe consenso sobre los mecanismos que operan detrás de estos efectos, pero los investigadores lo asocian a la incidencia del grupo de pares y la interacción con otros recursos institucionales, como la organización y gestión escolar (De Melo, 2011; Harker y Tymms, 2004; Hattie, 2002).

En segundo lugar, se ubican factores de nivel escolar que han recibido diversas denominaciones, que pueden englobarse como efectos “institucionales y contextuales” (Willms, 2010). Estos se basan en atributos de la institución escolar y refiere a los contextos específicos en los que se desarrolla la vida escolar. Se identifican aquí: la dotación de recursos para la enseñanza, los programas académicos, el tiempo escolar, las características del plantel docente, las prácticas de gestión institucional, entre otras. En general, buena parte de estas dimensiones suelen señalarse por parte de la literatura sobre eficacia escolar (Torrecilla, 2006). Aquí la hipótesis central es que estos aspectos pueden, en principio, lograr que la escuela “hagan la diferencia”, más allá de las características de los estudiantes y los factores composicionales. Los efectos institucionales, así entendidos, revisten particular importancia para la planificación educativa. Para el caso uruguayo, algunos de estos, como el formato institucional, pueden y, son de hecho, objeto de la política educativa, aun cuando dependen de definiciones centrales y no de las propias escuelas.

3. Datos y métodos

3.1 Fuentes de información

Los datos provienen de la muestra representativa de estudiantes de sexto grado de educación primaria de Uruguay, relevados para la evaluación TERCE de UNESCO en 2013. Se aplicó un diseño muestral aleatorio estratificado bi-etápico (UNESCO, 2016c), donde la unidad de muestreo es la escuela sorteada con probabilidad proporcional al tamaño de su matrícula y, en una segunda instancia, en cada escuela se elige aleatoriamente un grupo/

aula que se releva íntegramente (n=2642 estudiantes; 167 escuelas). Las variables de estratificación explícita son: dependencia administrativa (público/privado); área geográfica (rural/urbana) y convivencia de grados (Escuelas con solo tercero, con solo sexto y ambos grados).

Además, como es habitual en estos estudios, cada dominio cognitivo se estima a través de un conjunto de valores plausibles (von Davier *et al.*, 2009). En todos nuestros análisis se procedió a especificar este diseño muestral y considerar el set completo de valores plausibles.

Pese a que TERCE recopila datos sobre el aprendizaje en tercero y sexto grado, se trabajó únicamente con sexto debido a dos razones sustantivas. En primer lugar, se asume que el proceso educativo es, por definición, acumulativo. Es decir, lo demostrado por un estudiante en una prueba estandarizada es el resultado (entre otros factores) de lo que ha aprendido a lo largo de su vida escolar. Se considera, por tanto, apropiado considerar el nivel educativo superior. Segundo, sexto grado representa la finalización de un ciclo en la escolarización, ya que, en Uruguay, este es el último grado del nivel primario. El ciclo siguiente, correspondiente a la educación secundaria, implica importantes cambios institucionales, donde, por ejemplo, los contenidos curriculares se diversifican en una multiplicidad de asignaturas y docentes.

3.2 Los modelos jerárquicos lineales

Los modelos jerárquicos lineales o modelos multinivel (Raudenbush y Bryk, 1986, 2002; Hox, 2010) es —en su definición más elemental— un modelo estadístico que relaciona variables independientes (de diferentes niveles) para explicar la variación de una variable dependiente. Son, en esencia, una extensión de los modelos de regresión lineal múltiple con algunas peculiaridades (Cervini, 2002).

Los modelos multinivel son adecuados para analizar datos que presentan una estructura jerárquica o anidada. En la mayoría de las evaluaciones de aprendizaje a gran escala, esta jerarquía se observa fácilmente, ya que los estudiantes se agrupan en clases o grupos y estos, a su vez, se organizan dentro de las escuelas. Las ventajas de considerar una estructura jerárquica, presente en los datos de las evaluaciones educativas, radica en que permite considerar separadamente los efectos del individuo y el grupo al que pertenece, lo que permite identificar y distinguir el efecto de las escuelas en el rendimiento de los estudiantes.

Para Soares (2004), el uso de un modelo que incorpora la estructura jerárquica presente en los datos presenta las siguientes ventajas: (1) obtener mejores estimaciones para los parámetros relacionados con unidades específicas, porque a través del modelo jerárquico es posible obtener una ecuación para cada escuela, haciendo así uso eficiente de toda la información presente en la muestra; (2) posibilidad de formular y probar hipótesis sobre los efectos entre niveles; (3) partición de la varianza en componentes, por ejemplo, para un modelo con dos niveles, existe la varianza del nivel 1 (estudiantes), la varianza del nivel 2 (aulas o escuelas).

Otra ventaja de utilizar el análisis multinivel, en este caso, es que permite estudiar, al mismo tiempo, los efectos de las variables individuales y grupales en los resultados individuales y, además, incorpora la consideración de la dependencia de las observaciones dentro de cada grupo. Así, el uso de modelos multinivel permite obtener mejores estimadores de los coeficientes de regresión y su variación, en comparación con los modelos tradicionales. Además, las pruebas estadísticas estándar dependen en gran medida del supuesto de independencia de las observaciones y, si esto se viola (como en las estructuras multinivel), las estimaciones de error estándar de las pruebas estadísticas convencionales son mucho más bajas y sus resultados son espurios (Hox, 2010).

4. Resultados

La desigualdad en los desempeños educativos puede concebirse como el resultado agregado de dos componentes complementarios: la desigualdad que, en promedio, se constata *entre escuelas* y la desigualdad observada *entre los estudiantes* que concurren al mismo establecimiento educativo. Para dar cuenta de estos procesos, el análisis se organiza en dos etapas. En primer lugar, se estima la proporción de la varianza total de la variable de interés que se puede atribuir a cada uno de los niveles en el análisis (Raudenbush y Bryk, 1986). En segundo lugar, se ajustaron sucesivos modelos de efectos fijos con las variables del nivel individual y escolar que la literatura considera relevantes.

4.1 Descomposición de la varianza

El Modelo Nulo (MO), sin variables explicativas, evalúa la influencia de cada nivel, es decir, cuánto influye la escuela (nivel 2) en el rendimiento del estudiante y la influencia del propio estudiante (nivel 1) en su escuela. Por lo tanto, la descomposición de la varianza nos permite saber qué proporción de la variación corresponde a cada nivel de análisis.

La proporción de variación explicada por la escuela (Coeficiente de Correlación Intraclase, CCI) puede interpretarse como la porción de responsabilidad atribuida a las características internas de las escuelas en el desempeño de sus estudiantes. O, en otras palabras, ¿cuál es la magnitud de la variación promedio entre escuelas (*between schools*) que se refiere a propiedades de la institución escolar?

El CCI es una medida de heterogeneidad en las escuelas, conceptualmente relevante ya que permite distinguir empíricamente el impacto de varios niveles teóricos en la determinación de un resultado individual. Operativamente MO se utiliza para estimar la varianza explicada *versus* la varianza inexplorada, en comparación con los modelos condicionales estimados *a posteriori*.

164

La Tabla 1 presenta que las mayores diferencias en el rendimiento en Lectura se deben a las características individuales de los estudiantes. Es decir, el resultado sugiere que la mayor parte de la varianza total corresponde a las diferencias entre estudiantes. La varianza entre escuelas representa una proporción importante ($CCI = 0.2615$), aunque menor. Los resultados se ajustan a los antecedentes de las evaluaciones nacionales e internacionales para el nivel de educación básica (ANEP, 2009; 2015a; Blanco, 2011; Fernández, 2004; UNESCO, 2016a).

Grosso modo, una cuarta parte de la variabilidad de las competencias de los estudiantes en Lectura se debe principalmente a las diferencias entre las escuelas, lo que la literatura denomina el *efecto escolar*, es decir, que la variabilidad puede explicarse por atributos del nivel 2. Su complemento, aproximadamente las tres cuartas partes restantes de la variación, se debe a diferencias entre los estudiantes dentro de las escuelas, que, en principio, pueden estar relacionadas con factores de naturaleza individual.

Tabla 1. Modelo incondicional (Modelo Nulo) y partición de la varianza.

	Prueba de lectura
Varianza de nivel 2 – Escuela	3163.6
Varianza de nivel 1. Estudiantes	8936.4
Total de varianza	12100.0
CCI: Coeficiente de Correlación Intraclass	0.2615

Fuente: Elaboración propia en base a datos TERCE - UNESCO 2013.

Sin embargo, según Blanco (2008), existe una tendencia a interpretar el CCI como un indicador de la influencia directa de las escuelas en el aprendizaje de los estudiantes. Esto constituye, por dos razones, una atribución imprecisa. Primero, porque no considera que, al menos, parte de estas diferencias entre los promedios escolares se deba a las características de selectividad y contexto, sobre las cuales las escuelas no ejercen control. Segundo, porque no considera que las escuelas puedan influir en las diferencias entre los estudiantes, incrementándolos o disminuyéndolos de acuerdo con los procesos escolares.

Como señala Soares (2004), es necesario controlar estadísticamente el efecto de las variables relacionadas con la selectividad y composición de la escuela; de lo contrario, atribuiríamos a las escuelas un efecto que no es propio, sino del capital económico y cultural de las familias de los estudiantes.

Para evaluar la validez de esta hipótesis, pueden introducirse en el modelo diferentes factores para observar si, efectivamente, parte de esta variación pueda explicarse por características composicionales de las escuelas (Cervini, 1999; 2002; Fernández, 2004; Willms, 2006).

El modelo 1 (M1) ajusta la composición socioeconómica y cultural de la escuela que resulta del promedio de las características socioculturales de los estudiantes. El centrado de los predictores es una técnica recomendada por varios autores (Enders y Tofighi, 2007; Raudenbush y Bryk, 2002), que tiene por utilidad una mayor capacidad interpretativa de los coeficientes y ayuda a reducir la colinealidad entre niveles. Específicamente, para ajustar modelos composicionales, donde el nivel 2 es un agregado del nivel 1, implica incluir los puntajes individuales y los promedios de los grupos de referencia, realizando el centrado del nivel 1 en el cluster de referencia, en este caso, cada escuela. Al “restar” el efecto del nivel 1, se estima con mayor precisión la proporción de la variación atribuible al nivel 2. De manera que, por ejemplo, el efecto composicional del nivel socioeconómico corresponde a la diferencia entre el promedio grupal estimado para la variable NSE y el coeficiente estimado para la variable individual, centrada en relación a la escuela.

Consecuentemente, siguiendo a Willms (2010), en M2 amplía el set de variables composicionales. El cuadro 1 sintetiza los 3 modelos propuestos para el estudio de descomposición de la varianza.

Cuadro 1. Modelos con variables composicionales.

Modelo	Descripción	Variables
M0	Modelo Nulo	Vacío
M1	Modelo de composición socioeconómica de la escuela	Promedio del Índice de NSE de la escuela/grupo (a) Índice de NSE de la familia de los estudiantes
M2	Modelo composicional ampliado	Promedio del Índice de NSE de la escuela (a) Índice de NSE de la familia de los estudiantes Proporción de estudiantes con integración familiar completa (a) Integración familiar completa Proporción de estudiantes con rezago en la escuela (a) Rezago individual del estudiante Proporción de estudiantes con educación pre-escolar (a) Estudiantes con educación pre-escolar Proporción de padres con altas expectativas educativo de sus hijos (a) Padres con altas expectativas educativas de sus hijos Proporción de estudiantes que trabajan (a) Estudiantes que trabajan

Fuente: Elaboración propia en base a datos TERCE-UNESCO 2013.

Nota: (a) Indica variables agregadas para nivel 2 - escuela/grupo.

La Tabla 2 muestra que en M1, el coeficiente CCI se reduce drásticamente. Los datos muestran que el solo efecto composicional del entorno sociocultural es suficiente para explicar “estadísticamente” el 58.9% de la variación entre las escuelas. En otras palabras, solo el 10.75% de la variación del desempeño en lectura puede atribuirse, específicamente, a las escuelas.

M2 reduce aún más la variabilidad atribuible a factores escolares, confirmando la hipótesis de que las diferencias entre escuelas se deben principalmente a factores composicionales. Está claro, entonces, que no todas las diferencias entre los resultados escolares son atribuibles a los procesos y prácticas de las propias escuelas.

Los factores de composición escolar representan una parte muy significativa de la variación propia de nivel 2, dejando relativamente poco espacio para la acción escolar. La evidencia muestra que alrededor del 8% de las diferencias en el rendimiento entre las escuelas pueden considerarse estadísticamente independientes de la selectividad y la composición social

de los estudiantes. Es decir, que existe una proporción relativamente baja de la variación en los desempeños que obedece a un conjunto de factores que son estrictamente escolares, es decir, de la organización, la gestión y las prácticas educativas que operan dentro de las escuelas.

Tabla 2. Descomposición de la varianza, controlado por modelos de composición socioeconómica y cultural de la escuela y factores contextuales.

	MO		M1		M2
	Var.	Var	Reducción (vs M0)	Var	Reducción (vs M0)
Varianza de nivel 2 – Escuela	3163.6	1076.3	66.0%	796.0	74.8%
Varianza de nivel 1. Estudiantes	8936.4	8939.5	0.0%	8941.9	-0.1%
Total	12100.0	10015.8	17.2%	9737.9	19.5%
CCI (p)	0.2615	0.1075	58.9%	0.0817	68.7%

Fuente: Elaboración propia en base a datos TERCE - UNESCO 2013.

4.2 Ajuste de modelos multinivel: efectos fijos

La Tabla 3 presenta los resultados de los modelos multinivel ajustados para describir los efectos fijos sobre el dominio de Lectura. Se ajustaron los modelos M3 a M5, con las variables de composición que fueron significativas en los modelos de estudio de la descomposición de la varianza (M0 a M2). Por lo tanto, M3 incorpora variables del nivel 1 (estudiante) y M4 las del nivel 2 (escuela). El modelo 5 presenta las variables significativas ajustadas para el modelo final. En los modelos intermedios (M1 a M4) se adoptó la estrategia conservadora de eliminar solamente los términos no significativos al 90% debido al contexto exploratorio de la propia fase de ajuste.

Como se presentó en la sección anterior, los modelos M0 a M2 fueron utilizados para mostrar el peso de los factores composicionales en la varianza de escuelas. No obstante, en estos análisis no todas las variables fueron significativas (ver Anexo). Por lo tanto, en M3 se parte de las variables significativas para el estudio de los factores composicionales y se incorporan los factores individuales de nivel 1.

Sobre el ajuste de M3 es necesario realizar una serie de comentarios. En primer lugar, destacar el efecto de los factores composicionales, aun incorporadas las variables de nivel 1. Además, es evidente que no todo el efecto composicional se expresa a través de lo que capta el Índice de NSE. De hecho, las expectativas parentales sobre la educación de sus hijos, así como los hogares integrados por ambos padres presentan un fuerte efecto composicional. Además, se observa que el desempeño de los estudiantes cuyos padres afirman tener altas expectativas para la educación futura de

Tabla 3. Modelos multinivel de efectos fijos en lectura. Coeficientes, errores estándar y estadístico T y significación.

	Modelo 3 (M3)				Modelo 4 (M4)				Modelo 5 (M5)			
	Coef.	e.e.	T	Sig	Coef.	e.e.	T	Sig	Coef.	e.e.	T	Sig
Efectos fijos												
Intercepto	657.18	14.50	45.31		663.23	16.32	40.64		664.54	11.70	56.79	
Factores composicionales												
[1] Índice NSE (a)	46.91	7.31	6.41	(***)	41.78	9.53	4.38	(***)	49.94	7.36	6.79	(***)
[2] Prop. padres altas expectativas educ. (a)	59.56	18.56	3.21	(***)	57.55	19.99	2.88	(***)	71.35	19.22	3.71	(***)
[3] Prop. integración familiar completa (a)	43.75	18.80	2.33	(**)	25.32	19.20	1.32					
[4] Prop. Estudiantes que trabajan (a)	-21.08	12.16	-1.73	(*)	-10.75	12.86	-0.84					
Factores individuales (nivel 1)												
[5] Índice NSE	16.02	4.15	3.87	(***)	17.48	4.27	4.10	(***)	18.31	4.08	4.48	(***)
[6] Rezago (repite al menos 1 vez)	-57.17	5.66	-10.10	(***)	-56.10	5.79	-9.69	(***)	-58.23	5.65	-10.31	(***)
[7] Padres con altas expectativas educativas	31.46	5.75	5.47	(***)	30.19	6.00	5.03	(***)	29.30	5.83	5.03	(***)
[8] Uso familiar de información escolar	18.43	6.03	3.06	(***)	15.98	6.06	2.64	(***)	18.83	6.13	3.07	(***)
[9] Trabaja	-16.20	4.74	-3.42	(***)	-15.69	4.72	-3.32	(***)	-17.37	4.61	-3.77	(***)
[10] Es niña	12.03	4.86	2.48	(**)	12.34	4.68	2.64	(***)	13.02	4.72	2.76	(***)
[11] Educación pre-escolar	10.45	5.64	1.85	(*)	8.67	5.88	1.47					
[12] Ausentismo (1 vez a la semana o más)	4.04	5.72	0.71									
Factores escolares (nivel 2)												
[13] Montevideo (vs. Interior)					10.52	7.47	1.41					
[14] Colegio privado (vs. Público)					9.02	10.01	0.90					
[15] Índice de infraestructura escolar					1.59	4.52	0.35					
[16] Experiencia como Director/a (en años)					-0.01	0.25	-0.04					
[17] Índice de liderazgo directriz					-3.12	3.10	-1.01					
[18] Índice de clima escolar					5.56	3.02	1.84	(*)				
[19] Experiencia como Maestro/a (en años)					-0.22	0.26	-0.82					
[20] Formación docente reciente					2.64	6.63	0.40					

(*) P<0.10; (**) P<0.05; (***) P<0.01.

Fuente: Elaboración propia en base a datos TERCE-UNESCO 2013. Nota: (a) Indica variables agregadas para nivel 2 - escuela/grupo.

sus hijos, es un factor estadísticamente relevante. Algo similar sucede con la variable sobre el uso de la información escolar (en fila [8]), donde se observa que, en relación con la información que los padres reciben de la escuela, aquellos que apoyan a sus hijos parecen ser un factor positivamente asociado y muy relevante.

En segundo lugar, con excepción de Ausentismo, entendido, como la ausencia frecuente de clases (ubicado en la fila [12]), todas las variables incorporadas de nivel 1 fueron significativas, al menos al 90% de confianza. En particular, la condición de rezago (en fila [6]), presenta el coeficiente más alto de todos los ajustados. El rezago en al menos un año escolar, disminuye el desempeño promedio en medio desvío, en una escala con media 700 y desvío 100. El otro término negativo lo representa la condición laboral, que presenta un efecto negativo en relación a su desempeño en lectura. En tercer término, los datos muestran, en línea con lo que plantea la literatura (UNESCO, 2016a) que las niñas presentan, en promedio, una fuerte asociación positiva con el desempeño en el área de Lectura.

5. Discusión

169

Los análisis mostraron el fuerte efecto de la composición sociocultural de la escuela, capturada por el valor promedio de la escuela del índice NSE de los estudiantes. El efecto composicional del nivel socioeconómico presenta un efecto de magnitud considerablemente mayor que el Índice de NSE individual de la familia del estudiante, un resultado que ha sido documentado repetidamente en estudios anteriores para diferentes países y para diferentes niveles educativos.

Asimismo, los datos muestran la fuerte incidencia de las expectativas parentales sobre la educación de sus hijos, tanto como factor composicional como individual. Se encontró evidencia sólida sobre la importancia de generar altas expectativas en relación con el aprendizaje. La incidencia de este factor puede interpretarse como una dimensión de tipo cultural de este mismo aspecto.

En la misma dirección, se encontró que los resultados de los estudiantes, cuyos padres parecen usar la información recibida desde las escuelas (boletines o similar), para estimular a sus hijos, muestran en promedio un mejor desempeño en el área evaluada.

Aunque la importancia de la composición es casi incuestionablemente reconocida, los mecanismos subyacentes que explican los impactos de la composición sociocultural de la escuela en el desempeño individual son menos claros. Como indica De Melo (2011), el efecto de los pares podría operar de diferentes maneras, por ejemplo, en el capital social. En cualquier caso, es importante destacar que la composición escolar, son un “atributo” fijo para la escuela, es decir, no son parte del alcance acción escolar y la política educativa. En Uruguay, particularmente en el sector público, reflejan los mecanismos institucionales para asignar estudiantes a las escuelas, basados esencialmente en un criterio territorial. En este sentido, la segmentación escolar es, en gran medida, producto de la segmentación social y residencial.

Por otra parte, el ajuste de los posteriores modelos sugiere que las niñas se desempeñan mejor en lectura que los varones. Esta información refuerza la evidencia internacional (OECD, 2016; UNESCO, 2016b) que muestra que existen desigualdades en el rendimiento académico entre niños y niñas, que estarían influenciadas por el tipo de relaciones de género predominantes que impregnan los procesos de socialización y educación.

En tercer lugar, la evidencia muestra la fuerte condición que ejerce realizar tareas que, desde la percepción del estudiante, se consideran trabajo, ya sea dentro o fuera del hogar y con pago o no, presenta una relación negativa con los desempeños.

170

En relación con los aspectos relacionados con la trayectoria escolar y la repetición de grado, los factores están fuertemente asociados y operan en el sentido esperado. La evidencia indica el poderoso efecto de la repetición sobre los desempeños. En otras palabras: el hecho de repetir al menos un grado escolar tiene efectos negativos, significativos y de importante magnitud sobre el rendimiento educativo. Una segunda interpretación refiere es que los repetidores son, por así decirlo, un grupo autoseleccionado de “malos estudiantes”, por lo que no debería sorprender que obtendrían un rendimiento promedio más bajo. Aunque este razonamiento es, en principio, correcto, se carece de evidencia sobre cómo funciona otro conjunto de características extra-académicas, no observadas directamente, como hábitos de trabajo, esfuerzo, comportamiento en el aula, lo que podría explicar tanto la repetición rendimiento en las evaluaciones.

Finalmente, respecto a las características contextuales y factores institucionales de las escuelas, los resultados indican que, controlando las variables de nivel 1 y nivel 2, el desempeño de los estudiantes de escuelas privadas no difiere significativamente de las escuelas públicas. Este aspecto, en términos generales para el caso uruguayo, ha sido ampliamente documentado desde que existen evaluaciones estandarizadas de aprendizaje (ANEP, 2015a;

2017). Las diferencias en los desempeños observada en las escuelas privadas se deben, en mayor medida, a un efecto de la selectividad socioeconómica de los estudiantes que a cuestiones específicas de la gestión escolar.

6. Consideraciones finales

El estudio de la variación en los desempeños escolares trata de responder una pregunta simple pero muy relevante: ¿Por qué las escuelas presentan resultados diferentes entre sí? Este trabajo buscó mostrar el efecto de un conjunto de características que influyen los desempeños en Lectura en TERCE.

El análisis multinivel mostró la importancia de un conjunto de variables relacionadas con el desempeño del estudiante que, con matices, ha sido ampliamente documentado por otros estudios nacionales e internacionalmente. En esta línea, se evidenciaron particularidades de la estructura y funcionamiento del sistema educativo uruguayo. El análisis de la composición de la varianza mostró que una cuarta parte (26%) de la variación en los puntajes de la prueba de Lectura corresponden a diferencias entre las escuelas, lo que determina el posible “techo máximo” del “efecto escolar” en el aprendizaje. Sin embargo, la variación se reduce sensiblemente al incorporar factores contextuales y de composición. Esta baja variabilidad entre las escuelas evidencia la rigidez y centralidad del sistema educativo. Las escuelas uruguayas tienen márgenes relativamente bajos para influir en el aprendizaje y la mejora escolar. En términos generales, la capacidad de las escuelas para influir en el aprendizaje es considerablemente limitada en comparación con el peso de las condiciones socioculturales de los estudiantes y el entorno de la escuela. No obstante, esta capacidad, a pesar de ser restringida, debe dimensionarse y expandirse para mejorar la calidad y la equidad de la educación.

Con respecto a los factores asociados con el aprendizaje, el análisis evidenció la importancia del papel de la dimensión socioeconómica de la familia, expresada de forma individual como composicionalmente en la escuela. Este es un claro indicador de la distribución desigual de la educación, típicamente observada en los países latinoamericanos. En efecto, además de los problemas de calidad, la educación uruguaya presenta serios problemas en términos de equidad. La distribución de la desigualdad no ocurre solo entre individuos, sino que se segmenta en algunos grupos específicos de estudiantes en ciertos contextos y características específicas. Dicho de otra

manera, lo datos muestran la menor variabilidad entre el nivel socioeconómico de los estudiantes que asisten a la misma escuela que entre los estudiantes que asisten a diferentes escuelas.

Estos resultados muestran que la relación entre la posición social y el rendimiento académico es directa, pero también se configura a través de una serie de experiencias, expectativas y prácticas individuales y familiares que, eventualmente, pueden ser influenciadas por la dinámica escolar. En efecto, reducir este fuerte condicionamiento es también, al mismo tiempo, la principal preocupación de cualquier sistema educativo que busque mejorar la calidad y la equidad del aprendizaje de sus estudiantes.

Debe mencionarse también que la evidencia indica, en el sentido esperado, que la repetición constituye un instrumento ineficaz de política educativa. El fuerte impacto de la repetición en los desempeños, incluso después de controlar una amplia gama de factores, requiere al menos reconsiderar la efectividad de este instrumento en el objetivo final de lograr un aprendizaje de calidad para todos los estudiantes.

Con respecto a los factores institucionales, no se observaron efectos significativos para las dimensiones consideradas. No se encontró una relación positiva entre el desempeño y la infraestructura escolar, ni con la experiencia de los directores y docentes. Ninguna de las variables de gestión, prácticas y liderazgo es estadísticamente significativa. En este sentido, las relaciones, representaciones y prácticas de los docentes no mostraron efectos directos. Tampoco el clima escolar, que según la teoría es un factor ampliamente asociado con el aprendizaje, no tuvo un efecto significativo en el aprendizaje, una vez ajustado el modelo final.

Finalmente, es importante mencionar que aun cuando los datos del TERCE están en línea con los antecedentes encontrados en estudios nacionales e internacionales, la evidencia no puede considerarse definitiva. El debate sobre la relevancia e influencia de la escuela no se zanja con un análisis de este tipo. En efecto, aunque el relevamiento de los factores asociados a los aprendizajes en este artículo ha sido amplio en base a los datos disponibles, no es posible afirmar que los factores aquí mencionados sean, de hecho, los únicos que inciden sobre los aprendizajes. En este sentido, parece claro que la investigación educativa debe profundizar las hipótesis, sobre los factores asociados para el desarrollo de instrumentos de medición que capten procesos cualitativamente más complejos. En esa dirección, parece necesario continuar profundizando sobre los elementos institucionales que inciden con eficacia sobre el rendimiento del estudiante.

Referencias

- Albernaz, A., Ferreira, F.H.G. y Franco, C. (2002). *Qualidade e equidade na educação fundamental brasileira*. Rio de Janeiro: PUC-RIO, (Texto para discussão, n. 455).
- Andrade, R. J.; Soares, J. F. (2008). O efeito da escola básica brasileira. *Estudos em avaliação educacional*, 379-406.
- ANEP (2009). *Uruguay en el segundo estudio regional comparativo y explicativo (SERCE)*. Informe nacional. Montevideo: ANEP-CODICEN.
- ANEP (2015a). *Evaluación Nacional de 6° año en Matemática, Ciencias y Lengua 2013*. Primer informe, ANEP, Montevideo.
- ANEP (2015b). *Proyecto de Presupuesto Período 2015-2019*. Tomo I, Exposición de Motivos, Montevideo: ANEP- CODICEN.
- ANEP (2017). *Primer Informe Uruguay en PISA 2015*. Montevideo: ANEP-PISA.
- Atkinson, A. (2015). *Inequality. What can be done?* Harvard: Harvard University Press.
- Blanco, E. (2008). Factores escolares asociados a los aprendizajes en la educación primaria mexicana: un análisis multinivel. *REICE - Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 6(1), 1–8.
- Blanco, E. (2011). Los límites de la escuela: educación, desigualdad y aprendizajes en México. México, D.F.: El Colegio de México, Centro de Estudios Sociológicos.
- Blossfeld, H.-P., Schneider, T. y Doll, J. (2009). Methodological Advantages Of Panel Studies: Designing The New National Educational Panel Study (Neps) In Germany. *Journal For Educational Research Online*, 1(1), 10–32.
- Cardozo, S. (2016). *Trayectorias educativas en la educación media PISA-L 2009-2014*, INEEd - Grupo de estudios sobre Transiciones Educación-Trabajo (TET), Montevideo.
- Castro, R. F. (2009). Fatores associados ao desempenho escolar na 4ª série do ensino fundamental. Em JAC Lordêlo & MV. Dazzani (orgs). *Avaliação educacional: desatando e reatando nós*[online] (pp.265-295). Salvador: EDUFBA. Disponível em <https://bit.ly/2FIDxq4>
- Carnoy, M. (2015). International test score comparisons and educational policy: A review of the critiques. <https://nepc.colorado.edu/publication/international-test-scores>.
- Cervini, R. (1999). Calidad y equidad en la educación básica de argentina. *Factores asociados al logro escolar*, 5, 83-citation_lastpage.
- Cervini, R. (2002). Desigualdades en el logro académico y reproducción cultural en Argentina. *Revista mexicana de investigación educativa*, 7(16).
- Coleman, J., Campbell, B., Hobson, C., McPartland, J., Mood, A., & Winefield, F. (1966). *Equality of Education Opportunity Report*: Washington DC: US Government Printing Office.
- Cortés, F. y Rubalcava, R. M. (1984). *Técnicas estadísticas para el estudio de la desigualdad social*. México: El Colegio de México-Flacso.
- Cowell, F.A. (1998). *Measurement of inequality*. Londres: STICERD, London School of Economics.
- De Melo, G. (2011). Peer effects identified through social networks. Evidence from Uruguayan schools, *Documento de Trabajo*, Montevideo: Decon-UdelaR.

- Enders, C. K., y Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: A new look at an old issue. *Psychological Methods*, 12(2), 121–138.
- Fernández, T. (2004). *Distribución del conocimiento escolar: clases sociales, escuelas y sistema educativo en América Latina*. México, D.F: El Colegio de México.
- Fernández, T., y Cardozo, S. (2009). Tipos de desigualdad educativa, regímenes de bienestar e Instituciones en América Latina: un abordaje con base en PISA 2009, *Páginas de Educación*, 4(1), 33-55.
- Hanushek, E.A. (2011). The Economic Value Of Higher Teacher Quality. *Economics Of Education Review*, 30(3), 466-479. <https://doi.org/10.1016/J.Econedurev.2010.12.006>
- Hanushek, E. A. (2013). Economic Growth In Developing Countries: The Role Of Human Capital. *Economics Of Education Review*, 37, 204–212. <https://doi.org/10.1016/J.Econedurev.2013.04.005>
- Harker, R. y Tymms, R. (2004). The effects of student composition on school outcomes. *School Effectiveness and School Improvement*, 15(2), 177-199.
- Hattie, J. (2002). Classroom composition and peer effects. *International Journal of Educational Research*, 37 (5), 449-481.
- Hox, J.J. (2010). *Multilevel Analysis: Techniques and Applications*. 2ª ed. New York: Routledge.
- INEED (2014). *Informe Sobre El Estado De La Educación En Uruguay 2014*. Montevideo: Ineed.
- INEED (2015). *Uruguay En El TERCE: Resultados y Prospecciones*. Montevideo: Ineed.
- INEED (2017). *Informe Sobre El Estado De La Educación En Uruguay 2015-2016*. Montevideo: Ineed.
- Kirsch, I., y Braun, H. (2020). Changing times, changing needs: enhancing the utility of international large-scale assessments. *Large-scale Assessments in Education*, 8(1), 1-24.
- Torrecilla, F. J. M., y de Becerra, F. C. (2006). *Estudios sobre eficacia escolar en Iberoamérica: 15 buenas investigaciones*. Convenio Andrés Bello.
- OECD (2013). PISA 2012 Results: What Makes Schools Successful? Resources, Policies and Practices (Volume IV). PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264201156-en>
- OECD (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education*. Paris: PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264266490-en>
- Raudenbush, S. y Bryk, A. S. (1986). A Hierarchical Model For Studying School Effects Source. *American Sociological Association*, 59(1), 1–17.
- Raudenbush, S. y Bryk, A. (2002). *Hierarchical Linear Models*. Second Edition. Ed. Sage. Thousand Oaks. CA.
- Raudenbush, S. W., & Willms, J. (1995). The estimation of school effects. *Journal of educational and behavioral statistics*, 20(4), 307-335.
- Ritzen, J. (2013). International Large-Scale Assessments As Change Agents. In Y. K. Von Davier M., Gonzalez E., Kirsch I. (Org.), *The Role Of International Large-Scale Assessments: Perspectives From Technology, Economy, and Educational Research*. Dordrecht: Springer. <https://doi.org/10.1007/978-94-007-4629-9>

- Roemer, J. E., & Trannoy, A. (2016). Equality of opportunity: Theory and measurement. *Journal of Economic Literature*, 54(4), 1288-1332.
- Soares, J.F. (2003). Quality And Equity In Brazilian Basic Education: Facts And Possibilities. *Seminário iasi de estatística aplicada: estatística na educação e educação em estatística*, 9, 2003, Rio De Janeiro. Conferência Convidada: Quality And Equity In Brazilian Basic Education: The Hlm Answers. Rio De Janeiro: IASI.
- Soares, J. F. (2004). O efeito da escola no desempenho cognitivo de seus alunos. *REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 2(2), 6.
- UNESCO (2016a). *Informe De Resultados Del Tercer Estudio Regional Comparativo Y Explicativo. Factores Asociados*. Santiago De Chile: UNESCO.
- UNESCO (2016b). *Informe De Resultados Del Tercer Estudio Regional Comparativo Y Explicativo. Logros De Aprendizaje*. Santiago De Chile: UNESCO.
- UNESCO (2016c). *Reporte Técnico Tercer Estudio Regional Comparativo y Explicativo. TERCE*. Santiago de Chile: UNESCO.
- Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. *IERI monograph series*, 2(1), 9-36..
- Von Davier, M., Gonzalez, E., Kirsch, I., y Yamamoto, K. (Orgs.). (2013). *The Role Of International Large-Scale Assessments: Perspectives From Technology, Economy, And Educational Research*. Dordrecht: Springer Netherlands. <https://doi.org/10.1007/978-94-007-4629-9>
- Wagemaker, H. (2014). International large-scale assessments: From research to policy. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 11–36). Boca Raton: CRC Press.
- Willms, D. (2006). *Learning Divides: Ten Policy questions about the performance and equity of schools and schooling systems*. Institute for Statistics, Montreal: UNESCO.
- Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record*.

Anexo 1

	Modelo Nulo (M0)				Modelo 1 (M1)				Modelo 2 (M2)			
	Coef.	e.e.	T	Sig.	Coef.	e.e.	T	Sig.	Coef.	e.e.	T	Sig.
Efectos fijos												
Intercepto	731,04	5,11	143,10		720,78	3,18	226,75		681,01	18,34	37,13	
Efectos composicionales												
[1] Índice NSE (a)					73,11	4,44	16,45	***	45,92	7,74	5,94	***
[2] Prop. padres con altas expectativas educativas (a)									53,91	20,27	2,66	***
[3] Prop. integración familiar completa (a)									44,24	19,31	2,29	**
[4] Prop. Estudiantes que trabajan (a)									-21,05	12,27	-1,72	*
[5] Prop. rezago (a)									-16,83	26,41	-0,64	
[6] Prop. educación pre-escolar (a)									7,80	21,54	0,36	
Factores individuales (nivel 1)												
[7] Índice NSE					35,77	3,71	9,65	***	18,28	4,25	4,30	***
[8] Rezago (repite al menos 1 vez)									-57,93	5,39	-10,75	***
[9] Padres con altas expectativas educativas									31,00	5,57	5,57	***
[10] Educación pre-escolar									11,18	5,43	2,06	**
[11] Uso familiar de información escolar									-19,08	4,60	-4,15	***
[12] Trabaja												
[13] Es niña												
[14] Ausentismo (1 vez a la semana o más)												
[15] Integración familiar completa									1,52	4,49	0,34	
Factores escolares (nivel 2)												
[16] Montevideo (vs. Interior)												
[17] Colegio privado (vs. Público)												
[18] Índice de infraestructura escolar												
[19] Experiencia como Director/a (en años)												
[20] Índice de liderazgo directriz												
[21] Índice de clima escolar												
[22] Experiencia como Maestro/a (en años)												
[23] Formación docente reciente												

Nota: (*) P<0.10; (**) P<0.05; (***) P<0.01

A regulação supranacional operada pelo PISA: construindo credibilidade internacional

Edite Maria Sudbrack ¹  ; Dora Ramos Fonseca ² 

¹ Universidade Regional Integrada do Alto Uruguai e das Missões (URI), Brasil; ² Universidade de Aveiro (UA), Portugal

Resumo. Este ensaio de natureza qualitativa reflete acerca da Avaliação Internacional, denominada PISA e seu potencial para induzir outras políticas educacionais, na esteira das regulações supranacionais. Com efeito, o tema da Avaliação em Larga Escala no caso, o PISA encontra sua gênese no esforço de racionalização das políticas sociais, entre as quais, a educação, a partir da crise do Estado-Providência nos países centrais. Almeja-se a análise dos expedientes de avaliação em larga escala, cuja característica estandarizada traduz sua matriz econômica ao gosto mercadológico, baseada na regulação supranacional, na publicização de resultados pela mídia, acirrando a competitividade entre escolas e redes de ensino, espelhando a esfera mercadológica e produzindo o ranqueamento entre países. Sob o pretexto de monitorar a qualidade, o PISA é portador do signo da reestruturação produtiva do capitalismo e de suas transformações. E, ainda, entende-se que o exame acaba por produzir a hierarquização dos países pela via do conhecimento. O Certame opera um processo de regulação supranacional, reforçado pelos processos de globalização hegemônica. A escolha do PISA como foco de análise, situa-se, portanto, em sua macroinfluência, determinando em tese, a formulação de políticas educacionais no Brasil e Portugal. Com efeito, as políticas educativas gestadas em nível internacional têm impactado em diferentes países, independentemente de sua cultura ou de suas idiosincrasias, como a Avaliação em Larga Escala de característica estandarizada e padronizada. O estudo se dá pela abordagem sociocrítica, nomeadamente a dos Estudos Comparados de natureza crítica.

Palavras-chave: avaliação em larga escala; regulação supranacional; política educacional; PISA.

La regulación supranacional operada por PISA: construyendo credibilidad internacional

Resumen. Este ensayo de carácter cualitativo pretende reflexionar sobre las evaluaciones internacionales y su potencial para inducir otras políticas educativas a raíz de las regulaciones supranacionales. Así, en lo que respecta a la cuestión de la evaluación a gran escala del informe PISA, nace del esfuerzo por racionalizar las políticas sociales, incluyendo la educación, a partir de la crisis del estado de bienestar en los países centrales. El objetivo es analizar los expedientes de evaluación a gran escala, cuya característica normalizada traduce su matriz económica al modo del mercado, basándose en la regulación supranacional, la publicación de los resultados por los medios de comunicación, el aumento de la competitividad entre las escuelas y las redes educativas, como reflejo de la esfera del mercado, elaborando así una clasificación entre los diferentes países. Con el pretexto de supervisar la calidad, el informe PISA porta el sello de reestructuración productiva del capitalismo y sus transformaciones. Asimismo, se entiende que el examen provoca en última instancia la jerarquización de los países a través del conocimiento. El certamen se produce en un proceso de regulación supranacional, reforzado por los procesos de globalización hegemónica. Por tanto, la elección del informe PISA como foco de análisis radica en su macroinfluencia, que determina, en resumen, la formulación de políticas educativas en Brasil y Portugal. De este modo, las políticas educativas gestionadas a nivel internacional han repercutido en diferentes países, independientemente de su cultura o idiosincrasia, como la evaluación a gran escala caracterizada por la estandarización y la normalización. El estudio se basa en el enfoque sociocrítico, es decir, los estudios comparados de carácter crítico.

Palabras clave: evaluación a gran escala; regulación supranacional; política educativa; PISA.

The supranational regulation operated by PISA: building international credibility

Abstract: This test of qualitative nature reflects on the International Assessment, called PISA and its potential to induce other educational policies, in the wake of supranational regulations. With result, the theme of the evaluation on a large scale in the case, the PISA finds its genesis in the effort to rationalize of social policies, among them, education, from the crisis of the Providence State in the central countries. It's hoped the analysis of dossiers of evaluation on a large scale, whose Standardized feature translates it's economic matrix to the interest, marketing, based on the supranational regulation, in the publication of results by the media, boosting competitiveness between schools and education networks, reflecting the marketing sphere and producing rankings between countries. The effect the pretext of monitoring quality, the PISA carries the sign of the productive restructuring of capitalism and its transformations. And, still, it is understood that the

exam ends up producing the hierarchy of countries through knowledge. The Certame operates a process of supranational regulation, reinforced by the processes of hegemonic globalization. The choice of PISA as the focus of analysis is situated, therefore, in its macro influence, determining in thesis, the formulation of educational policies in Brazil and Portugal. With result educational policies gestated in the international level have impacted in different countries, regardless of their culture or their idiosyncrasies, such as the evaluation on a large scale of characteristic standardized and standardized. The study is based on the sociocritical approach, namely the Comparative Studies of a critical nature.

Keywords: evaluation on a large scale; supranational regulation; educational politics; PISA.

1. Primeiras palavras

Este ensaio de natureza qualitativa reflete acerca da Avaliação Internacional, denominada PISA e seu potencial para induzir outras políticas educacionais, na esteira das regulações supranacionais. Com efeito, o tema da Avaliação em Larga Escala no caso, o PISA (*Programme for International Student Assessment*) encontra sua gênese no esforço de racionalização das políticas sociais, entre as quais, a educação, a partir da crise do Estado-Providência nos países centrais.

Segundo leciona Afonso “trata-se de tentar conciliar o Estado-avaliador-preocupado com a imposição de um currículo nacional comum com o controlo dos resultados e- a filosofia do mercado educacional [...]” (2000, p. 119). Este retrato da realidade portuguesa também se aplica ao cenário brasileiro.

Nesta dimensão, a crise do Estado-Nação, notadamente nos países centrais, tem impactado sobre a formulação de políticas educacionais, face aos determinantes da globalização e do novo capitalismo. Tal impacto também se alastra sobre os países capitalistas periféricos, parecendo seguir uma intencionalidade de controle e intervenção, na esteira da regulação supranacional. Conforme vaticina Afonso (2009), o estado desvincula-se das funções de produtor de bens e serviços para se transformar em regulador do processo de mercado.

Neste artigo almeja-se a análise dos expedientes de avaliação em larga escala, cuja característica estandardizada traduz sua matriz econômica ao gosto mercadológico, baseada na regulação supranacional, na publicização de resultados pela mídia, acirrando a competitividade entre escolas e redes de ensino, espelhando a esfera mercadológica e produzindo o ranqueamento entre países.

Diversos estudos revelam que se regista uma propagação do ideário neoliberal no campo educativo traduzido na expansão das ideias de *mercado* ou *quase mercado educacional* uma *faceta gestonária* utilizada com uma *linguagem bélica* ao nível do discurso político na área da educação, como afirmou contundentemente Formosinho & Machado (2007). São vários os investigadores (Seixas, 2001; Sá, 2002; Barroso & Viseu, 2003; Lima, 2011) que nos alertam para essa forte orientação nas políticas educativas em muitos países. É nosso pressuposto que as políticas educacionais estão profundamente marcadas pelas produções e pelas decisões de entidades supranacionais que hoje influenciam claramente as políticas à escala supranacional (Castro, 2011). São vários os organismos desse tipo que contribuem formal ou informalmente para a regulação (transnacional) da educação nos diferentes países (Azevedo, 2000; Seixas, 2001; Teodoro, 2001; Neto-Mendes, 2004; Dias, 2004; Barroso, 2006; Nóvoa, 2009; Castro, 2011). As instituições como o Banco Mundial, o Fundo Monetário Internacional (FMI), a Organização para a Cooperação e o Desenvolvimento Económico (OCDE), a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO), a União Europeia e o Conselho da Europa, entre outras, têm marcado a direção das políticas educativas nos diferentes países do mundo. Na esteira de António Teodoro (2001), dizemos que o *gerencialismo global* faz deslocar, progressivamente, o poder dos vários Estados-Nação para organizações supranacionais as quais aos poucos assumem o controlo das agendas políticas em educação.

Numa perspectiva comparada entre Portugal e Brasil, projeta-se a análise dos mecanismos de Avaliação em Larga Escala, nomeadamente o PISA, na indução de políticas públicas, com potencial, em tese, para induzir políticas educativas. Tal proposta merece um esforço analítico por ser portadora de influência, sobretudo no modo de pensar o campo político (Correia, 2010). Os mecanismos de regulação à escala transnacional impõem as “lentes” para a observação e interpretação das realidades educacionais definindo à priori e de forma uniformizante as grelhas de análise e lógicas de pensamento e ação. Como aponta Nóvoa (2009, p.29), temos assistido a uma “série de classificações internacionais que não se limitaram a descrever uma situação, mas que constituíram categorias de pensamento, que definiram modalidades de ação”.

As narrativas da Avaliação Externa ausentaram do debate a ideia de processo, em outras palavras, há um deslocamento da promoção da qualidade, reduzindo-a a um produto (Nóvoa, 2009, p. 457). A primazia na agenda dos governos de diferentes países tem incorporado uma cultura avaliativa massificante, a partir da adesão das nações às aferições internacionais, como é o caso do PISA, justificando as reformas educacionais de toda a ordem.

Sob o pretexto de monitorar a qualidade, o PISA é portador do signo da reestruturação produtiva do capitalismo e de suas transformações. E ainda, entende-se que o exame acaba por produzir a hierarquização dos países pela via do conhecimento. Sob tal vértice, o Certame opera um processo de regulação sobre as diferentes nacionalidades, reforçado pelos processos de globalização hegemônica.

Ao problematizar os exames do PISA e do TIMSS, depreende-se que os mesmos têm identificado fragilidades nos sistemas escolares e têm frequentemente oferecido aos governos argumentos para a reformas educacionais rápidas. Em contexto social de redefinição do papel do Estado e da lógica mercantil que tem preponderado em vários países, a Avaliação parece assumir uma mudança nos modos de planejamento e gestão de políticas educacionais, sob a égide da regulação supranacional.

A escolha do PISA como foco de análise, situa-se, portanto, em sua macroinfluência, determinando em tese, a formulação de políticas educacionais no Brasil e Portugal. Com efeito, as políticas educativas gestadas em nível internacional têm impactado em diferentes países, independentemente de sua cultura ou de suas idiossincrasias. Um exemplo marcante é a Avaliação em Larga Escala de característica estandardizada e padronizada, nomeadamente o PISA e o caráter regulador que opera.

180

Está presente na narrativa da exacerbada valoração dos processos avaliativos, a hierarquização do conhecimento, o que leva a indagação sobre que conhecimento vale mais. Sob esta racionalidade se estabelecem comparativos entre os países, a partir dos produtos aferidos nos exames internacionais, como o PISA.

O poder atribuído aos resultados do Certame tem facultado aos governos, independente de corrente ideológica, realizar reformas educacionais assentadas nos escores produzidos. Vale dizer, governa-se por números, desconsiderando os processos democráticos de eleição de prioridades, tomada de decisões e de participação.

2. Avaliação no mundo globalizado

A globalização hegemônica produziu uma macroeconomia pautada no exacerbado individualismo, no endeusamento do mercado, gerando grande contingente de excluídos numa relação pautada pela competitividade em todos os níveis. Segundo Bauman (1999), é o destino irremediável do mundo, um processo irreversível; é também um processo que nos afeta, a todos, na

mesma medida e da mesma maneira. De acordo com Santos (2004), o que caracteriza os tempos globalizados é a confusão dos espíritos que impede o nosso entendimento do mundo, do lugar, da sociedade e de cada um de nós mesmos.

O individualismo arrebatador presente na ordem social e individual acaba por construir o outro como objeto. Instaura-se uma nova lei sobre o valor das coisas, em que vai desaparecendo o apreço pela saúde, emprego, educação, dentro outros, enquanto bens inalienáveis do ser humano. Ao referir-se à globalização, Santos (2004, p. 18) fala da influência de três mundos em um só: “O primeiro seria o mundo tal como fazem vê-lo: a globalização como fábula; o segundo seria o mundo tal como ele é: a globalização como perversidade; e o terceiro, o mundo como ele pode ser: outra globalização”.

Face a tantos reptos impostos às políticas educacionais, novas exigências são dirigidas à educação, o representando terreno fértil para acolher as avaliações em larga escala, nomeadamente as legitimadas pela OCDE e suas narrativas.

Sob este cenário, o projeto social transforma-se e influencia as políticas educacionais de maneira insofismável. Tributário da globalização hegemônica amplia a financeirização da economia enquanto modelo socioeconômico e cultural, cuja agenda determina novos sentidos à cultura e à educação, alçando-as à condição de mercadoria, cujo compromisso maior está no lucro de quem a gerencia. O incremento da globalização econômica, política e cultural exerce impactos significativos sobre a avaliação da educação e sugere maior relação entre o setor produtivo e suas necessidades. Tal ambiente busca modelos comparativos entre países. Assim, a OCDE apela para as reformas nos serviços públicos, calcadas, no desempenho eficaz. (Cury, 2017).

Entre os desdobramentos deste ideário (Santos, 2004) aludem à aplicação das teorias da *new public management*, cujo *modus operandi* considera que é preciso atuar segundo padrões empresariais de eficiência e eficácia e com novas formas de gestão, baseadas em contratos e serviços terceirizados, na mesma lógica econômica. Esta agenda situa-se na esteira de Dardot e Laval (2016, p. 17), quando afirmam “o neoliberalismo antes de ser uma ideologia ou uma política econômica, é em primeiro e fundamentalmente uma “racionalidade” e, como tal, tende a estruturar e organizar não apenas a ação dos governantes, mas a própria conduta dos governados”.

A globalização, no entender de Estêvão (2002) é multifacetada, ainda que sobressaia-se a perspectiva econômica neoliberal. No caso da OCDE, através do PISA, acentua-se seu caráter vertical, e o poderio das

nações dominantes transnacionalizando sua influência. Eis que a “globalização competitiva” (Arruda, 2000), encaixa-se hermeticamente nos propósitos do Inquérito internacional, enredando as nações em seus valores e na homogeneização de políticas e discursos.

Ainda que este modelo não deva levar ao fatalismo de que de não há outra saída, também não se pode ser ingênuo, desconhecendo a grande sobre determinação da globalização pela ideologia e práticas neoliberais. O cenário agudiza-se com a fragilidade e a vulnerabilidade dos estados nacionais, fruto das desigualdades sociais.

Há que estar atento, problematizando esta dinâmica global, ou como ensina Estêvão (2016, p. 97, 98):

[...] não se pode deixar de questionar a “prosperidade patológica” do *boom* consumista a que vimos assistindo; [...] ficando indiferentes as injustiças resultantes de um capitalismo multipolar sujeito a movimento de suas “placas tectônicas” de impacto nem sempre previsível.

O ambiente narrado é o território que fomenta a avaliação, baseado no desempenho, eficiência, eficácia e performatividade.

3. O fetiche da avaliação

À semelhança do fetichismo do mercado, expressão cunhada por Karl Marx, de que a mercadoria opera como se fosse portadora de vontade própria, a avaliação parece ter-se transmutado em alvo de adoração dos planejadores das políticas educacionais dos países, como se fossem detentoras de um poder mágico de por si só atribuir qualidade à educação.

O tema da avaliação externa em todos os níveis da educação tem sua gênese no esforço de racionalização das políticas sociais, inclusive a educação, partir, da crise do Estado Previdência. O discurso do Estado avaliador encontra sua fonte também numa racionalidade instrumental/reguladora, na qual a imposição dos valores da eficiência e eficácia forjam um pensamento único, traduz como legítimo. Nestas lógicas mercantis, a verificação e o controle dos resultados educativos tornaram-se aspetos centrais na ação/regulação do Estado em relação às escolas, reforçando a sua faceta de *Estado-avaliador*. Seguindo as ideias de Estêvão (2008, p. 506), o Estado acaba por entrar “na lógica dominante, a mercantil, “desmonumentalizando-se” e deixando o seu antigo poder imperial, mas assegurando a difusão das ideias da eficiência e eficácia, dando assim o seu contributo para o funcionamento do mercado”.

Os novos modos de regulação das políticas públicas em âmbito transnacional assentam-se no conhecimento que é hierarquizado pelas avaliações e legitimado por decisões políticas dos diferentes países. O PISA é um instrumento por excelência a serviço de tal propósito.

Esta narrativa ajusta-se ao desenvolvimento econômico e tecnológico para transmutar-se em uma racionalidade instrumental que hierarquiza o conhecimento. Perdem-se, com efeito, os espaços de deliberação democrática, de pensamento crítico e autônomo, substituindo-os por valores do mercado. Estes contextos, de traços predominantemente reguladores, conduzem a valorizar opções baseadas na meritocracia e não na ocorrência de uma avaliação mais democrática. Em estudo realizado por Afonso e Costa (2009), os mesmos apontam para a existência de uma regulação baseada no conhecimento, denominada de KRT (*Knowledge Based Regulation-Tools*), a qual define cada vez mais a posição dos países no cenário educacional e econômico.

A avaliação transfigura-se em fenômeno político, eis que envolve programas, redes de ensino, projetos, faz uso de recursos públicos. Seu impacto político nos sistemas educacionais se alarga para a sociedade como um todo. Apoia-se em reformas educacionais, seja nos currículos, na gestão, no financiamento, na assistência técnica, nas políticas e programas, entre outros.

183

As características da avaliação em larga escala dizem respeito a planos de longo prazo, aplicadas em períodos de tempo delimitados que permitem a análise temporal de evolução ou não-evolução. Mais que produzir novos dados, os sistemas de ensino devem gerar debates e reflexões sobre os indicadores existentes, na perspectiva da tomada de decisões, de interpretação dos resultados, não se detendo apenas ao produto, mas no processo. Há que se imprimir significação a estes elementos, à luz das realidades vividas. Nesta perspectiva, pode inserir-se a promoção da equidade da oferta educativa, induzindo novas políticas ou programas para fazer frente às necessidades diagnosticadas.

A carga de intervenção nos Estados nacionais se dá pelo uso das informações praticadas, redundando em uma “forma de governação” (Afonso & Costa, 2009, p. 1040). De forma análoga, os dados estatísticos, os rankings e os comparativos publicados são meios de pressão sobre os sistemas, exigindo-lhe a prestação de contas, compondo, a partir da instrumentação abundante, um quadro de interpretação do mundo, segundo sua ótica.

A contribuição de Afonso (2009) sinaliza que a avaliação estandarizada criterial apresenta-se como aliada ao controle do Estado sobre os sistemas educativos e seus profissionais. Vale dizer que, o predomínio de

testes e exames é uma realidade presente em vários países (Afonso, 2009). O autor prossegue informando que a avaliação produz rankings, ou seja, efeitos de mercado. Tais resultados, segundo o autor (Afonso, 2009) são assumidos de forma acrítica, como se fossem técnica, cultural e politicamente neutros.

As reflexões levadas a termo pela academia e pelo coletivo dos docentes podem imprimir tensões às visões dominantes, através da proposição de avaliações que envolvam múltiplas dimensões. Para o autor (Afonso, 2009) impõe-se a construção de modelos democráticos e transparentes de avaliação, prestação de contas e de responsabilização (*accountability*), implica igualmente a valorização social, cultural e política dos processos de participação, negociação e justificação, e, adesão de modelos exemplificados de justiça e equidade social, educacional e avaliativa.

Ventura (2016) refere que a avaliação deve ser um meio de “partilhar e transferir conhecimento, competência e poder” (p. 15).

4. O PISA

184

O Programa Internacional de Avaliação de Alunos (PISA), criado em 1997 e aplicado em 2000, afere conhecimentos e habilidades dos alunos na faixa etária de 15 anos de idade, 7ª série ou 8º ano. Realizado nos países da OCDE e em países convidados, acontece a cada 3 anos, avaliando as áreas de leitura, matemática e ciências.

O Inquérito constitui-se em dispositivo internacionalmente aceito, que compara desempenhos escolares entre países, revelando-se num importante mecanismo de ações da OCDE que se justifica, segundo seus argumentos, para medir a qualidade da educação. Para a OCDE “O PISA é um esforço colaborativo, que reúne a expertise científica dos países participantes”. Além da dimensão técnica, o PISA amealha capital político, pela validação internacional que lhe é conferida, não só pelos países membros da OCDE, mas também pelos que participam do Certame como convidados.

Ainda que a OCDE não expresse literalmente em sua missão institucional a funcionalidade da regulação (Jakobi, 2012), são visíveis suas interferências na regulação das políticas públicas, notadamente, as educativas, nas duas últimas décadas. Entre as Políticas Públicas a educação tem recebido profunda carga de influência dos organismos internacionais, acentuando-se a regulação que opera em caráter transnacional (Lemos, 2014)

Não por acaso, a definição de políticas públicas de educação tem interessado a estes organismos deixando de ser atributo exclusivo dos estados nacionais. Com efeito, os grandes destinos da educação são gestados em nível supranacional, além fronteiras. No momento em que o conhecimento é a grande moeda, passa a ocupar um lugar no coração da competição econômica e- por extensão- a transmissão desses conhecimentos, designadamente, a educação (Moutsios, 2009).

Torna-se, assim, estratégico para as agências multilaterais participarem tão avidamente na produção de diretrizes para conduzir a educação. Para Carvalho (2011, p. 15), o PISA configura-se como um dispositivo técnico e social, pelo fato de que “associa (i) procedimentos de medida e meios de conhecimentos a (ii) interpretações acerca do campo educacional e da acção educativa, bem como acerca dos modos através dos quais devem ser regulados”.

O autor (ID) prossegue afirmando que o PISA, não apenas aponta capacidades de aprendizagem dos países, mas, sobretudo, sinaliza para quais modos de aprender, os que possuem potencial competitivo mundial. Fabrica e circula, portanto, uma determinada tecnologia de conhecimento e políticas. Tal contexto para a política, objetiva designar que saberes podem orientar os gestores em seu planejamento e ação.

185

Carvalho e Costa (2011) alertam que o PISA é muito mais do que um grande Inquérito. Mas que: “O Programa é um espaço” de organização que põe em cooperação mundos sociais muito diversos- investidores de centros de pesquisas públicos ou privados, peritos, profissionais da OCDE, decisores políticos, burocratas e tecnoestruturas da administração- e que recorre a sofisticados meios de coordenação entre produtores de conhecimento pericial e entre estes e os atores que operam nos campos da política, da administração, da educação e da investigação (p.41). Estas alianças atuam no sentido de que o PISA, não só se mantenha em evidência, e, com prestígio, como faz ampliar seu raio de ação.

A OCDE, cujo foco destacamos é uma organização com alto grau de sofisticação. Em tese recente Lemos (2014) descreve o grau de aprimoramento da direção de educação da OCDE, a qual, por exemplo, é composta por vários programas, entre os quais o Comitê do PISA (OCDE, 2014). A manutenção de sua influência se dá também pela literatura que produz. Os discursos da OCDE são veiculados através de documentos, textos, livros, base de dados, papers, entre outros, destacando-se como uma das maiores editoras do mundo neste campo.

A OCDE tem consciência desta ingerência que causa e a monitora através de levantamentos, que apontam seus impactos nas políticas educativas dos países. No caso deste estudo que envolve o PISA, o exame é legitimado, pela adesão dos países membros, como Portugal, e também dos países convidados como o Brasil, os quais levam em conta seus resultados na formulação de medidas educacionais.

As críticas dirigidas ao PISA, pelo seu caráter estandardizado é resumidamente apontada por Fernandes (2004), quando assinala os contextos, os currículos, o tipo de amostragem, entre outras discrepâncias entre países, como dificuldades em estabelecer comparativos.

São muitas as ressalvas atribuídas ao Certame, para efeito deste trabalho, concentramos o esforço na análise da categoria KRT com base em Lemos (2014), que, sem minimizar outras interferências representa nossa lente de análise, em consonância com nossos objetivos de investigação, de que a mesma parece estabelecer uma nova governação, uma regulação supranacional. A avaliação em Larga Escala, denominada PISA (*Programme for International Student Assessment*), idealizada pela OCDE (Organização para a Cooperação e Desenvolvimento Econômico) acontece a cada três anos nos países membros da organização. Na construção da proposta, vislumbra-se uma resposta da OCDE às próprias exigências/necessidades dos países membros de disporem de dados sobre conhecimentos e competências dos sistemas educacionais (Afonso & Costa, 2009).

Para alguns, esta avaliação tem o mérito de divulgar uma base de dados e informações sobre o desempenho cognitivo, propiciando o cruzamento de indicadores entre países, escolas e estudantes testados, acompanhando assim a evolução das performances dos países envolvidos. Fica visível a diferença entre os melhores e os piores desempenhos, acirrando as desigualdades educacionais (Brasil, 2016). Segundo a mesma fonte (Brasil, 2016), no Brasil, a média do PISA está em ascensão, a evolução mais positiva foi obtida por Portugal, Itália, entre outros países.

O PISA inscreve-se na perspectiva da regulação da educação, como política pública, atendendo aos propósitos de uma economia baseada no conhecimento. Neste particular, combinam-se a regulação e os elementos de mercado ao domínio público, que, segundo Afonso (1999), aumenta o controle sobre as escolas, bem como inserem dispositivos de publicização dos resultados, promovendo a competitividade entre as redes de ensino.

O uso do Indicador Internacional enquanto processo de decisão política, difunde um tipo especial de conhecimento, no caso do exame de avaliação de desempenho, na leitura, nas ciências e na matemática e suas

literacias. Assim, a OCDE acaba por regular a política educativa, a partir de instrumentos de avaliação em larga escala, como o PISA, comparando desempenhos entre países e sistemas de ensino.

Pelo fato de ser coordenado pela OCDE e sua experiência em estudos comparativos, é portador de uma lógica que o legitima. Assim, segundo Afonso e Costa (2009), esta visão legitimadora é veiculada pela credibilidade, pela maleabilidade, pela relevância simbólica e pela pertinência (Afonso e Costa, 2009. p. 1043), que lhe é atribuída.

Segundo os mesmos autores, a OCDE apresenta grande credibilidade de seus métodos; a maleabilidade envolve a possibilidade que o instrumento possui de ser adaptado a diferentes contextos nacionais; a relevância simbólica reside no simbolismo dos processos de adesão dos países, dando-lhes visibilidade e integração às agendas internacionais; a pertinência dá ao PISA, grande relevância política, já que conscientiza, indica problemas e valoriza os bons indicadores dos sistemas avaliados. Para Afonso (1999), esta modalidade de avaliação estandardizada é congruente com o mercado, com o caráter normativo da política, e não formativo ou emancipatório.

5. Questão metodológica

Este ensaio inscreve-se como um estudo de natureza qualitativa. A abordagem do problema “A regulação supranacional operada pelo PISA: construindo credibilidade internacional”, deu-se pela abordagem sociocrítica, nomeadamente a dos Estudos Comparados de natureza crítica.

Os estudos comparados precisam ter em mente a influência da globalização sobre os sistemas de ensino nos distintos países, neste caso, entre Portugal e Brasil. Não é possível estabelecer uma relação casual à moda das agências internacionais com o processo de globalização. Sabe-se que estas agências têm presente os princípios da teoria econômica, ou seja, de que a educação é fator de competitividade internacional. Segundo Sousa, Castro e Rothes (2013).

“Diferentemente do passado, no presente século a Educação Comparada não se encontraria necessariamente associada aos processos de uniformização ou homogeneização dos sistemas de ensino”, mas antes se deslocando “dos sistemas de ensino para as escolas, das estruturas para os atores sociais, do plano das ideias para o do discurso, dos fatos para a dimensão política (Nóvoa, 1997)”. Tal como entendem (Ferreira, 2009; Madeira, 2009; Schriewer, 2009) é necessário encontrar e identificar novas

problemáticas, adotar novos modelos de análise de forma a produzir novos sentidos para os processos educacionais. Estes elementos acompanham nossa reflexão, eis que nosso estudo pauta-se pelo comparatismo crítico. Esta abordagem metodológica realiza uma ruptura com o comparatismo tradicional em educação, seja no plano político, seja opondo-se a formas consensualistas etnocentradas, que tomam os países colonialistas como superiores.

Os desafios da globalização impõem novas respostas da educação comparada. A uniformização dos discursos passa a exigir “a necessidade de ver reconhecidos e tomados em consideração espaços identitários mais restritos que o espaço nacional” (Mallet, 2004, p. 1313).

Ao categorizar a educação como um serviço, a OCDE, contribui para ampliar sua subserviência ao mercado, traduzindo-se num processo de internacionalização, o qual amplifica a responsabilização dos sistemas e dos atores, via avaliações em larga escala. Neste contexto, os estudos comparados são chamados a rever suas óticas de análise, eis que tais espaços pelo volume de dificuldades tencionam para novos posicionamentos.

188

A categoria definida, a priori, como geral, a Regulação transnacional, justifica-se, na tese de que o PISA consubstancia-se como instrumento regulador das políticas em nível transnacional, estabelecendo-se como carro-chefe das políticas educativas da OCDE. Seus argumentos remetem à credibilidade que ostentam perante os países membros, caso de Portugal e também os convidados como é o caso do Brasil, bem como às relações entre conhecimento e política (Carvalho, 2013). A narrativa do PISA difunde regras e normas, estimulando um espaço ora de competição, ora de suposta colaboração. Complementarmente ao ideário que veicula, estabelece normativas de governação para os gestores educativos. A máxima da OCDE revela-se emblemática: “E lembre-se: sem dados, você é apenas mais uma pessoa com opinião” (Powerpoint, Tóquio, 2005 *apud* Carvalho, 2013, p.66) Sem nenhum pudor, o discurso é enfático: “há que pensar a educação pela racionalidade econômica, a qual se subordina à competitividade” (ID, 2013, p.66).

A Regulação Supranacional opera pela veiculação de um léxico em que o gestor é “um solucionador de problemas, o aprendiz, o decisor político razoável que governa com e através da monitorização [...]”.

6. A guisa de conclusão preliminar

Este ensaio de natureza qualitativa buscou refletir acerca da regulação supranacional operada pela OCDE através do PISA, passando pela análise de seu prestígio, credibilidade, bem como, pela capacidade de induzir reformas educativas em diferentes países. Há um conjunto de mecanismos, por parte da OCDE, que impulsiona e administra estes países sobre a política que produz sua estratégia de comunicação e trocas com atores fulcrais do campo, imprimindo uma forte razão pragmática para o Inquérito.

Com efeito, o PISA mantém-se. Nas palavras de Carvalho e Costa (2011, p. 73) “O PISA vive e sobrevive não para resolver problemas, mas porque cria nos seus utilizadores a percepção de que lhe é útil para resolverem problemas que imaginam serem os seus”.

À pesquisa em educação compete deslindar esta narrativa, propondo outros usos para a Avaliação baseada em processos de tomada de decisões e emancipações, para além do cariz econômico.

Referências

- Afonso, A. J. (1999). Estado, mercado, comunidade e avaliação: esboço para uma rearticulação crítica. *Educação & Sociedade*, 20(69), 139-164.
- Afonso, A.J. (2000) *Avaliação Educacional: regulação e emancipação*. São Paulo: Cortez.
- Afonso, A.J. (2009). Nem tudo o que conta é mensurável ou comparável. Crítica à accountability baseada em testes estandardizados e rankings escolares. *Revista Lusófona de Educação*, 13(1), 13-29.
- Afonso, N. & Costa, E. (2009). A influência do Programme for International Student Assessment (PISA) na decisão política em Portugal: o caso das políticas educativas do XVII Governo Constitucional Português. *Sísifo: revista de ciências da educação*, 10. Disponível em <https://bit.ly/33vRtvC>.
- Arruda, M. (2000). Globalização e sociedade civil: repensando o cooperativismo no contexto da cidadania activa. In: M. Arruda & L. Boof. *Globalização: desafios socioeconômicos, éticos e educativos*. Petrópolis: Vozes.
- Bauman, Z. (1999). *Globalização: as consequências humanas*. RJ: Jorge Zahar.
- Barroso, J. (2006). *A Regulação das Políticas Públicas de Educação: Espaços, dinâmicas actores*. Lisboa: Educa.
- Barroso, J. & Viseu, S. (2003). A emergência de um mercado educativo no planeamento da rede escolar: de uma regulação pela oferta a uma regulação pela procura. *Educação, Sociedade*, 24(84), 897-921

- BRASIL (2016). INEP. *Censo de Educação Superior 2015*. Brasília. Disponível em <https://bit.ly/32wlsTQ>.
- Carvalho, L.M. (coord.). (2011). *O espelho do perito: Inquéritos internacionais, conhecimento e política em educação- o caso do PISA*. Fundação Manuel Leão.
- Carvalho, L. M. (2013). Mútua vigilância organizada. *Educação: Temas e Problemas*, (12_13), 61-74. Disponível em <https://bit.ly/2ZDWPnm>.
- Carvalho, L.M. & Costa, E. (2011). Fabricando o espelho do perito: a construção de conhecimento pericial no mundo do PISA. In: L.M. Carvalho [coord.]. *O espelho do perito: Inquéritos internacionais, conhecimento e política em educação- o caso do PISA*. Fundação Manuel Leão.
- Castro, D.P. (2011). *Auditoria, contabilidade e controle interno no setor público: integração das áreas do ciclo de gestão: contabilidade, orçamento e auditoria e organização dos controles internos como suporte à governança corporativa*. 4. ed. São Paulo: Atlas.
- Correia, J.A. (2010). Paradigmas e cognições no campo da administração educacional: das políticas de avaliação a avaliação como política. *Revista Brasileira de Educação*, 15(45).
- Cury, C.R.J. (2017). A globalização e os desafios para os sistemas nacionais: agenda internacional e práticas educacionais nacionais. *RBPAE*, 33(1), 15-34.
- Dardot, P. & Laval, C. (2016). *A nova razão do mundo: ensaio sobre a sociedade neoliberal*. São Paulo: Boitempo.
- Dias, M. (2004). As políticas “locais” de educação e a pro:ssão de professor: novos contextos de trabalho, novas identidades pro:ssionais. In Jorge Adelino Costa, Neto-Mendes, António e Ventura, Alexandre (org.). *Políticas e gestão local da educação*. Aveiro: Universidade de Aveiro, pp. 225-265.
- Estêvão, C. (2008). Educação, conflito e convivência democrática. *Ensaio Avaliação e Políticas Públicas em Educação*, 16, 503-512
- Estêvão, C.A.V. (2016). Educação, justiça e direitos humanos. *Educação e Pesquisa*, 32(1), 85-101. Disponível em <https://bit.ly/2FmEsws>.
- Estêvão, C.A.V. (2002). Justiça complexa e educação: uma reflexão sobre a dialectologia da justiça em educação. *Revista Crítica de Ciências Sociais*, 64, 107-134. Disponível em <https://bit.ly/3iy7mYz>.
- Fernandes, D. (2004). *Avaliação das aprendizagens: uma agenda, muitos desafios*. Texto Editora. Disponível em <https://bit.ly/3hwJpig>.
- Ferreira, A. (2009). O sentido da educação comparada: uma compreensões sobre a construção de uma identidade. In: Souza, Donaldo Bello de, Martinez, Silvia Alícia. (Orgs.). *Educação comparada: Rotas de além mar*. São Paulo: Xamã.
- Formosinho, J. & Machado, J. (2007). A Modernidade, Burocracia e Pedagogia. In Jesus Maria Sousa e Carlos Nogueira Fino (Org.). *A Escola Sob Suspeita*. Porto: ASA, pp. 97-119
- Jakobi, T. (2012). *Regulating regulation? The regulatory of the OCDE*. Conference, exeter.
- Lemos, V.V. (2014). *A influência da OCDE nas políticas públicas de educação em Portugal*. Instituto Universitário de Lisboa. Disponível em <https://bit.ly/3bYQoK5>.
- Lima, L. (2011). Políticas educacionais, organização escolar e trabalho dos Profesores. *Educação: Teoria e Prática*, 21(38), 1-18. <https://bit.ly/3d7RBpw>

- Madeira, A. (2009) O campo da educação comparada: do simbolismo fundacional à renovação das lógicas de investigação. In: Souza, Donaldo Bello de, Martinez, Silvia Alícia. (Orgs.). *Educação comparada: Rotas de além mar*. São Paulo: Xamã.
- Mallet, R. (2004). Do estado-nação ao espaço-mundo: as condições históricas da renovação da educação comparada. Campinas: *Educação e Sociedade*, 25(89), 1301-1332. Disponível em <https://bit.ly/35DmLmM>
- Moutsios, S. (2009). *The transnationalization of. Educacion policy-making: dominant institutions and policies*. Saarbrucken, Germany: UDM.
- Neto-Mendes, A. (2004). Regulação estatal, auto-regulação e regulação de mercado: subsídios para o estudo da pro;ssão docente. In Jorge Adelino Costa, António Neto-Mendes e Alexandre Ventura (org.). *Políticas e Gestão Local da Educação*. Aveiro: Universidade de Aveiro, pp. 23-33
- Nóvoa, A. (2009). *Professores: Imagens do futuro presente*. Lisboa: Educa.
- Nóvoa, A. (coord.). (1997). *Os professores e sua formação*. Lisboa: Dom Quixote.
- OCDE (2014). *Relatório Nacional PISA 2012-Resultados Brasileiros*. Disponível em <https://bit.ly/3c7c7WQ>
- Sá, V. (2002). As políticas de escolha da escola pelos pais: da bondade das intenções à desilusão das realizações, ou talvez não! In Jorge Adelino Costa, António Neto-Mendes e Alexandre Ventura. *Avaliação de Organizações Educativas*. Aveiro: Universidade de Aveiro, pp. 69-92
- Santos, B.S. (2004). Do pós-moderno ao pós-colonial: e para além de um e outro. In: *Congresso Luso Afro Brasileiro de Ciências Sociais*. Conferência de Aberta/Coimbra, 16 de setembro de 2004.
- Schriewer, J. (2009). Aceitando os desafios da complexidade: metodologias de educação comparada em transição. In: Souza, Donaldo Bello de, Martinez, Silvia Alícia. (Orgs.). *Educação comparada: Rotas de além mar*. São Paulo: Xamã.
- Seixas, A. M. (2001). Políticas educativas para o ensino superior: A globalização neoliberal e a emergência de novas formas de regulação estatal. In Stephen Stoer, Luísa Cortezão, José Correia (org.). *Transnacionalização da Educação. Da crise da educação à "educação" da crise*. Porto: Afrontamento, pp. 209-238
- Souza, D. B. D., Castro, D. F., & Rothes, L. (2013). Políticas de descentralização da educação no Brasil e em Portugal: avanços e recuos da desconcentração de poderes. *Revista Portuguesa de Educação*, 26(2), 7-33.
- Teodoro, A. (2001). Organizações internacionais e políticas educativas nacionais: a emergência de novas formas de regulação transnacional, ou uma globalização de baixa intensidade. In: S.R. Stoer, L. Cortezão & J.A. Correia (org.). *Transnacionalização da Educação: da crise da educação à "educação" da crise*. Porto, Portugal: Edições Afrontamento.
- Ventura, A. (2016). *Entrevista à Info Geekie- a avaliação externa sob olhar de três especialistas*. Disponível em <https://bit.ly/2FyUi6K>

España vs. Portugal en educación. Una aproximación sistémica

Francisco López Rupérez¹  <https://orcid.org/0000-0003-2613-9652>; Isabel García García¹  <https://orcid.org/0000-0002-9884-6148>

¹ Universidad Camilo José Cela (UCJC), España.

Resumen. Como consecuencia de las exigencias del contexto y de su rápida transformación, las sociedades modernas se hallan confrontadas a la necesidad imperiosa de mejorar sus sistemas educativos. Una práctica de utilidad para la mejora de las organizaciones consiste en compararse con otras que, partiendo de situaciones inferiores, han conseguido avanzar de un modo sustantivo; para, posteriormente, tratar de averiguar cómo lo hicieron. El presente trabajo asume esa estrategia y adopta Portugal como elemento de comparación. Para ello, parte de una aproximación sistémica a la educación sobre la base de la estructura: contexto, *inputs*, políticas, *outputs*, y *outcomes*, que sirve como marco conceptual para seleccionar un conjunto amplio de indicadores sobre el que efectuar una comparación suficientemente sistemática entre los dos países, en materia educativa. Del análisis de las diferencias emergen las siguientes conclusiones: (a) en materia de variables de contexto, España se sitúa por encima de Portugal; (b) en materia de *inputs*, Portugal supera a España en gasto acumulado por alumno y está retrasada en dotación de recursos educativos y humanos; (c) en cuanto a *outputs* y a *outcomes*, Portugal supera a España en todos los indicadores considerados; (d) finalmente, en relación con las políticas, la posición de Portugal está claramente más adelantada que la de España. Estas evidencias se traducen en un conjunto de recomendaciones, inspiradas en el consenso internacional y validadas por el éxito del caso portugués.

Palabras clave: educación comparada; enfoque sistémico; educación española; educación portuguesa; evaluación.

Espanha vs. Portugal na educação. Uma abordagem sistêmica

Resumo. Como consequência das exigências do contexto e de sua rápida transformação, as sociedades modernas são confrontadas com a necessidade urgente de melhorar seus sistemas educacionais. Uma prática útil para o aprimoramento das organizações consiste em comparar-se com outras que, partindo de situações inferiores, conseguiram avançar de forma substancial; para, depois, tentar descobrir como o fizeram. O presente trabalho assume essa estratégia e adota Portugal como elemento de comparação. Para isso, parte de uma abordagem sistêmica da educação baseada na estrutura: contexto, *inputs*, políticas, *outputs* e *outcomes*, que serve como uma estrutura conceitual para selecionar um amplo conjunto de indicadores sobre o qual efetuar uma comparação suficientemente sistemática entre os dois países, em matéria de educação. Da análise das diferenças, emergem as seguintes conclusões: (a) em termos de variáveis de contexto, a Espanha está acima de Portugal; (b) em termos de *inputs*, Portugal supera a Espanha na despesa acumulada por aluno e fica atrás na oferta de recursos educativos e humanos; (c) em relação a *outputs* e a *outcomes*, Portugal supera a Espanha em todos os indicadores considerados; (d) finalmente, em relação às políticas, a posição de Portugal está claramente à frente da Espanha. Estas evidências traduzem-se num conjunto de recomendações, inspiradas em consensos internacionais e validadas pelo sucesso do caso português.

Palavras-chave: educação comparada; abordagem sistêmica; educação espanhola; educação portuguesa; avaliação.

Spain vs. Portugal in education. A systemic approach

Abstract. As a result of the demands of the context and its rapid transformation, modern societies have been confronted with the urgent need to improve their educational systems. A useful practice for the improvement of organizations is to compare oneself with others that, starting from inferior situations, have managed to advance in a substantive way; to later try to find out how they did it. This paper takes up this strategy and adopts Portugal as an element of comparison. To do this, it starts from a systemic approach to education based on the structure: context, *inputs*, policies, *outputs*, and *outcomes*, which serves as a conceptual framework for selecting a broad set of indicators on which to make a sufficiently systematic comparison between the two countries, in the field of education. The following conclusions emerge from the analysis of the differences: (a) regarding context variables, Spain is above Portugal; (b) in terms of *inputs*, Portugal exceeds Spain in accumulated expenditure per student and is behind in providing educational and human resources; (c) in terms of *outputs* and *outcomes*, Portugal exceeds Spain in all the indicators considered; (d)

finally, in relation to policies, the position of Portugal is clearly more advanced than that of Spain. This evidence is translated into a handful of recommendations, inspired by international consensus and validated by the success of the Portuguese case.

Keywords: comparative education; systemic approach; spanish education; portuguese education. evaluation.

1. Introducción

En el contexto de la llamada “cuarta revolución industrial” (Shwab, 2016) –una revolución tecnológica basada en sistemas físicos, sistemas biológicos, sistemas digitales y en sus interacciones– las sociedades modernas se hallan confrontadas a la necesidad imperiosa de mejorar sus sistemas educativos (Schleicher, 2018). Este consenso internacional emergente se ha visto reforzado por los efectos deletéreos de la pandemia COVID-19 (Reimers y Schleicher, 2020).

Aun cuando una práctica habitual en la gestión para la mejora de las organizaciones consiste en compararse con aquellas otras situadas en posiciones superiores (López Rupérez, 1994), es posible, asimismo, modificar la orientación y compararse con las que, partiendo de situaciones inferiores, han conseguido avanzar de un modo sustantivo, para, posteriormente, tratar de averiguar cómo lo hicieron.

Geográfica, cultural y lingüísticamente próximo, Portugal es un país hermano de un tamaño intermedio que ha padecido como España los efectos de una dictadura militar; ha arrastrado un retraso histórico superior; dispone de un inferior nivel de riqueza medida por el PIB *per cápita*, y un índice de nivel socioeconómico y cultural más bajo. Y, a pesar de todas esas desventajas comparativas, Portugal ha conseguido evolucionar a mejor a lo largo de las últimas décadas y obtiene resultados superiores a los de España en materia educativa, medidos por diferentes indicadores que son considerados clave en el panorama internacional.

Por tal motivo, procede efectuar una comparación sistemática entre los dos países en el ámbito de la educación escolar, intentar evidenciar algunas diferencias y apuntar a sus muy probables factores explicativos. Además, y bajo este objetivo, en el presente trabajo late la pretensión de arrojar alguna luz sobre la naturaleza de las reformas que debería emprender España en la próxima década en materia educativa y sus rasgos ineludibles, tomando como inspiración el caso portugués.

De acuerdo con este propósito, la metodología de análisis de este trabajo se apoya en una aproximación sistémica a la educación que servirá como marco conceptual para seleccionar un conjunto relativamente amplio de indicadores sobre el que efectuar una comparación sistemática entre los dos países, en materia educativa. Esta comparación reposará, principalmente, en la consideración de series cronológicas. De los resultados obtenidos y del estudio de las diferencias emerge un conjunto de conclusiones que se traducirá en un puñado de recomendaciones inspiradas en el consenso internacional, y validadas por el éxito del caso portugués.

2. Una aproximación sistémica a la Educación

La aproximación sistémica a la Educación es deudora del enfoque adoptado por el mundo de los indicadores internacionales de educación que sirvió como elemento estructurante del proyecto INES (*Indicators of Education Systems*) de la OCDE (CERI-OECD, 1992; 1994). La publicación anual *Education at a Glance. OECD Indicators* se articuló, en buena parte de sus ediciones, sobre la base de dicha aproximación sistémica (OECD, 2012).

En el presente trabajo, partiremos de la configuración básica de ese enfoque sistémico y lo ampliaremos, profundizando en sus relaciones internas y aproximándonos a su nivel de complejidad.

195

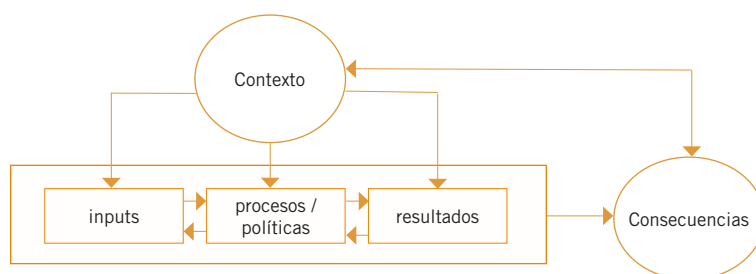
El sistema educativo propiamente dicho puede modelizarse como un conjunto de *inputs*, entrantes o insumos que, mediante los procesos y las políticas que se desarrollan en su seno, generan un conjunto de *outputs* o resultados propiamente educativos. Este esquema simple se complica (véase la figura 1) al considerar que no sólo la naturaleza y magnitud de los *inputs* permiten el desarrollo de los procesos y las políticas y, con toda probabilidad, los condicionan, sino que, inversamente, las políticas operan sobre los *inputs* en tanto que actuaciones de los gobiernos que los proporcionan y administran. En un sentido análogo, la aplicación de políticas sobre el sistema producirá resultados educativos; pero, en cualquier sistema moderno e inteligente, los resultados actuarán retroactivamente sobre las políticas –a partir de los mecanismos de *feedback* formativo– permitiendo reforzarlas o corregirlas (López Rupérez *et al.*, 2019).

Esta complejidad se incrementa cuando se consideran dos elementos externos al sistema propiamente dicho, pero en interacción con él y entre sí. Por un lado, está el entorno o contexto socioeconómico y cultural en el que opera el sistema educativo y que alude a un sistema más amplio; y, por otro, los *outcomes*, consecuencias o impacto que tiene la educación

más allá del sistema educativo. El contexto condiciona la magnitud de los *inputs*, incide sobre las políticas y afecta a los resultados, por una vía tanto directa –por efecto de factores diversos propios de ese sistema más amplio del que la educación es tan solo un subsistema–, como indirecta, mediante su incidencia sobre la definición e implementación de las políticas educativas que operan sobre los resultados.

Por otra parte, las consecuencias de los *outputs* del sistema educativo derivan obviamente –al menos en parte– de dichos resultados, aunque esta vinculación entre resultados y consecuencias esté modulada por la influencia del contexto. Pero, además, esos cambios producidos en variables sociales y económicas reflejan, a la postre, cambios del contexto; de modo que la educación influye sobre el contexto y es influido por él, de acuerdo con un bucle causal que se extiende en el medio y largo plazo, y que definirá un horizonte de progreso si el sistema tiene éxito.

Figura 1. Modelo conceptual sistémico de la Educación ampliado.



Fuente: Elaboración propia (López Rupérez y García García, 2020).

Disponemos, por tanto, de un modelo conceptual integrado, como el que se representa en la figura 1, que nos servirá de referente para articular el presente estudio comparado.

3. Una comparación sistemática entre España y Portugal en materia educativa

Por su potencia, a la hora de realizar una descripción completa de los sistemas educativos, el anterior modelo conceptual resulta especialmente apropiado para una comparación sistemática entre España y Portugal en materia educativa. En lo que sigue, efectuaremos la comparación sobre la base de esos cinco elementos que aparecen representados en la figura 1: Contexto, *inputs*, *outputs*, *outcomes* y políticas.

3.1 Desde la perspectiva del contexto

3.1.1 Nivel socioeconómico y cultural

Una de las variables contextuales ampliamente reconocidas, por su impacto sobre los resultados de los alumnos, es el nivel socioeconómico y cultural. Desde su primera edición, PISA ha tomado en consideración el índice correspondiente (ESCS: *Economic, Social and Cultural Status*) y lo ha medido de un modo sistemático a partir del cuestionario dirigido a los alumnos (OECD, 2019). La tabla 1 –o cuadro de yuxtaposición– muestra los resultados comparados España/Portugal en relación con dicho índice (ISEC en español). El ISEC (Índice Socio Económico y Cultural) es un indicador, normalizado tomando como base los países miembros de la OCDE, con un valor medio igual a 0 y una desviación estándar igual a 1, de modo que valores positivos de dicho índice significan que el país se sitúa por encima de la media, y valores negativos significan que se sitúa por debajo de ella.

3.1.2 Nivel de riqueza

El nivel de riqueza de un país es otra de las variables contextuales clásicas que se mide habitualmente mediante su PIB *per cápita*. Al tratarse de un indicador que toma en consideración la población, resulta mucho más preciso, a efectos de comparación, que el PIB total. La tabla 1 recoge los resultados para España y Portugal de dicho indicador de nivel de riqueza.

Con relación tanto a un indicador contextual, como al otro España se encuentra, con respecto a Portugal, en una posición más favorable, lo que refleja un mayor nivel de desarrollo.

3.2 Desde la perspectiva de los inputs

3.2.1 Gasto acumulado por alumno

El gasto acumulado por alumno en establecimientos escolares en las edades comprendidas entre los 6 y los 15 años es uno de los indicadores empleados por PISA para cuantificar el nivel de gasto educativo en términos comparables, cuando se corrige del efecto de las diferencias entre países en cuanto a sus respectivos poderes adquisitivos (paridad de poder adquisitivo, *ppa*). En la tabla 1 se muestran los valores correspondientes al ejercicio de 2015, que es el último consolidado y disponible para la comparación internacional (OCDE, 2019). Se advierte en ella que, a pesar de su inferior nivel de riqueza, Portugal gasta bastante más por alumno que España.

3.2.2 Dotación en personal educativo

A partir de las respuestas al cuestionario dirigido a directores escolares, la edición 2018 de PISA ha introducido un “índice de escasez de personal educativo” (OECD, 2019) –también normalizado sobre la base de los países miembros de la OCDE– con un valor medio igual a 0 y una desviación estándar igual a 1. La tabla 1 recoge los valores correspondientes a este indicador para España y para Portugal.

3.2.3 Dotación en recursos educativos

Un tratamiento similar al anterior se ha efectuado en PISA 2018 para los recursos educativos (OECD, 2019) obteniéndose para España y para Portugal los valores que se muestran en la tabla 1. Tanto para el indicador de dotación de personal, como para el de recursos educativos, los datos revelan índices de escasez superiores a la media de la OCDE. En ambos casos, el grado de escasez es mayor en Portugal que en España, según la percepción de los directores de los centros de educación secundaria encuestados.

Tabla 1. Indicadores de contexto y de *inputs* considerados en el estudio

Indicador	España	Portugal	OCDE
- Nivel socioeconómico y cultural (ISEC)	-0,1	-0,4	0,0
- Nivel de riqueza (PIB pc en \$ USA ppa)	33.629	28.760	39.333
- Gasto acumulado/alumno (\$ USA ppa)	74.947	83.050	90.294
- Escasez en personal educativo	0,33	0,82	0,0
- Escasez en recursos educativos	0,22	0,48	0,0

Nota: Los datos de nivel de riqueza y de gasto acumulado por alumnos (de 6 a 15 años) corresponden al año 2015 que es el último reportado por la OCDE.

Fuente: Elaboración propia sobre la base de datos recopilados por la OCDE (OECD, 2016; 2019)

3.3 Desde la perspectiva de los outputs

En un sentido restringido –pero más fácil de medir– los *outputs* de un sistema educativo son sus logros académicos, que pueden ser considerados según diferentes perspectivas: se pueden abordar desde un enfoque más global o más analítico; basados en pruebas de rendimiento o en estadísticas administrativas; deteniéndonos en la educación o pensando en sus consecuencias; mirando al presente o preparando el futuro; analizando la evolución del propio sistema o comparándonos con otros. Todas estas perspectivas, no excluyentes entre sí, serán consideradas en lo que sigue en nuestro particular ejercicio de comparación internacional.

Por otro lado, es sabido que el impacto de las políticas educativas sobre los resultados del sistema, en su conjunto, se deja sentir en el medio y largo plazo; de ahí que, a efectos de comparación –sea de un país consigo mismo, sea con los demás–, resulte más seguro analizar series históricas, cuando menos a lo largo de una década. Este es el enfoque adoptado en el presente estudio para abordar la comparación España con Portugal en lo concerniente a los *outputs* de sus respectivos sistemas educativos.

3.3.1 Rendimiento en PISA

La tabla 2 –o cuadro de yuxtaposición– resume la comparación entre España y Portugal que permite la síntesis diagnóstica efectuada por PISA sobre la base de su historial –de cerca de dos décadas– relativo a los resultados de sus evaluaciones a gran escala (OECD, 2019). Del análisis de los datos cabe inferir las siguientes evidencias:

Tabla 2. Resumen comparado de las tendencias de rendimiento en PISA en España y en Portugal

Tendencias	Lectura		Matemáticas		Ciencias	
	España	Portugal	España	Portugal	España	Portugal
Trayectoria de rendimiento general	—	Constantemente	Plana	Positiva, pero menos en los últimos años	Más negativa en los últimos años	Positiva, pero menos en los últimos años
Tendencia promedio entre los estudiantes con mejor rendimiento (percentil 90)	1,2	4,1	-0,1	7,8	-0,2	6,0
Tendencia promedio entre los estudiantes de bajo rendimiento (percentil 10)	-3,2	5,2	0,3	2,6	0,4	1,7
Brecha en los resultados de aprendizaje entre los estudiantes con mayor y menor rendimiento	—	Estable	Estable	Brecha cada vez mayor	Estable	Estable

Fuente: elaboración propia a partir de *PISA 2018 Results* (Volume I): What Students Know and Can Do» © OECD 2019, Anexo D, p. 327.

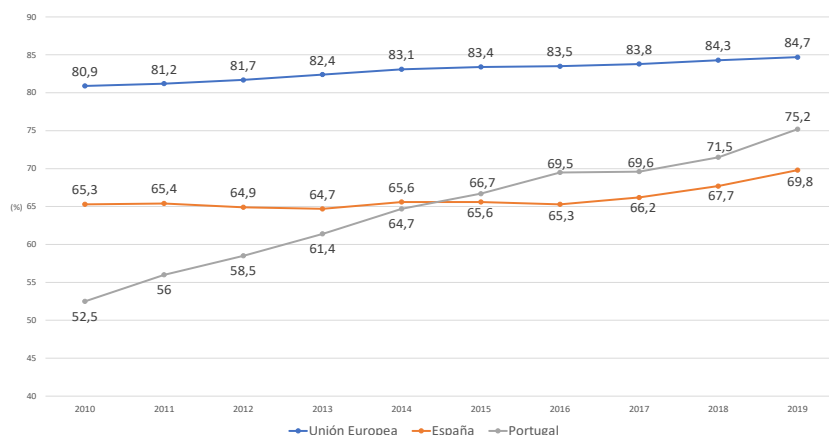
- a) España presenta una tendencia plana o ligeramente negativa en el periodo considerado, mientras que Portugal se ha comportado de forma consistentemente positiva en las tres áreas consideradas: Lectura, Matemáticas y Ciencias.

- b) En materia de excelencia (percentil 90), España está estancada mientras que Portugal ha mejorado significativamente en las tres áreas, particularmente en Matemáticas y en Ciencias.
- c) En materia de equidad (percentil 10), España apenas si ha variado en Matemáticas y en Ciencias y ha empeorado en Lectura, mientras que Portugal ha mejorado en las tres áreas.
- d) El comportamiento en Matemáticas se traduce en un cierto empeoramiento de Portugal en cuanto a la brecha existente entre los extremos, toda vez que la excelencia ha progresado más rápidamente que la equidad, lo que no sucede en España debido al carácter plano de su evolución.

3.3.2 Tasa de la población joven que ha completado la educación secundaria superior

La tasa de la población adulta joven (25-34 años) que ha completado, al menos, la educación secundaria superior es un indicador internacional de resultados que apunta al grado de acierto de un sistema educativo en la preparación del futuro, al referirse a un nivel formativo considerado imprescindible. La figura 2 fija la atención en la comparación de tendencias entre España y Portugal, tomando en ambos casos como referencia el comportamiento de la Unión Europea en su conjunto.

Figura 2. Evolución del porcentaje de jóvenes de 25 a 34 años que han completado al menos, el nivel educativo de la educación secundaria superior (CINE 3-8) en España, Portugal y Unión Europea. Años 2010 a 2019



Fuente: Elaboración propia a partir de Eurostat [edat_ifse_03].

Del análisis de los datos que se representan en dicha figura 2 cabe destacar las siguientes evidencias:

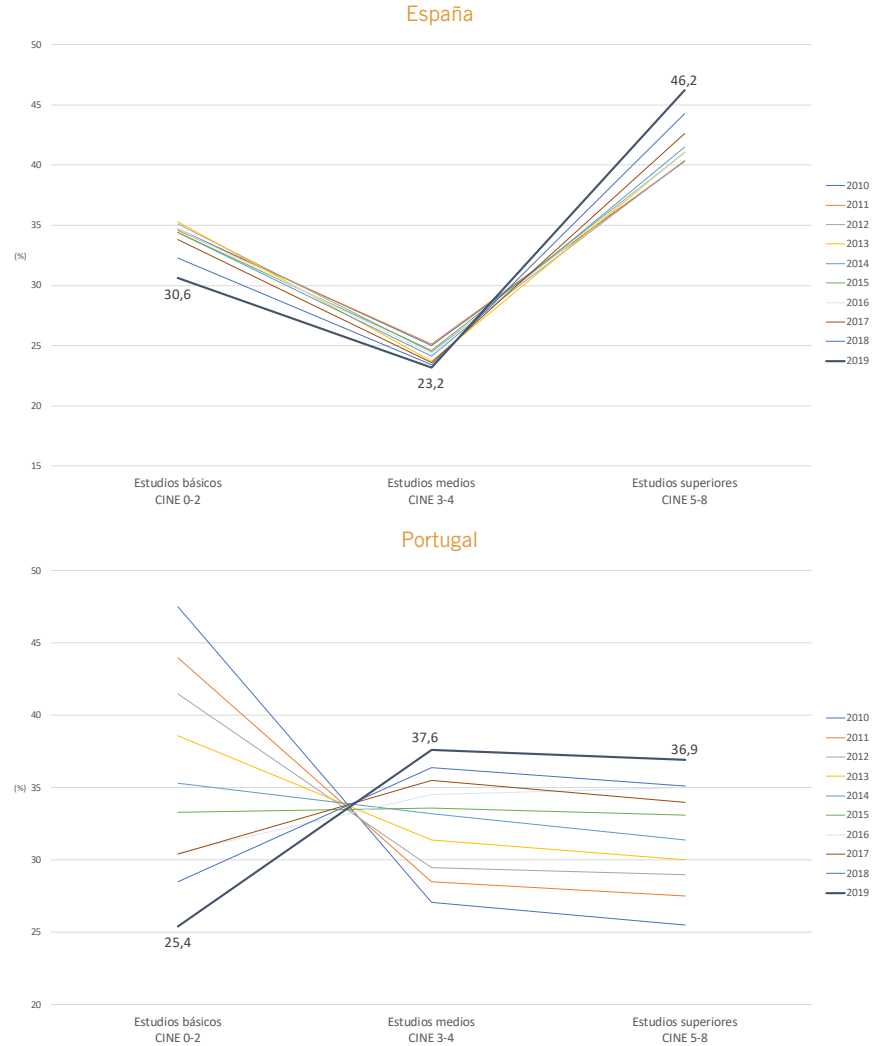
- a) Portugal, que partía en 2010 de una posición retrasada 12,8 puntos porcentuales con respecto a España, le superaba en 5,2 puntos porcentuales en 2019.
- b) Portugal, en este mismo periodo, ha reducido la brecha de los 28,4 puntos porcentuales que le separaban de la UE, a 10,1 puntos, habiendo logrado, por tanto, una convergencia de 18,3 puntos de magnitud. Mientras, España lo hacía de 15,6 a 15,5 puntos porcentuales, es decir, con una convergencia de 0,1 puntos, que no es significativa.

3.3.3 Brecha formativa de la población joven

Las figuras 3 y 4 muestran la evolución de los respectivos patrones de distribución de la población adulta joven (25-34 años) por niveles de formación, correspondientes a España y a Portugal. De su análisis cabe destacar los siguientes hechos:

- a) España ha corregido en parte esa brecha, pero a un ritmo lento e incapaz de alterar en una década el patrón de distribución —en forma de V— de la población joven por niveles de formación: básico, medio y superior.
- b) Portugal ha hecho avances mayúsculos en el mismo periodo de tiempo, reduciendo notablemente el porcentaje de la población joven con estudios básicos, y elevando a la par los porcentajes correspondientes a estudios medios y a estudios superiores. Esa mayor rapidez en los procesos de mejora se evidencia por la simple comparación gráfica de las anchuras de las correspondientes bandas de variación.
- c) Los progresos efectuados por Portugal, a lo largo del periodo de referencia, en materia de tasa de la población adulta joven (25-34 años) que ha conseguido completar, al menos, los estudios correspondientes a la educación secundaria superior, se ha traducido en la corrección del patrón de distribución de esa misma población por niveles de formación, que al principio de dicho periodo se parecía al de España y ahora se parece mucho más al de la Unión Europea, con forma de V invertida.
- d) El notable incremento producido en Portugal, en cuanto a porcentaje de población joven con un nivel intermedio de formación, le ha llevado a alcanzar los 37,6 puntos porcentuales frente a 23,2 de España.

Figuras 3 y 4. Evolución del nivel de educación de la población comprendida en el intervalo de edad de 25 a 34 años en España y en Portugal, respectivamente. Años 2010-2019



Fuente: Elaboración propia a partir de Eurostat [edat_lfse_03].

3.4 Desde la perspectiva de los outcomes

De conformidad con la orientación típica de los sistemas de indicadores internacionales de educación (CERI-OCDE, 1992; 1994), los *outcomes* de un sistema educativo equivalen al impacto de dicho sistema que se proyecta en los ámbitos social y económico (CERI-OCDE, 2007), más allá de su marco concreto de definición. Aun cuando la relación entre *outputs* y *outcomes* está

modulada por un conjunto diverso de factores intervinientes, de conformidad con la evidencia empírica disponible la intensidad de la asociación entre ambos es las más de las veces, aunque moderada, estadísticamente significativa (Xia y Kirby, 2009). Por tal motivo, procede considerar los *outcomes* como un producto, siquiera sea parcial, de los sistemas educativos.

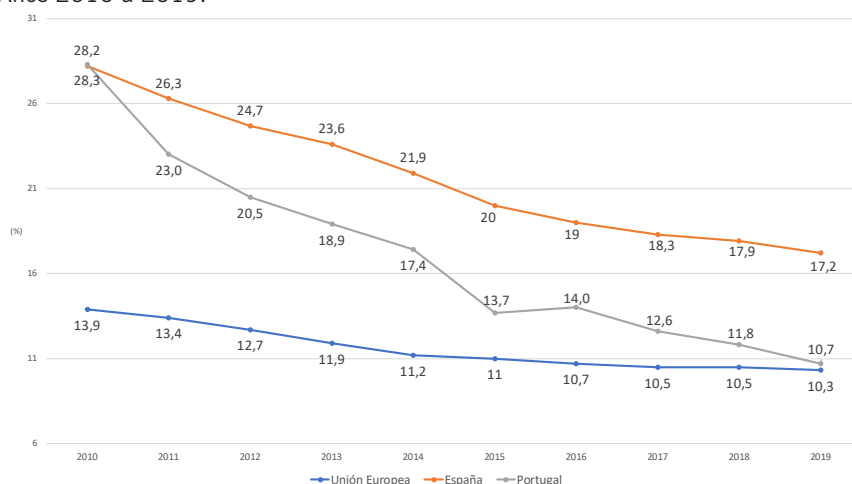
3.4.1 Abandono educativo temprano

Este indicador, que forma parte destacada del Monitor de Educación y Formación –herramienta empleada por la Comisión Europea para el seguimiento de los países miembros en materia educativa (European Commission, 2019)–, apunta directamente, por su propia definición, al ámbito de los *outcomes*. De hecho, antes de que pasara a formar parte del conjunto de indicadores para el seguimiento de la Estrategia de Lisboa en materia educativa, el abandono educativo temprano era un indicador europeo de cohesión social (SCO51).

En materia de abandono educativo temprano, la tendencia manifestada en los indicadores de *outputs* se hace, si cabe, aún más clara. La conclusión que se deriva de un modo evidente de los datos representados en la figura 5 es la siguiente: partiendo en 2009 de valores similares a los de España, Portugal ha logrado en el periodo de referencia converger prácticamente con la UE, mientras que España mantiene todavía una brecha de convergencia con la Unión de 6,9 puntos porcentuales y tendrá dificultades para alcanzar en 2020 el objetivo adaptado del 15 por ciento.

203

Figura 5. Evolución del abandono educativo en España, Portugal y Unión Europea. Años 2010 a 2019.

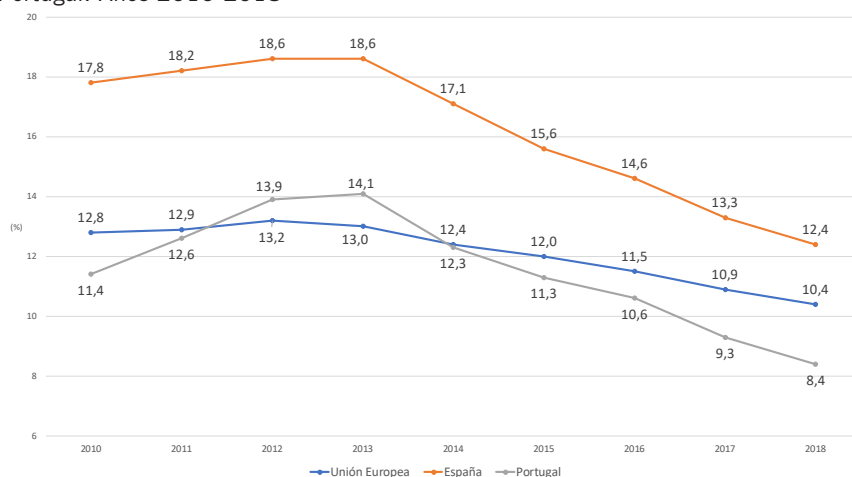


Fuente: Elaboración propia a partir de Eurostat [SDG_04_10].

3.4.2 Tasa de NiNis

Uno de los indicadores de la Unión Europea que integra, a la vez, situaciones de no-formación y de no-empleo en los jóvenes es el porcentaje de las personas de entre 18 y 24 años que ni estudia ni trabaja. Conocidos como NiNis, estos jóvenes son firmes candidatos a quedarse al margen del camino en su trayectoria vital lo que, al integrar sus efectos a nivel global, se convierte en una seria amenaza, tanto para la cohesión social como para el desarrollo social.

Figura 6. Evolución del porcentaje de la población joven con edades comprendidas entre 15 y 24 años que ni estudia ni trabaja en la Unión Europea, en España y en Portugal. Años 2010-2018



Fuente: Elaboración propia a partir de Eurostat [lfsi_neet_a].

El análisis de los datos correspondiente a la evolución de la tasa de NiNis para España y Portugal –y para la UE como referencia de fondo– se representan gráficamente en la figura 6. De la consideración de esas tres series cronológicas se infieren las siguientes evidencias:

- a) España se sitúa, en cualquier punto de la serie, sistemáticamente por encima de Portugal como consecuencia, muy probablemente, de la desventaja de España en materia tanto de abandono educativo temprano como de desempleo juvenil, así como debido a su interacción, por efecto de esa regla empírica que establece “a menor formación, mayor desempleo”.

- b) En el periodo central de la crisis de 2008 el ritmo de crecimiento de la tasa de NiNis fue muy superior en Portugal que en España; pero a la salida de aquélla Portugal había sido capaz de mejorar notablemente su situación hasta colocarse, frente al fenómeno NiNi, en una posición más ventajosa que la de la propia UE.

3.4.3 Riesgo de pobreza o exclusión social

De conformidad con la Estrategia Europa 2020, se consideran personas en riesgo de pobreza y/o exclusión social a la población que se encuentra en cualquiera de las situaciones siguientes: (a) Personas que viven con bajos ingresos (60% de la mediana del ingreso equivalente o por unidad de consumo en el año anterior a la entrevista); (b) personas que sufren privación material severa (4 de los 9 ítems definidos); (c) personas que viven en hogares con una intensidad de empleo muy baja (por debajo del 20% del total de su potencial de trabajo en el año anterior a la entrevista). En el caso de que las personas estén incluidas en más de una de estas tres condiciones, se contabilizan solo una vez (Eurostat, 2020).

El análisis de la evolución de este fenómeno por niveles de formación puede arrojar alguna luz sobre el comportamiento diferenciado entre España y Portugal a este respecto y sobre algunas de sus claves explicativas. Las figuras 7 y 8 permiten esa comparación más compleja. Un resultado surge con nitidez de su análisis y es que, aun cuando el riesgo de pobreza o exclusión social de la población joven (18-24 años) afecta en mayor medida a España, sea cual fuere el nivel de formación considerado, la diferencia más clara, tanto en magnitud como en tendencia, corresponde al sector de más bajo nivel de formación. Así, se advierte con toda claridad que conforme en Portugal la tasa correspondiente empezó a reducirse desde un valor de 51,9 puntos porcentuales a partir de 2013, en España sucedió dos años después y desde un valor de 60,9 puntos.

A la vista de lo anterior, parece como si, en comparación, el riesgo de pobreza o exclusión social se cebara especialmente en España con los sectores de más bajo nivel de formación; lo que, junto con el poder predictivo del abandono educativo temprano en relación con dicho indicador (López Rupérez y García García, 2020) parece sugerir –con todas las limitaciones que para este tipo de evidencias plantea la atribución causal– el promover, con cierta urgencia, un incremento del nivel de formación de esos jóvenes situados en el nivel inferior y una mejora de sus cualificaciones.

Figuras 7 y 8. Evolución del riesgo de pobreza o exclusión social de la población joven con edades comprendidas entre 18 y 24 años, según el nivel educativo, en España y Portugal, respectivamente. Años 2008-2018



Fuente: Elaboración propia a partir de Eurostat [ilc_peps04].

3.5 Desde la perspectiva de las políticas

Aun cuando la influencia de las políticas sobre los resultados de los alumnos está mediada por un conjunto amplio de niveles y de factores intermedios, su *intensidad* –o fuerza de la relación entre políticas y *outputs*– ha de considerarse complementada por su *extensión*, que se refiere a su alcance,

medido por la amplitud de la población a la que llega dicha influencia. Así que la relevancia de un factor, o de una política educativa que lo potencie, dependerá del producto *intensidad x extensión* (López Rupérez, 2020).

Por tal motivo, procede profundizar en los análisis de las políticas porque, con una muy elevada probabilidad, se encuentren en ellas algunas de las claves de las diferencias antes descritas en materia de resultados.

3.5.1 Algunas características de las reformas educativas portuguesas

Un análisis del caso portugués, apoyado en distintas fuentes (Crato, 2020; OECD, 2018), permite concluir que los sucesivos gobiernos, de signos políticos diferentes, han prestado a la educación un interés especial, y han introducido reformas bien alineadas con objetivos de mejora situados en el medio y largo plazo; mejoras que se han puesto de manifiesto, con bastante claridad, en el anterior análisis empírico de las series cronológicas.

Una de las reformas extensas que el Gobierno portugués emprendió al final de la primera década del presente siglo, y que en este punto conviene traer a colación, fue la extensión de la enseñanza obligatoria hasta los 18 años. El 27 de agosto de 2009, la Asamblea de la República aprobó la Ley nº 85/2009 que amplió la obligatoriedad y gratuidad de la enseñanza a un periodo de 12 años (6-18 años). Con ello, se modernizaba la reforma que en 1986 extendió la etapa obligatoria hasta los 16 años pero que tardó un decenio en ser implantada. Por su carácter masivo, por su impulso equalizador de los territorios y por su naturaleza de política de Estado, es altamente probable que algunos de los resultados observados en cuanto al incremento del nivel formativo de los jóvenes –particularmente la reducción del abandono educativo temprano y la reducción de la brecha formativa de la población joven– estén vinculados a esa valiente decisión política, que hubo de implementarse en un contexto de reducción drástica del gasto público por efecto de la intervención de la economía portuguesa derivada de la gran crisis de 2008. Con esta reforma –alineada con una recomendación genérica del Parlamento europeo del 1 de diciembre de 2011 (Parlamento Europeo, 2011)–, Portugal se incorporaba al grupo relativamente numeroso de países y de regiones de la Unión Europea que han adoptado una decisión semejante (Eurydice, 2015).

Esta iniciativa se suma a un conjunto amplio de reformas continuadas, acumulativas y coherentes, iniciadas con el comienzo del siglo XXI, cuyos principios mayores y sus concreciones pueden sintetizarse –a partir del análisis histórico efectuado por Crato (2020)– en los términos siguientes:

a) Prestar una atención prioritaria a los resultados.

- Difusión pública de los resultados de cada centro en las pruebas nacionales de final de la educación secundaria.
- Atención especial a los resultados en la orientación de los aprendizajes básicos.
- Definición de un sistema de estándares que especifiquen lo que los alumnos tienen que alcanzar.
- Establecimiento de un sistema integrado de evaluaciones externas e internas para la mejora.
- Fortalecimiento de la autorregulación del sistema mediante las evaluaciones externas.
- Creación de un Instituto de Evaluación Educativa como organismo autónomo e independiente.

b) Conciliar calidad con cantidad.

- Extensión de la enseñanza obligatoria hasta los 18 años.
- Extensión de la enseñanza del inglés.
- Incremento del gasto educativo hasta el 5,1% de su PIB, recuperando el nivel anterior a la crisis.

c) Conciliar excelencia con equidad.

- Asunción colectiva de la idea de que la escuela puede hacerlo mejor.
- Definición de un currículo bien estructurado y exigente, y el impulso de una educación ambiciosa para todos.
- Desarrollo de un plan para reducir el abandono educativo y apoyar a los alumnos con dificultades académicas y de entornos desaventajados.
- Apoyo en una educación seria y una evaluación fiable, de acuerdo con estándares nacionales.
- Implementación de una concepción flexible de la formación profesional, asociada a las empresas.

d) Hacer del conocimiento lo primero.

- Considerando que el conocimiento es importante.
- Asumiendo que, sin una base robusta de conocimiento, los alumnos no podrán desarrollar competencias avanzadas.

- Reconociendo que la amplitud de contenidos no puede lograrse a expensas de una renuncia a un conocimiento estructurado y profundo.

e) Priorizar los aprendizajes básicos.

- Establecimiento de un plan especial para la mejora en Matemáticas.
- Establecimiento de un plan especial para la mejora en Lectura.
- Definición de cuatro pilares formativos básicos: lectura y escritura fluidas; matemáticas, estadística y lógica; historia y geografía; y ciencias.
- Definición de dos áreas adicionales: artes en sus formas diversas; y literatura y lenguas.
- Orientación del currículo de un modo flexible y moderno.

f) Alinear los recursos con la reforma del currículo.

- Movilización de los recursos de conocimiento experto basados en los avances científicos al respecto.
- Establecimiento de una nueva normativa sobre libros de texto para convertirlos en uno de los pilares esenciales en los que apoyar la reforma curricular.

g) Reforzar la autonomía de los centros.

- Organización autónoma de las horas de dotación de profesorado asignado por la administración.
- Organización de las asignaturas de acuerdo con una lógica de curso o de ciclo.
- Asignación de sus recursos en la implementación de medidas para promover el éxito escolar.
- Ante la escasez de recursos durante la intervención de su economía, desarrollo de incentivos de orden moral, tales como premios o reconocimientos públicos.
- Reconocimiento de los centros que mejoren los resultados de sus alumnos, mediante la ampliación del número de horas de dotación de profesorado disponible para el centro.

h) Fortalecer la profesión docente.

- Introducción de un examen de entrada para los candidatos a profesor.
- Mejora de la formación inicial del profesorado, mediante la regulación del acceso y la reforma de los programas con el refuerzo de los conocimientos de las disciplinas básicas.

— Establecimiento de un máster de dos años con un amplio periodo de inducción.

- i) Adoptar un enfoque propio de un gradualismo reformador que, bajo gobiernos de signo diferente, otorga una continuidad constructiva a las reformas.

La OCDE (2018), a propósito de su análisis experto sobre la reforma que ha puesto en marcha Portugal para orientar sus currículos, de conformidad con el enfoque por competencias (Proyecto de Autonomía y Flexibilidad Curricular), efectúa la siguiente valoración general de la educación portuguesa:

Portugal ha adoptado un enfoque estratégico sólido para la reforma educativa. El país comenzó el proceso concibiendo los resultados que el sistema educativo debería buscar para sus alumnos, basándose en evidencias sobre las condiciones propias del contexto del siglo XXI. Formuló estos objetivos en un plan estratégico coherente, descrito con detalle en el documento de referencia, “Perfil de los estudiantes al finalizar la escolarización obligatoria”. Portugal ha logrado un acuerdo general sobre sus planes de reforma a través de consultas, debates y comunicaciones cuidadosas que han sido bien gestionadas y exitosas. Al buscar el asesoramiento de expertos, las aportaciones de los actores y la comunicación y el debate abiertos, el país ha invertido en la continuidad del plan de reforma por parte de los futuros gobiernos. El país parece estar abierto a la retroalimentación y al aprendizaje de las lecciones derivadas de las evidencias que han surgido sobre los éxitos y las debilidades del ‘Proyecto de Autonomía y Flexibilidad Curricular’ (OECD, 2018, p. 7).

Más allá de las decisiones puntuales –aunque no por ello menos relevantes– adoptadas por Portugal en materia de políticas educativas, hay un rasgo muy general que cabe inferir de lo anterior y que explicaría todo lo demás. Y es, pura y simplemente, la calidad de su gobernanza en educación. Gobernanza que contiene buena parte de los rasgos propios de un gobierno inteligente de los sistemas educativos (López Rupérez *et al.*, 2017).

3.5.2 Algunas características comparadas de las reformas educativas españolas

Con el propósito de facilitar la comparación, cabe tomar en consideración los 9 criterios descritos más arriba y que sintetizan los principios inspiradores de las reformas portuguesas desarrolladas a lo largo del presente siglo. No obstante, resulta difícil encontrar puntos de coincidencia. Así, por ejemplo, España no ha hecho de los resultados y de su mejora (a) una prioridad explícita de su sistema educativo, ni ha planteado en sus reformas –traducidas en las leyes orgánicas de 2006 (LOE) y 2013 (LOMCE)– esa ambición por el conocimiento (d) que alcance a todos, sea cual fuere su nivel socioeconómi-

co y cultural (c); ni tampoco ha hecho suya la recomendación de 2015 del Consejo Escolar del Estado de extender la “enseñanza básica” (obligatoria y gratuita) hasta los 18 años (b) (CEE, 2015; López Rupérez *et al.*, 2015).

Si se profundiza en las reformas propiamente curriculares (e y f), España se ha alejado de esas claves del éxito portugués que ha diagnosticado la OCDE (2018). Así, y como hemos destacado y justificado en otro lugar, al analizar la reforma curricular de la LOMCE y su enfoque por competencias (López Rupérez, 2020):

La contrastación del caso de España con las mejores prácticas internacionales ha permitido identificar cuatro tipos de errores:

- a) La apuesta por una opción de máximos y no gradualista del Marco Europeo de Referencia (todas las competencias, todas las materias, todas las etapas y al mismo tiempo), reflejo probablemente de una ambición política por alinearse con las orientaciones de la Unión Europea que ha resultado ser a la postre excesiva, simple y poco crítica.
- b) Una confianza excesiva, de inspiración burocrática, en el poder transformador de las normas, al considerar la publicación de los currículos en el Boletín Oficial del Estado como la culminación de la labor del Gobierno.
- c) La debilidad de un planteamiento estratégico que no contempló una aproximación sistémica o integrada de diferentes políticas facilitadoras del éxito en la implementación de la reforma curricular; ni se articuló ningún plan estratégico, siquiera fuere parcial, que las incorporara con realismo; ni se implicó suficientemente al profesorado, ni a los directores en el proceso.
- d) Un insuficiente liderazgo institucional, escasamente basado en las personas y en el conocimiento e incapaz de pilotar —desde la *auctoritas* y no sólo desde la *potestas*— los procesos de reforma del sistema en su conjunto, justamente en un difícil contexto político con riesgos permanentes de conflicto entre administraciones gobernadas por partidos de distinto signo, que suelen emplear la confrontación con el Gobierno como una estrategia de desgaste político del adversario.” (López Rupérez, 2020, p.145)

Algunos tímidos avances normativos se pueden apreciar, de la mano de la LOMCE, en materia de autonomía de los centros (g), así como en los procesos de selección de los directores escolares en los centros públicos. Pero ningún progreso normativo se ha producido en materia de fortalecimiento de la profesión docente (h), más allá de los debates —prolongados en el tiempo, a lo largo de la última década— sobre el llamado “MIR educativo” (López Rupérez, 2014; 2015).

En materia de calidad de la gobernanza del sistema educativo español, una evaluación de carácter exploratorio apoyada en un estudio Delphi (López Rupérez *et al.*, 2017; 2020) –sobre un panel de 21 expertos y tres rondas sucesivas de consultas– ha arrojado los siguientes resultados:

Para ninguna de las seis categorías consideradas en el modelo la puntuación total alcanza un valor igual o superior a 2,0 puntos –que corresponde a la calificación de mediano– lo que revela, a juicio de los expertos, una notable y consistente mediocridad de la gobernanza de nuestro sistema educativo. La dimensión n.º 4, «Gobernanza basada en el conocimiento, la evidencia empírica y la investigación», por su baja calidad, sobresale, de un modo consistente, sobre todas las demás. Su puntuación igual a 1,5 la sitúa en un nivel claramente mediocre, es decir, entre escaso y mediano”. (López Rupérez *et al.*, 2017, p. 102).

4. Conclusiones

Del análisis comparado entre los sistemas educativos de España y Portugal, efectuado a partir de un conjunto amplio de indicadores mayoritariamente cuantitativos, se derivan, en síntesis, las siguientes conclusiones empíricas:

- a) En relación con el *contexto*, España se sitúa consistentemente por encima de Portugal en factores tales como nivel de riqueza y estatus socioeconómico y cultural.
- b) En lo que concierne a los *inputs*, el gasto acumulado por alumno entre los 6 y los 15 años Portugal supera a España en un 11% aproximadamente, aunque la percepción de escasez de recursos –tanto educativos como humanos– por parte de los directores de los centros portugueses, es superior; probablemente como consecuencia de carencias procedentes del pasado y aún no cubiertas.
- c) En materia de *outputs*, en todos los indicadores de resultados educativos considerados –puntuaciones en PISA, tasa de la población adulta joven que ha completado al menos la educación secundaria superior, y brecha formativa de la población adulta joven– Portugal ha conseguido, a lo largo de la última década, superar a España.
- d) En lo relativo a los *outcomes* analizados –abandono educativo temprano, tasa de NiNis y riesgo de pobreza y exclusión social– Portugal se sitúa en posiciones sistemáticamente más favorables

que las de España y, en algún caso, que las de la media de los países de la UE.

- e) Finalmente, en lo que respecta a las *políticas* que transforman los *inputs* en *outputs*, la posición de Portugal está claramente más adelantada que la de España, según se infiere de un análisis comparado, de naturaleza cualitativa, basado tanto en descripciones de académicos portugueses, como en evaluaciones de la OCDE.

A la vista de los análisis más arriba descritos, cabe establecer como conclusión principal que España debería emprender reformas profundas y continuadas de su sistema educativo inspiradas en el consenso internacional emergente. Ello requeriría:

- a) Un planteamiento político y estratégico adecuado que garantice un programa consensuado de reformas, extendido sobre una década.
- b) Un apoyo claro de dicho programa en el conocimiento, la evidencia empírica y la investigación
- c) Una priorización de aquellas políticas que más impacto tienen sobre los resultados de los alumnos.
- d) Un enfoque integrado, entendido como aquél que toma en consideración esos factores críticos junto con sus interrelaciones.
- e) Una aportación de recursos adicionales al sistema con criterios de eficacia, eficiencia y equidad.
- f) Una implicación efectiva de los actores principales a fin de contribuir al éxito de las reformas.

Por su carácter general, estas recomendaciones son de aplicación en aquellos países de Iberoamérica que hayan de emprender procesos de reforma de sus sistemas educativos.

Procede en este punto final traer a colación la conclusión de Ben Levin (2010) –basada no sólo en la reflexión académica, sino también en su experiencia en la acción política– que resumía las lecciones de los últimos 50 años en materia de reformas educativas en los siguientes términos: “Si hay una cosa que hemos aprendido sobre política educativa es que ordenar a la gente que lo haga mejor no podrá tener éxito sin comprometer sus corazones y sus mentes” (Levin, 2010, p.742).

Referencias

- CEE (2015). *Informe 2015 sobre el estado del sistema educativo*. Consejo Escolar del Estado. Ministerio de Educación Cultura y Deporte. Madrid. Disponible en <https://bit.ly/347Gwwr>
- CERI-OECD (1992). *The OECD International Education Indicators. A Framework for analyses*. Paris: OECD Publishing.
- CERI-OECD (1994). *Évaluer l'enseignement. De l'utilité des indicateurs internationaux*. Paris: Éditions OECD.
- CERI-OECD (2007). *Comprendre l'impact social de l'éducation*. Paris: Éditions OECD.
- Crato, N. (2020). Curriculum and Educational Reforms in Portugal: An Analysis on Why and How Students' Knowledge and Skills Improved. *Audacious Education Purposes. How Governments Transform the Goals of Education Systems*. Reimers, F.M. (Ed). Chum, Switzerland: Springer. Disponible en <https://bit.ly/36frag5>
- European Commission (2019). *Education and Training Monitor 2019*. Disponible en <https://bit.ly/3jgr9Ms>
- Eurostat (2020). *People at risk of poverty or social exclusion by sex*. Disponible en <https://bit.ly/3mXOqVL>
- Eurydice (2015). *Compulsory Education in Europe, 2014/15*. Disponible en <https://bit.ly/2S87geH>
- Levin, B. (2010). Governments and education reform: some lessons from the last 50 years. *Journal of Educational Policy*, 25:6, 739-747. <https://doi.org/10.1080/02680939.2010.523793>
- López Rupérez, F. (1994). *La gestión de calidad en educación*. Madrid: La Muralla.
- López Rupérez, F. (2014). *Fortalecer la profesión docente. Un desafío crucial*. Madrid: Narcea Ediciones
- López Rupérez, F. (2015). "MIR Educativo" y profesión docente. Un enfoque integrado. *Revista Española de Pedagogía*, LXXIII(261), 283-299. Disponible en <https://bit.ly/30gOTbP>
- López Rupérez, F. (2020). *El currículo y la educación en el siglo XXI. La preparación del futuro y el enfoque por competencias*. Madrid: Narcea Ediciones.
- López Rupérez, F., García García, I. y Sanz Labrador, I. (2015). *La extensión de la enseñanza básica hasta los 18 años. Beneficios y costes*. Madrid: Fundación para los análisis y los estudios sociales. Disponible en <https://bit.ly/34bYfXPf>
- López Rupérez, F., García García, I. y Expósito Casas, E. (2017). *La calidad de la gobernanza del sistema educativos español. Un estudio empírico*. Madrid: Universidad Camilo José Cela. Disponible en <https://bit.ly/30ihP3l>
- López Rupérez, F., García García, I. y Expósito Casas, E. (2019). *Liderazgo de la dirección y feedback formativo. Dos pilares básicos de la gobernanza escolar*. Madrid: Universidad Camilo José Cela. Disponible en <https://bit.ly/33a6KTG>
- López Rupérez, F., García García, I. y Expósito Casas, E. (2020). Un marco analítico para la evaluación de la calidad de la gobernanza de los sistemas educativos. *Revista Iberoamericana de Educación*, 83(1), 53-76. <https://orcid.org/0000-0003-2613-9652>

- López Rupérez, F. y García García, I. (2020). *A vueltas con la equidad en educación. Una aproximación empírica en la perspectiva de las consecuencias*. Madrid: Universidad Camilo José Cela.
- OECD (2012). *Education at a Glance. OECD Indicators*. Paris: OECD Publishing.
- OCDE (2016). Résultats du PISA 2015 (Volume I): L'excellence et l'équité dans l'éducation. PISA. Paris: Éditions OCDE. <https://doi.org/10.1787/9789264267534-fr>
- OECD (2018). *Curriculum Flexibility and Autonomy in Portugal. An OECD review*. Disponible en <https://bit.ly/33cX2jr>
- OECD (2019). *PISA 2018 Results (Volume I): What Students Know and Can Do*. PISA. Paris: OECD Publishing. <https://doi.org/10.1787/5f07c754-en>
- Parlamento Europeo (2011). Abordar el abandono escolar prematuro. *Diario oficial de la Unión Europea. Resolución del Parlamento Europeo, de 1 de diciembre de 2011*. Disponible en <https://bit.ly/3mXRjWB>
- Reimers, F. & Schleicher, A. (2020). *A framework to guide an education response to the COVID-19 Pandemic of 2020*. Paris : OECD. Disponible en <https://bit.ly/3kT9efp>
- Schleicher, A. (2018). *Primera Clase. Cómo construir una escuela de calidad para el siglo XXI*. Madrid: Fundación Santillana.
- Shwab, K. (2016). *Cuatro principios de liderazgo de la cuarta revolución industrial*. Foro Económico Mundial. Disponible en <https://bit.ly/33anQ3S>
- Xia, N. & S.N. Kirby (2009). Retaining Students in Grade: A Literature Review of the Effects of Retention on Students' Academic and Nonacademic Outcomes. RAND Technical report. Disponible en <https://bit.ly/3cleShi>

Do julgamento professoral às avaliações externas, que fazemos com as crianças e jovens na escola? Um estudo de caso na cidade de Manaus-AM

Camila Ferreira da Silva ¹  <http://orcid.org/0000-0002-2348-9350>; Cecília Acácia da Silva Sampaio ¹  <http://orcid.org/0000-0002-3527-3207>; Thiago Felipe Paranaingá da Cunha ¹  <http://orcid.org/0000-0002-7795-6102>

¹ Universidade Federal do Amazonas (UFAM), Brasil

Resumo. O presente artigo ocupa-se das avaliações educacionais externas de larga escala e seus efeitos no cotidiano escolar. O cenário material no qual nosso estudo se situa diz respeito às transformações que marcaram as últimas três décadas no cenário educativo brasileiro responsáveis pela mudança na tônica da avaliação da aprendizagem no interior das escolas em direção à onipresença das avaliações externas – com especial atenção aos movimentos de constituição de um macrosistema nacional de avaliação da educação e da multiplicação de outros sistemas locais. Com o objetivo de refletir sobre as implicações das avaliações de larga escala para os educandos, ocupamo-nos dos sistemas avaliativos estadual e municipal do estado do Amazonas e de sua capital: trata-se do Sistema de Avaliação do Desempenho Educacional do Amazonas (SADEAM) e da Avaliação do Desempenho do Estudante (ADE). A partir de um estudo de caso de uma escola estadual localizada na Zona Oeste de Manaus, com base nos resultados alcançados no SADEAM (2011-2015), foi possível compreender a relação existente entre a realização das avaliações de larga escala e o processo de ensino-aprendizagem das crianças nas escolas manauaras.

Palavras-chave: avaliação educacional; avaliação de larga escala; SADEAM; ADE.

De la evaluación de los profesores a las evaluaciones externas, ¿qué hacemos con los niños y jóvenes en la escuela? Un estudio de caso en la ciudad de Manaus-AM

Resumen. Este artículo versa sobre las evaluaciones educativas externas a gran escala y sus efectos en la vida diaria escolar. El escenario material en el que se enmarca nuestro estudio se refiere a las transformaciones que han marcado las tres últimas décadas en el escenario educativo brasileño, responsables del cambio de énfasis de la evaluación de los aprendizajes en las propias escuelas hacia la omnipresencia de las evaluaciones externas, con especial atención a los movimientos de constitución de un macrosistema nacional de evaluación de la educación y la multiplicación de otros sistemas locales. Con el fin de reflexionar sobre las implicaciones de las evaluaciones a gran escala para los estudiantes, vamos a analizar los sistemas de evaluación estatales y municipales del estado de Amazonas y su capital: el Sistema de Evaluación del Desempeño Educativo de Amazonas (SADEAM) y el Sistema de Evaluación del Desempeño de los Estudiantes (ADE). A partir de un estudio de caso de una escuela estatal situada en la zona occidental de Manaus, basado en los resultados obtenidos en el SADEAM (2011-2015), pudimos comprender la relación entre las evaluaciones a gran escala y el proceso de enseñanza-aprendizaje de los niños en las escuelas de Manauara.

Palabras clave: evaluación educativa; evaluación a gran escala; SADEAM; ADE.

From teacher judgment to external assessments, what do we do with children and young people at school? A case study in the city of Manaus-AM

Abstract. This article deals with large-scale external educational assessments and their effects on school life. The material scenario in which our study is located concerns the transformations that marked the last three decades in the Brazilian educational scenario responsible for the change in the tonic of the evaluation of learning inside schools towards the ubiquity of external evaluations – with special attention to the movements of constitution of a national macrosystem for the evaluation of education and the multiplication of other local systems. In order to reflect on the implications of large-scale evaluations for students, we are concerned with the state and municipal evaluation systems of the state of Amazonas and its capital: this is the Educational Performance Evaluation System of Amazonas (SADEAM) and Student Performance Assessment (ADE). From a case study of a state school located in the West Zone of Manaus, based on the results achieved in SADEAM (2011-2015), it was possible to understand the relationship between the conduct of large-scale assessments and the teaching process and learning in schools in Manaus.

Keywords: education assessment; large scale assessment; SADEAM; ADE.

1. Introdução

A avaliação ocupa um lugar central nos processos educativos desde a concepção de escola moderna, faz parte do cotidiano das instituições educacionais e neste espaço expressa-se pela avaliação da aprendizagem historicamente realizada pelos professores, pelos conselhos de classe e pelas autoavaliações que passaram a ser promovidas pela gestão escolar em cenário recente. Os debates internacionais em torno dos sistemas educativos na transição das décadas de 1980 para 1990 fizeram emergir a questão da necessidade de avaliar tais sistemas no sentido de diagnosticá-los e, com isso, conseguir realizar um planejamento capaz de imprimir a noção de qualidade que se desenhava na literatura educacional especializada naquele momento. Destaca-se que esta noção de qualidade da educação estava aliada à concepção que alia educação formal ao desenvolvimento econômico e social dos países e que, portanto, toma a educação como um elemento diretamente proporcional aos processos de desenvolvimento ou subdesenvolvimento.

Este panorama de época desencadeou, para além das Conferências e Fóruns mundiais de Educação, a estruturação de políticas públicas nacionais especificamente voltadas para a avaliar a educação, bem como o desenvolvimento de sistemas de avaliação educacional próprios (Coelho, 2008). No caso do Brasil, data exatamente de 1990 o estabelecimento do Sistema Nacional de Avaliação da Educação Básica (SAEB), o qual, segundo Oliveira (2015), tem como principal função oferecer subsídios concretos para os processos de formulação, reformulação e monitoramento das políticas públicas para a Educação Básica no país. Os desdobramentos do SAEB na última década têm na construção de sistemas de avaliação educacional estaduais e municipais uma de suas expressões mais bem acabadas no sentido do alcance dos processos avaliativos de larga escala no país. Este é, pois, o cenário que serve de antessala para nosso artigo, posto que nos interessa compreender os contextos de emergência e as consequências do Sistema de Avaliação do Desempenho Educacional do Amazonas (SADEAM) e da Avaliação de Desempenho do Estudante (ADE), esta última sob gerência da Secretaria Municipal de Educação (SEMED) de Manaus.

Com o objetivo de refletir sobre as implicações das avaliações de larga escala para os educandos, este artigo se situa na Sociologia Política da Educação, e metodologicamente se caracteriza pela realização de um estudo de caso de uma escola estadual localizada na cidade de Manaus – Zona Oeste, com base em uma abordagem quanti-qualitativa dos resultados obtidos nas avaliações do SADEAM no período 2011-2015. Do ponto de vista da organização argumentativa, o texto encontra-se estruturado da seguinte maneira: em um primeiro momento tratamos de demarcar o descolamento histórico da avaliação da aprendizagem que ocorria no interior das escolas para a ava-

liação da larga escala que se dá de forma externa às instituições escolares; logo depois ocupamo-nos de situar o leitor no âmbito das especificidades dos sistemas próprios de avaliação do estado do Amazonas e de sua capital; e, por fim, a última seção do texto traz a análise da escola estadual sobre a qual se desenvolveu o estudo de caso, apresentando os resultados que esta instituição obteve nas últimas edições do SADEAM, bem como colocando a questão das consequências deste tipo de avaliação para os educandos.

2. Do julgamento professoral às avaliações externas: o que está em jogo?

A ampla literatura acadêmica sobre a avaliação da aprendizagem escolar aponta o papel determinante que o ato de avaliar possui para a mediação pedagógica e para o acompanhamento do desenvolvimento dos educandos no cotidiano da relação que se estabelece entre professores e estudantes. Esta mesma literatura, em complementaridade a este entendimento, problematiza os sentidos que a avaliação acaba por ganhar no cotidiano escolar, em função de uma usual reprodução de concepções e práticas avaliativas tradicionais que se baseiam na dominação, no autoritarismo e tomam a avaliação como sinônimo de “exame” (Freitas, 2003; Leite & Kager, 2009; Bourdieu & Saint-Martin, 2015; Hoffmann, 2017; Oliveira & Medeiros, 2018). Nesse sentido, esta bifurcação que se opera no âmbito de como se compreende e como se materializa a avaliação da aprendizagem na instituição escolar é reveladora das tensões que se colocam no debate sobre a avaliação educacional e escolar.

Luckesi (2000) assevera que o ato de avaliar, como todo e qualquer ato de conhecer, inicia-se pela constatação, que nos oferece a garantia de que o objeto é como é. A relação entre avaliação e a construção de um diagnóstico situacional ganha força nesta perspectiva e nos fala do significado de qualidade educacional para as escolas, uma qualidade que figura no encaixo da aprendizagem e do desenvolvimento dos educandos. Contudo, as tensões entre as diferentes significações dos processos avaliativos em contexto escolar e educacional estão na base de duas questões que se correlacionam mutuamente e que nos servem aqui e ponte para pensar as correlações entre avaliações internas e externas à escola, nomeadamente: 1) a perspectiva de avaliação como “exame”, que ganhou tanto lastro na sociedade que nos permite afirmar que, das crianças aos adultos, se perguntarmos sobre o que é a avaliação no contexto da escola obteremos respostas que tratarão de elementos como “provas”, “testes”, “notas”, “aprovação” ou “reprovação”; e 2) a reprodução deste mesmo olhar examinatório nas avaliações de larga

escala, a partir dos mecanismos de hierarquização de instituições e sujeitos escolares, e de sistemas escolares e suas respectivas cidades, estados, regiões e países.

Para compreender esta avaliação como exame, primeiramente se faz necessário olhar para o trabalho de avaliação dentro das escolas, no qual se destaca a ação do professor quando, em função do modo como historicamente tem se estabelecido o ato de avaliar nessas instituições, precisa reunir todos os dados e estabelecer uma nota e, assim, julgar a criança, o jovem ou o adulto que desempenha papel de estudante no ambiente educacional. Este movimento deu origem ao termo “juízo professoral”, que busca sintetizar o poder simbólico que a avaliação implica. Tal juízo que recai sobre os alunos, como sabemos, não se limita somente a analisar o aprendizado e desenvolvimento individual, mas também se importa, de maneira tácita e até inconsciente, em ponderar as capacidades intelectuais, físicas e até sociais de cada educando, ao passo que extrapola a sala de aula, a relação pedagógica e os objetivos propriamente educativos. Para Bourdieu e Saint-Martin (2015, p. 188):

[...] as taxinomias, que as fórmulas rituais, dos considerandos do julgamento professoral (“as apreciações”) revelam e que se pode supor estruturam o julgamento professoral na medida em que o exprimem, podem ser colocadas em relação com a sanção numerada (a nota) e com a origem social dos alunos que fazem o objeto dessas duas formas de avaliação.

220

O juízo que o professor realiza parte de sua compreensão do sujeito aluno e da forma que este deve agir e se apresentar perante ao corpo docente e à escola, mas essa imagem perfeita e idealizada de aluno, na qual todos os educandos buscam a duras penas se enquadrar, é fruto de uma ideologia de homogeneidade aplicada, ensinada e reproduzida diretamente ou indiretamente pelas experiências e pela formação docente, visando um padrão de educandos que se afasta quase que completamente da realidade. O educador exprime um juízo, neste modelo hegemônico de avaliação, como parte do processo educacional e é dele a responsabilidade de analisar o desempenho de seus alunos, tarefa realizada não somente contemplando a performance do indivíduo que está em avaliação no âmbito dos saberes construídos, mas também as características da *hexis* corporal, que remetem necessariamente à história, à classe social, às disposições culturais incorporadas pelos educandos (Bourdieu & Saint-Martin, 2015).

Vale a pena destacar que este processo de julgamento professoral não revela apenas o sistema de classificações que se materializa no interior das instituições escolares na relação pedagógica entre professores e estudantes, ele nos fala de uma questão mais ampla, posto que é uma das formas de

juízo social a que estamos submetidos hodiernamente. Trata-se, então, de uma questão social, mais especificamente de relações sociais construídas necessariamente sobre processos de distinção e hierarquização sociais: os sujeitos são apartados socialmente, para além das questões objetivas que são mais facilmente observáveis a olhos nu na figuração social, por sistemas de classificação simbólicos que atravessam as diferentes instituições sociais, inclusive a escola e as demais instituições educacionais (Bourdieu, 2007).

Interessa-nos demarcar um movimento pendular entre as avaliações da aprendizagem que ocorrem no contexto escolar e aquelas que passaram a ser praticadas a nível das políticas públicas e que visam escrutinar os sistemas educacionais, as avaliações de larga escala. O movimento de juízo que historicamente partira das mãos do professor em direção ao aluno (avaliação interna) ganha continuidade e novas roupagens no julgamento no qual as escolas passam a ser avaliadas de modo geral por órgãos internacionais, nacionais, estaduais e municipais (avaliação externa). Os sistemas de avaliação da educação, que têm se espalhado globalmente desde a década de 1990, ao selecionarem níveis, etapas e anos da educação básica e superior para estudar o desempenho, satisfação e cumprimento de metas previamente estabelecidas, incorrem na reprodução dos sistemas de classificação das instituições, sujeitos escolares e até de diferentes contextos e sistemas educativos.

221

Segundo Alavarse; Machado e Arcas (2017, p. 1358),

As avaliações externas têm como um de seus principais componentes a aferição do desempenho dos alunos em provas padronizadas e contempla um amplo contingente de participantes, por isso, também, denominada na literatura como avaliação em larga escala, fornecendo em conjunto de informações que pode orientar ações das mais variadas ordens nas políticas educacionais [...].

Levando-se em consideração o papel crucial deste tipo de avaliação para o planejamento de políticas educacionais – já destacado na literatura (Klinger, Deluca & Miller, 2008; Castro, 2009) –, o uso que tem sido feito de seus resultados segue, por vezes, uma lógica fortemente classificatória. Falar, pois, sobre juízo neste primeiro momento é crucial para nosso debate no sentido de se entender o movimento ao qual educação escolar está submetida, e que muitas vezes ocorre de forma um tanto invisível ao hierarquizar alunos, escolas, zonas geográficas, municípios, estados, países e assim por diante. Para Oliveira e Medeiros (2018, p. 721):

[...] as classificações deles decorrentes [dos exames escolares], constituem para Bourdieu uma ocasião privilegiada para se observar a arbitrariedade pedagógica, bem como a ocultação de tal arbitrariedade. Revelam a lógica

de funcionamento do sistema de ensino, cujas avaliações demonstram, ao olhar mais atento, o que está em jogo são não somente os conhecimentos e habilidades efetivamente ensinados, mas também um conjunto de normas implícitas, de critérios difusos e inconscientes da percepção social, que configuram um verdadeiro julgamento de classe e da relação com a cultura, a partir dos parâmetros da classe dominante.

Com isso, evidencia-se que este cenário privilegia um processo de conformação das instituições, dos professores e dos estudantes às avaliações, internas e externas à escola, o que se expressa pela importância delegada aos números, às metas, às posições que cada um vai ocupar e à comparação entre diferentes instituições e contextos. O caso do Brasil pode ser tomado à análise como um retrato de todo este processo, constituindo apropriada antessala para o caso do estado do Amazonas e de sua capital, Manaus.

Com múltiplas influências de debates internacionais, fóruns e conferências e ainda de organismos internacionais (Lima & Gandin, 2019), o Brasil estrutura em 1990 o Sistema Nacional de Avaliação da Educação Básica (SAEB) e, desde então, promove avaliações sistemáticas da educação escolar no país e, com isso, compõe o Índice de Desenvolvimento da Educação Básica (IDEB). Ao longo do tempo este sistema foi sendo modificado de acordo com as demandas de cada panorama de época, oscilando entre as séries e as áreas sobre as quais suas avaliações incidiam – como exemplo temos que em 2019, seguindo as exigências da Base Nacional Comum Curricular (BNCC), passou a avaliar, além de Português e Matemática, os componentes curriculares das áreas de Ciências da Natureza e Ciências Humanas no 9º ano (Oliveira, 2015).

A literatura nacional evidencia que a avaliação de larga escala no Brasil, da década de 1990 aos dias atuais, acabou por consolidar sistemas de avaliação decorrentes de suas experiências federal, estadual e municipal, o que nos permite ratificar, juntamente com Castro (2009), que existe, então, um macrossistema de avaliação da qualidade da educação brasileira. Este macrossistema comporta frentes heterogêneas de avaliação nos dois níveis educacionais, Educação Básica e Educação Superior, e oferece suporte para o delineamento e monitoramento de políticas públicas educacionais (Souza, 2011). Os limites do conjunto dessas avaliações também figuram na literatura especializada, demonstrando que os inegáveis avanços têm esbarrado na timidez do uso dos resultados das avaliações para as escolas, para os docentes e para os educandos (Gatti, 2014).

No caso do estado do Amazonas, a exemplo de outros estados brasileiros, as avaliações internacionais e nacionais desembocaram na necessidade de construção de um sistema próprio para avaliar a educação. O Sistema de

Avaliação do Desempenho Educacional do Amazonas (SADEAM), criado em 2008 pela SEDUC-AM, e a Avaliação de Desempenho do Estudante (ADE), criada em 2014 pela SEMED-Manaus, constituem expressões desse processo de descentralização no âmbito da política de avaliação da educação no país. O nível de gradação que acaba por se materializar no conjunto das avaliações nacional, estadual e municipal coloca as instituições e os sujeitos escolares no estado do Amazonas em constantes atividades didático-pedagógicas de preparação para as diferentes avaliações, o que naturalmente modifica o cotidiano escolar e, conseqüentemente, impacta os processos de ensino e aprendizagem – é, pois, sobre o cenário manauara que nos debruçaremos nos próximos tópicos com vistas a compreender os principais desdobramentos desses processos avaliativos.

O movimento pendular que destacamos no início deste tópico nos mostra que o alcance dos processos avaliativos na educação sofreu um deslocamento: de dentro para fora da escola. As conseqüências deste movimento expressam-se tanto na centralidade quanto no alcance da avaliação educacional de larga escala, se antes o foco estava no processo pedagógico e na relação professor-estudante para pensarmos a aprendizagem e a construção de conhecimento, agora deparamo-nos com uma onipresença da avaliação, que tem gerado, entre outras questões, uma necessidade de preparação para os testes que compõem estes sistemas avaliativos, um desvirtuamento do cotidiano escolar e a subjugação dos objetivos escolares, sobretudo aqueles ligados à aprendizagem e ao desenvolvimento dos educandos, em benefício de uma preparação direcionada unicamente para a obtenção de bons resultados nas avaliações externas.

E, do ponto de vista do julgamento tácito dos processos de avaliação educacional que destacamos inicialmente, tem-se que ele se alastra e alarga-se na mesma medida do alcance dessas avaliações de larga escala. Se antes a literatura preocupava-se com os efeitos do julgamento professoral sobre os educandos, hoje, para além disso, coloca-se imperativamente a necessidade de refletir acerca das conseqüências de um julgamento social generalizado sobre as escolas, os profissionais da educação que nela trabalham e, mais uma vez, sobre os educandos.

3. Situando as avaliações educacionais externas no cenário amazonense

Os estudos acerca da implementação das avaliações externas no Brasil começaram a ser realizados no final dos anos 1980, por influência de outros países que também iniciaram o processo de introdução de sistemas

de avaliações educacionais, o foco neste momento residia na economia e na relação entre Educação e Desenvolvimento. Tais iniciativas voltadas para a educação estavam fundamentadas nas políticas neoliberais que começaram a ganhar força no final da década de 1970 em países como Inglaterra e Estados Unidos, que possuíam governos de direita.

Essas políticas neoliberais surgem como estratégias para limitar as funções do Estado e tornar o mercado econômico o centro da vida social, valorizando os processos de privatizações de setores públicos e a competitividade a nível internacional. Na educação, enquanto setor privilegiado nesta perspectiva de relação direta entre o investimento em educação e o nível de desenvolvimento das nações, o mesmo caminho fora trilhado e, para tal, dentre outras frentes, há a implementação de sistemas avaliativos em larga escala com vistas a garantir uma qualidade competitiva, a nível global, dos sistemas de educação nacionais.

Ao aderir a esse tipo de estratégia, os sistemas educacionais em diferentes contextos geográficos, sociais, econômicos e políticos ao redor do mundo colocam-se na fronteira da noção de “quase-mercado”, que, segundo define Julian Le Grand (1991), é a substituição do monopólio do Estado por vários outros fornecedores concorrentes e autônomos que pretendem não necessariamente obter o lucro através de tais medidas, mas sim conseguir cada vez mais clientes (Afonso, 1999). Consequentemente, a adoção do quase-mercado dentro dos sistemas educacionais públicos é vantajoso somente para o mercado, já que há a diminuição de gastos para este, havendo um controle maior das contas públicas e a redução dos gastos destinados à educação. Além disso, por meio das avaliações externas, o mercado pode controlar o currículo e os conteúdos ministrados nas escolas, com a ênfase no incentivo à competição e à concorrência entre as instituições em níveis nacionais e internacionais na busca por melhores desempenho e resultados. Desse modo, as escolas acabam se tornando verdadeiras empresas, “[...] inclui-se todo o processo educacional na esfera do mercado e generalizam-se os procedimentos e valores típicos do capitalismo competitivo na gestão dos sistemas e das instituições educacionais” (Souza & Oliveira, 2003, p.874).

A primeira experiência com as avaliações em larga escala no Brasil ocorreu em 1990, quando foi criado pelo Ministério da Educação (MEC) o SAEB. “O sistema visava, inicialmente, conhecer as falhas e ou deficiências do contexto educacional brasileiro e, assim, obter um diagnóstico” (Nogueira, 2018, p.25). A partir daí, novas mudanças em relação ao objetivo dessa avaliação começaram a ser realizadas, como por exemplo, em 1993 a avaliação passou a conferir um foco maior ao perfil dos estudantes e às práticas dos professores, assim como também à gestão escolar; ou mesmo as reformulações mais recentes em decorrência da BNCC, que citamos no tópico anterior.

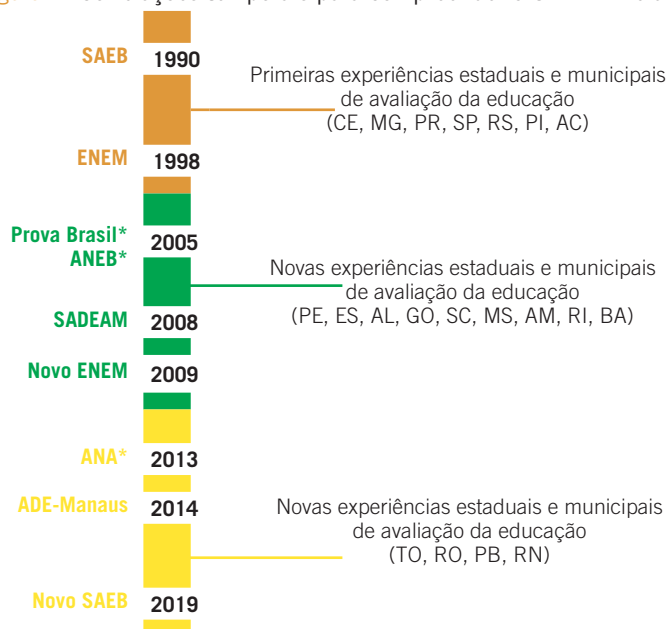
Após a criação do SAEB, outras avaliações externas passaram a figurar também no Brasil, a exemplo de muitos outros países do Sul global. Tais avaliações foram supostamente elaboradas visando medir o nível de desempenho dos estudantes brasileiros das escolas públicas e privadas e, por meio desses resultados, traçar novas estratégias que possibilitassem a melhoria da qualidade educacional no país. Autores como Lima e Gandin (2019) e Coelho (2008) alertam para este cenário de influências externas ou transnacionais nos processos de consolidação de uma hegemonia de discursos e práticas em favor das avaliações de larga escala no Brasil – destacam-se elementos como: o estabelecimento de compromissos com os parâmetros internacional para a educação, a influência de organismos internacionais e a dependência transnacional tanto monetária quanto de modelo de sistema educacional.

É exatamente neste cenário que o SAEB se consolida e passa a fazer parte do cotidiano das instituições escolares em todo o Brasil, figurando como uma instância que determina, dentre outros fatores, a organização didática e pedagógica para o cumprimento das metas estabelecidas. Um dos desdobramentos deste macrossistema de avaliação da educação no país reside na instituição de políticas educacionais de avaliação de larga escala a níveis estadual e municipal.

No caso do Amazonas, o primeiro sistema de avaliação deste tipo fora criado em 2008 com o intuito de avaliar o desempenho dos estudantes amazonenses, estamos a falar do Sistema de Avaliação do Desempenho Educacional do Amazonas (SADEAM). Posteriormente, outras avaliações foram sendo implementadas no estado, inclusive na capital, Manaus, que passou a ter seu próprio sistema de avaliação para averiguar o desempenho dos estudantes manauaras matriculados nos anos iniciais e finais do ensino fundamental e EJA da rede municipal. Esse sistema foi denominado de Avaliação de Desempenho do Estudante (ADE), criado em 2014 pela SEMED-Manaus e tem sido considerado o “[...] projeto pioneiro da SEMED, com o objetivo de avaliar o desempenho dos estudantes e auxiliar diretores, pedagogos e professores no monitoramento dos processos pedagógicos de ensino e aprendizagem a partir da análise dos resultados obtidos” (Breves, Rocha & Monteconrado, 2017, p.16223-16224).

O SADEAM e a ADE, juntamente com os outros sistemas de avaliação em larga escala do país, formam um macrossistema de avaliações voltado para a performance da Educação Básica no Brasil. Desse modo, pode-se observar uma relação histórico-política entre esses diferentes sistemas de avaliação, o que nos auxilia a compreender o lugar e o papel dos sistemas estadual e municipal do Amazonas e de Manaus:

Figura 1. Correlações temporais para compreender o SADEAM e a ADE



* Componentes do SAEB antes da renovação de 2019

Fonte: Elaboração própria.

Antes da renovação do SAEB, anunciada em junho de 2019 e ainda em curso, a Avaliação Nacional de Rendimento Escolar (ANRESC) – também conhecida como Prova Brasil –, a Avaliação Nacional da Educação Básica (ANEB) e a Avaliação Nacional da Alfabetização (ANA) compunham o SAEB. Juntas estas avaliações fornecem o IDEB, que, por sua vez, é considerado um importante indicador censitário da qualidade educacional do país, no entanto, há contradições acerca do seu real objetivo para Educação Básica. Jesus *et al.* (2017, p. 219) afirmam que:

Ao mesmo tempo que a avaliação é necessária como parte do processo pedagógico em sala de aula, é tema controverso quando se trata de avaliação do sistema de ensino porque pode ser tratada desde o ponto de vista didático até questões sociais e políticas que ultrapassam as instituições de ensino e tomam dimensões internacionais de relações de poder. Em síntese, ao Ideb é concedida atribuição de sumarizar a qualidade de ensino oferecida a partir desses dois fatores não parece viável, já que incapaz de refletir a realidade das instituições.

Tais relações de poder nos levam a destacar os limites de um índice como o IDEB, posto que não consegue abarcar em seu espectro quantitativo a complexidade, a diversidade e as especificidades das escolas brasileiras, ao passo que, com a divulgação das médias e resultados finais, acaba por ratificar a cultura classificatória no âmbito das instituições de ensino. Dessa

maneira, percebe-se que a preocupação economicista ganha centralidade, a exemplo do cenário inicial da relação educação-desenvolvimento que tratamos inicialmente neste tópico, e a relação com o mercado se sobressai em detrimento de uma formação cidadã na qual a qualidade [ou o educando e a escola] não se reduz a um índice. Prova disso é que as avaliações que servem até então de base para o IDEB medir o nível educacional no país centram-se somente nas disciplinas de Português e Matemática. Em relação a isso, para Sousa (2014, p.412), a ênfase nestas duas disciplinas “[...] induz a um movimento de homogeneização do que se ensina em todo o Brasil”.

Tomando o SADEAM à análise, e aqui pensando a especificidade do estado do Amazonas, tem-se que este sistema avaliativo avalia os estudantes não só em Língua Portuguesa e Matemática, mas também nas áreas de Ciências Humanas e Ciências da Natureza. Entretanto, é importante ressaltar que as avaliações feitas nas áreas de Ciências da Humanas e da Natureza são destinadas apenas para os alunos do Ensino Médio e Educação de Jovens e Adultos (EJA). No primeiro ano de aplicação do SADEAM, em 2008, foram realizadas provas com os estudantes matriculados nos 5º e 9º anos do Ensino Fundamental, sendo avaliados em Português e Matemática e para os estudantes do 3º ano do Ensino Médio e EJA foram realizadas avaliações nas disciplinas de Biologia, Química, Geografia, História, Língua inglesa e Física (Caed/UFJF, ca 2016). A última edição do SADEAM foi realizada em 2019, e nesta edição as provas foram aplicadas para os alunos do 4º e 7º anos do Ensino Fundamental I e para o 1º ano do Ensino Médio – contudo, destaca-se que os resultados disponibilizados na página eletrônica oficial do SADEAM¹ abarcam somente o período de 2011 a 2015, o que provavelmente está atrelado à parceria estabelecida com o Centro de Políticas e Avaliação da Educação (CAEd), instituição que faz parte da Universidade Federal de Juiz de Fora (UFJF), a partir de 2011.

No caso da ADE, e aqui falamos especificamente da cidade de Manaus, há também esta marca de contemplar áreas do saber para além da Língua Portuguesa e da Matemática,

Nos anos de 2014 a 2017 foram avaliados por meio da ADE, estudantes do ensino fundamental dos anos iniciais e finais (3º ao 9º anos), nos componentes curriculares de Língua Portuguesa e Matemática, além de Ciências da Natureza e Humanidades (para o 8º/9º anos), com três aplicações ao ano, tendo como elementos estruturantes dos testes os documentos supracitados e outros afins. Em 2018, optou-se por avaliar o desempenho dos estudantes dos 3º, 4º, 6º e 8º anos em três momentos, da Educação de Jovens e Adultos-EJA, da 3ª e 4ª fases em duas aplicações e do Projeto Itinerante uma aplicação no decorrer do ano letivo somente

¹ <http://www.sadeam.caedufjf.net/>.

nos componentes curriculares de Língua Portuguesa e Matemática, por entender que o MEC/Inep construirá as matrizes de referências concernentes aos componentes curriculares Ciências da Natureza e Ciências Humanas, alicerçada a Base Nacional Comum Curricular, com aplicação para 2019 (Semed-Manaus, ca 2019, para 4-5).

Observa-se, então, que a ADE atualmente não abarca os 1º e 2º anos do Ensino Fundamental, centrando-se, pois, no intervalo compreendido entre 3º e 9º anos desta etapa de ensino. Destaca-se ainda que em sua página eletrônica oficial da Divisão de Avaliação e Monitoramento², onde encontram-se as informações sobre a ADE, não são disponibilizados os resultados das edições desta avaliação (2014-2019) – sendo este um elemento concreto que incidiu sobre a nossa escolha para a realização do estudo de caso nesta pesquisa, pelo que tivemos que trabalhar com os dados do SADEAM, os quais encontram-se disponíveis atualmente.

Ademais, compreendendo as relações históricas e políticas que servem de antessala para a construção do SADEAM e da ADE, em um cenário de consolidação do próprio SAEB no Brasil e das ingerências internacionais sobre a noção de qualidade dos sistemas educativos em países do Sul global, selecionamos uma escola estadual da capital amazonense, localizada na Zona Oeste e que atende estudantes do 1º ao 5º ano do Ensino Fundamental, para a realização de um estudo de caso. Nosso objetivo com tal estudo de caso é construir uma reflexão aprofundada sobre a instituição em questão e tomá-la como expressão das consequências das avaliações de larga escala para o cotidiano das crianças e jovens nas escolas do estado do Amazonas.

4. Estudo de caso de uma escola da Zona Oeste de Manaus e as implicações das avaliações de larga escala para crianças e jovens nas escolas

Tomar uma escola à análise, com base nos resultados obtidos em uma avaliação de larga escala implica esclarecer dois elementos que se colocam de forma apriorística em nossa análise, a saber: i) não realizaremos uma leitura presa somente aos dados quantitativos, tampouco com uma mirada hierarquizante comparando diferentes escolas, nosso interesse centra-se em, a partir dos dados, problematizar os processos de ensino e aprendizagem nas escolas de Manaus no cenário de onipresença das avaliações de larga escala; ii) o estudo de caso aqui realizado busca, por intermédio da experiência de uma escola, pensar o conjunto mais amplo das relações pedagógicas que acabam por se transformar neste cenário.

² <https://bit.ly/3jEbLtH>.

De maneira geral, quando tomamos a realidade do Amazonas, e mais especificamente de Manaus, à análise, vale a pena chamar atenção para a tentativa de – com os sistemas estadual e municipal de avaliação da educação – melhorar os índices do estado e de sua capital em avaliações outras, nos cenários nacional e internacional. Nesse sentido, a espiral de diferentes avaliações que incidem sobre as escolas, os gestores, os professores e sobre os educandos e suas famílias merece reflexões no sentido de compreender seus efeitos.

Henchen e Brasil (2017), ao analisarem os resultados das avaliações externas na cidade de Manaus, apontam algumas questões que se revelam importantes para nosso debate, a saber: elementos como infraestrutura das escolas, acesso, problemas sociais enfrentados pelos estudantes e pela comunidade escolar de maneira geral provocam reflexos no cotidiano dos alunos e, por sua vez, incidem direta ou indiretamente nos resultados obtidos; vem ocorrendo um movimento de padronização nos registros das ações por parte das escolas avaliadas, o que significa uma acentuada preocupação com os índices e um descompasso com a realidade vivenciada em função do objetivo de se alcançar um resultado considerado satisfatório; as escolas com jornada ampliada lidam com metas altas e têm conseguido responder à altura, o que coloca em xeque os casos das instituições com jornadas menores, esta discrepância acaba por produzir um discurso em prol de maior dedicação de tempo para a preparação dos alunos para essas avaliações – percebe-se, portanto, que se adota uma preocupação com os resultados e não necessariamente com a qualidade do ensino, já que o tempo de aplicação das atividades escolares passa ser reduzindo aos objetivos das avaliações externas e passa a atender somente uma ordem sistêmica de cumprimento de rankings.

229

Aliado a isso, os estágios curriculares e aqueles não-obrigatórios, bem como as experiências com o Programa Institucional de Bolsa de Iniciação à Docência (PIBID) e com a Residência Pedagógica, ao colocarem os acadêmicos das licenciaturas em contato direto com a realidade das escolas estaduais de Manaus, têm promovido uma observância em torno do lugar que as avaliações de larga escala têm ocupado no cotidiano dessas escolas. Nota-se uma ênfase nos processos de preparação das crianças e jovens para as provas que compõem as avaliações locais e nacionais, pelo que as necessidades de aprendizagem dos educandos, muitas vezes, são deixadas de lado, especialmente em tempos em que essas avaliações se aproximam.

O deslocamento pendular da aprendizagem e desenvolvimento dos estudantes para o treinamento para responder a essas avaliações externas vem modificando o cotidiano escolar e, conseqüentemente, a mediação e a relação pedagógica. Quando a avaliação de larga escala provoca uma ruptura com os processos de ensino e aprendizagem escolar e gera uma preparação

quase que frenética para a feitura de uma prova pontual, é preciso, então, avaliar os próprios pressupostos que compõem as políticas públicas de avaliação educacional.

No caso da *Escola Estadual Amarela* – nome fictício que utilizaremos a partir deste momento para tratar da escola estudada –, localizada no bairro São Jorge na Zona Oeste de Manaus e atendendo do 1º ao 5º ano do Ensino Fundamental (EF), promovemos uma correlação entre os resultados obtidos no SADEAM 2011-2015 nas diferentes disciplinas. O público atendido por esta escola é composto por crianças que se encontram matriculadas nos anos iniciais do Ensino Fundamental, a avaliação do SADEAM incidiu sobre os 3º e 5º anos, sempre nas disciplinas de Língua Portuguesa – no caso do 3º ano esta disciplina está caracterizada como Escrita e Leitura – e Matemática. Nossa análise toma como recorte as avaliações realizadas com o 3º ano do EF, o qual, no caso da *Escola Estadual Amarela*, foi avaliado nas edições 2011, 2012 e 2015.

Inicialmente, tratemos do nível de envolvimento dos estudantes da *Escola Estadual Amarela* ao SADEAM, o qual pode ser observado pela taxa de participação nas diferentes edições com as quais estamos a trabalhar:

230

Tabela 1. Taxa de participação dos estudantes no SADEM (2011-2015): o caso da Escola Estadual Amarela

2011			
Disciplinas avaliadas	Participação		
	Previsto	Efetivo	%
Língua Portuguesa (Escrita e Leitura)	167	127	76%
Matemática	166	139	83,7%
2012			
Disciplinas avaliadas	Participação		
	Previsto	Efetivo	%
Língua Portuguesa (Escrita e Leitura)	177	144	81,4%
Matemática	177	163	92,1%
2015			
Disciplinas avaliadas	Participação		
	Previsto	Efetivo	%
Língua Portuguesa ¹	138	119	86,2%
Escrita	138	118	85,5%
Matemática	138	123	89,1%

¹. Em 2015, a área anteriormente denominada “Leitura” passou a ser chamada de “Língua Portuguesa” e a ter seu instrumento aplicado em dia diferente da Escrita.

Fonte: Elaboração própria, com base em Caed/UFJF (ca 2016).

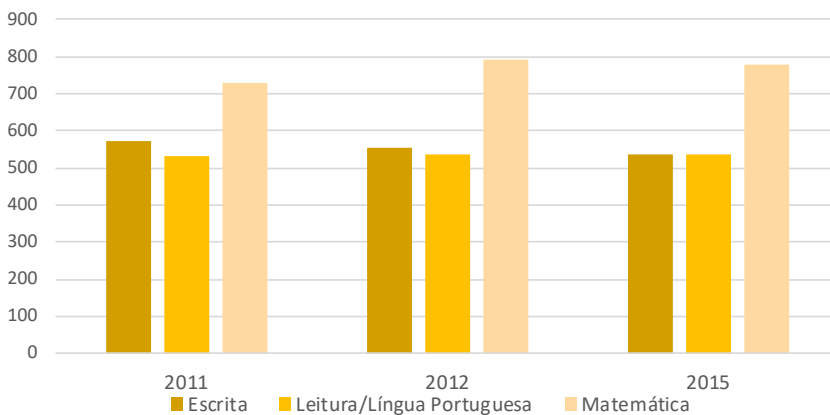
Observa-se que, ao longo do período analisado há oscilações distintas quanto à participação dos estudantes nas áreas do conhecimento avaliadas na escola em questão: enquanto em Língua Portuguesa a adesão foi crescendo nesses anos, em Matemática não há uma linearidade progressiva; por outro lado, mesmo com maior grau de irregularidade nos anos analisados, é na área de Matemática que sempre há maior participação quantitativa ano a ano. De modo geral, pode-se afirmar que a participação das crianças matriculadas no 3º ano do EF na *Escola Estadual Amarela* no SADEAM tem se revelado bastante expressiva, o que pode ter como elemento explicativo o apelo que as escolas passaram a fazer no sentido de garantir a presença dos estudantes quando da aplicação das provas.

Sabe-se que o SADEAM utiliza cinco elementos para compor os resultados das escolas analisadas, são eles: a matriz de referência, a teoria de resposta ao item (TRI), o caderno de testes, a escala de proficiência e os padrões de desempenho (Oliveira, 2015). As escolas são classificadas, pois, com relação ao desempenho de seus estudantes – “Abaixo do Básico”, “Básico”, “Proficiente” e “Avançado” constituem os padrões de desempenho desenhados no SADEAM. Nesse sentido, os resultados apontam o percentual de educandos que se encontra em cada um desses padrões, sendo o conjunto da performance dos estudantes determinante para o estabelecimento do índice de cada escola em cada uma das disciplinas avaliadas.

231

Os resultados da *Escola Estadual Amarela* podem ser sintetizados da seguinte forma:

Gráfico 1. Resultados SADEAM 2011-2015: 3º ano do Ensino Fundamental da Escola Estadual Amarela



Fonte: Elaboração própria, com base em Caed/UFJF (ca 2016).

A discrepância entre as duas grandes áreas do saber representadas no gráfico acima é notória, vale a pena destacar que ela se mantém constante para todas as escolas do estado do Amazonas que foram avaliadas no âmbito do SADEAM, e acaba por nos falar tanto das diferenças das respectivas matrizes de referência de cada uma dessas áreas, quanto do desafio histórico colocado, no Amazonas, em Manaus e no Brasil, sobre a Língua Portuguesa, a alfabetização, a leitura, escrita e a compreensão e interpretação. Com base nas matrizes de referência e nos padrões de desempenho de cada área, é curioso que em Escrita e Leitura/Língua Portuguesa tenhamos mais estudantes na categoria “Avançado”, enquanto em Matemática o que se tem é que a maior parte dos estudantes do 3º ano do EF da escola analisada encontram-se na categoria “Básico” (Caed/UFJF, ca2016).

Chama atenção ainda as variações nas três áreas representadas no *Gráfico 1*: i) vê-se que, no caso da Escrita, os anos analisados viram desenhar-se um movimento decrescente de quase 40 pontos entre 2011 e 2015; ii) em Leitura/Língua Portuguesa temos a área que manteve maior estabilidade nos três anos que aqui figuram em uma média de 5 pontos para mais ao final deste período; iii) já a disciplina de Matemática apresentou um salto de pouco mais de 20 pontos entre 2011 e 2012, porém, desceu quase 14 pontos de 2012 para 2015. No cenário mais amplo, faz-se mister destacar que a *Escola Estadual Amarela*, a exemplo de muitas outras escolas estaduais que vêm sendo avaliadas pelo SADEAM desde 2008, apresenta índices bastante próximos daqueles alcançados pelas médias do estado do Amazonas e da própria Coordenadoria Distrital de Educação (CDE) na qual está inserida³.

Diante desses dados, ao pensarmos a participação das crianças do 3º ano do Ensino Fundamental nas avaliações de larga escala e a construção dos índices de uma escola específica, emerge a necessidade de interrogar os processos que envolvem a política de avaliação educacional, sobretudo no que concerne aos seus desdobramentos no contexto educacional e escolar do Amazonas. Objetivamente falando, pode-se apontar as seguintes ações encabeçadas pela SEDUC-AM na sequência ao SADEAM: a criação do Fundo Estadual de Incentivo ao Cumprimento de Metas da Educação Básica e a ligação com a política de bonificação para as escolas que atingem as metas ou, ao menos, melhoram seus índices – a exemplo do “Prêmio Escola de Valor”; a criação e implementação, em 2013, do “Programa de Intervenção Pedagógica do Amazonas” (PIPA), cujo objetivo consiste em oferecer assessoramento didático-pedagógico para as escolas com baixo desempenho no

³ Trata-se da Coordenadoria Distrital de Educação 04, que abarca escolas das Zonas Oeste e Centro-Oeste de Manaus.

SADEAM e no SAEB; em 2011, nesta mesma linha, a secretaria constituiu o “Programa Criando Oportunidades”, com vistas a garantir reforço escolar nas escolas com índices baixos; dentre outras ações (Oliveira, 2015).

Se o discurso que se legitimou desde 2008 [ano de criação do SADEAM] tem girado em torno da melhoria da educação no estado do Amazonas, faz-se mister interrogar tanto este sistema de avaliação de larga escala, quanto as iniciativas de reforço que o seguiram no sentido de compreender suas potencialidades e seus limites. Que impactos o SADEAM, a ADE e todas as ações supracitadas têm desenhado no cotidiano das escolas?

Cabe-nos, portanto, uma reflexão em torno do que estamos a fazer com as crianças e com os jovens nas escolas neste cenário de supervalorização dos índices e das hierarquias que são construídas a partir deles. Nossos educandos da Educação Básica têm se deparado com uma formação que os coloca desde muito cedo em um processo de auto adequação às provas e testes previstos nas avaliações de larga escala. O brincar, a socialização, a própria interação com os pares, o trabalho colaborativo são exemplos de elementos que têm sido secundarizados, posto que a lógica do treinamento para sair-se bem nas avaliações de larga escala prevê um trabalho solitário de memorização do formato, do conteúdo e das respostas que devem ser dadas às perguntas postas nesse tipo de avaliação.

Em que pesem os avanços que a política de avaliação educacional trouxe para os processos de elaboração de novos caminhos e de monitoramento da educação nacional – e aqui a literatura especializada já destaca tais avanços e sua importância (Castro, 2009; Fontanive, 2011) –, faz-se mister compreender, por outro lado, os limites que estão postos no que usualmente ocorre quando da ampla publicação dos resultados das avaliações externas, nomeadamente: a hierarquização das instituições escolares, dos bairros, das cidades, estados e até dos países em função de sua performance nos testes aplicados, o que naturalmente implica uma hierarquia dos sujeitos que compõem essas instituições; a ratificação da cultura classificatória no campo educacional, que reproduz um sistema de classificações que distingue escolas “boas” e “ruins”; a responsabilização e culpabilização dos sujeitos escolares (gestores, professores e educandos e suas famílias) pelos resultados considerados negativos; entre outros.

5. Considerações Finais

A educação vem sendo tomada, desde a Modernidade, como um elemento crucial para a ideia de desenvolvimento social e econômico dos países, o que significa necessariamente que as ações e os direcionamentos dos Estados nacionais no âmbito das políticas de educação são reveladores dos projetos de nação. As políticas públicas de avaliação da educação inserem-se nesta lógica, ao passo que historicamente ligam-se a processos amplos como a mundialização, a competitividade e mercadorização da educação, bem como às influências supranacionais nos sistemas educativos dos países.

No cenário educacional brasileiro, com a criação do SAEB em 1990 houve a necessidade de diagnosticar o desempenho dos estudantes brasileiros, bem como pensar a questão da qualidade da educação no país. Posteriormente, a implementação de tal sistema avaliativo influenciou governos estaduais e municipais a criarem seus próprios sistemas de avaliação educacional de larga escala. Contudo, ao lado dos objetivos ligados à monitorização e aprimoramento da educação, o real sentido de uma avaliação diagnóstica para identificar os possíveis problemas educacionais que necessitam ser sanados levando-se em consideração as especificidades de cada instituição de ensino, tem sido secundarizado. As avaliações externas de larga escala tornaram-se, pois, ferramentas políticas elaboradas com o intuito de atingir as metas pré-estabelecidas pelo governo e por agências especializadas, acompanhando o desenvolvimento de múltiplos instrumentos avaliativos [locais, nacionais e até internacionais] para a construção de uma imagem educacional paradigmática e, por vezes, alienada da realidade e, em contrapartida, a competição, a hierarquização e a classificação, a partir dos resultados amplamente divulgados dessas avaliações externas, ganham força dentro das escolas e no imaginário social.

A adoção de tais sistemas exerce grande impacto direto na organização didático-pedagógica das instituições de ensino, uma vez que, mediante a eles, há um foco maior para o treinamento de conteúdos e habilidades que são exigidos exclusivamente por essas avaliações, o objetivo do processo de ensino-aprendizagem acaba, pois, sendo subvertido no interior das escolas, ao passo que elas são bombardeadas com múltiplas avaliações externas atualmente, passando a privilegiar os produtos dessas avaliações: os resultados, os índices e a posição frente a outras escolas. A aprendizagem e o desenvolvimento dos educandos são, portanto, secundarizados.

O caso da escola estadual aqui estudada acabou por demonstrar a influência de tais avaliações no cotidiano escolar das crianças matriculadas nas escolas públicas de Manaus. Nesse sentido, ratificamos a relevância de se enfrentar o debate acerca dos efeitos das avaliações externas para o

processo educativo. Ressalta-se que o tema abordado – seja em Manaus, em outros municípios amazonenses, e mesmo em outros estados da federação – necessita de mais estudos e pesquisas para continuarmos a refletir sobre o alcance das avaliações de larga escala e sua relação com a noção socialmente referenciada de qualidade da educação.

Referências

- Afonso, A. J. (1999). Estado, mercado, comunidade e avaliação: Esboço para uma rearticulação crítica. *Educação & Sociedade*, 69, 139-164.
- Alvarse, O. M., Machado, C. & Arcas, P. H. (2017). Avaliação externa e qualidade da educação: formação docente em questão. *Rev. Diálogo Educ.*, 17(54), 1353-1375.
- Bourdieu, P. (2007). *A Distinção: crítica social do julgamento*. São Paulo/Porto Alegre, EDUSP/Zouk.
- Bourdieu, P. & Saint-Martin, M. (2015). As categorias do juízo professoral. In Maria Alice Nogueira & Afrânio Catani (Org.). *Escritos de educação* (pp. 205-241). Petrópolis, RJ: Vozes.
- Brasil. *Ideb*. Portal do Mec. Recuperado de <http://portal.mec.gov.br/conheca-o-ideb>.
- Breves, N. S. P., Rocha, S. E. A., & Monteconrado, G. M. (2017). Sistema de avaliação de desempenho educacional - SADEM: experiência na rede pública municipal de ensino de Manaus. *13º Congresso Nacional de Educação/EDUCERE*, Curitiba. Recuperado de <https://bit.ly/34Y7PxU>.
- Caed/UFJF. *O SADEAM*, Juiz de Fora, ca 2016. Recuperado de <https://bit.ly/2FpaeZF>.
- Castro, M. H. G. (2009). Sistemas de avaliação da educação no Brasil: avanços e novos desafios. *São Paulo em Perspectiva/Fundação Seade*, 23(1), 5-18.
- Coelho, M. I. M. (2008). Vinte anos de avaliação da educação básica no Brasil: aprendizagens e desafios. *Ensaio: Aval. Pol. Públ. Educ.*, 16(59), 229-258.
- Fontanive, Nilma. (2011). O uso pedagógico dos testes. In Alberto de M. Souza (Org.). *Dimensões da Avaliação educacional* (pp. 139-173). Petrópolis: Vozes, 2011.
- Freitas, L. C. (2003). *Ciclos, seriação e avaliação: confronto de lógicas*. São Paulo: Moderna.
- Gatti, B. (2014). Avaliação: contexto, história e perspectivas. *Olh@res*, 2(1), 08-26.
- Henchen, A. K. & Brasil, D. S. B. (2017). Os desafios ambientais das escolas do campo municipais fomentados pelo Instituto Águila nos resultados das avaliações externas municipais da cidade de Manaus/AM. *Revista Eletrônica Pesquisaeeduca*, 9(17), 201-211.
- Hoffmann, Jussara. (2017). *Avaliação: mito e desafio: uma perspectiva construtivista*. Porto Alegre: Mediação.
- Jesus, E. L. et al. (2017). Avaliação Educacional em Manaus-AM: o IDEB em questão. *Areté-Revista Amazônica de Ensino de Ciências*, 10(21), 216-229.

- Klinger, D., Deluca, C. & Miller, T. (2008). The evolving culture of large-scale assessments in Canadian education. *Canadian Journal of Educational Administration and Policy*, 3, 1-34.
- Le Grand, J. (1991). Quasi-markets and social policy. *The Economic Journal*, 101, 1256-1267.
- Leite, S. A. S. & Kager, S. (2009). Efeitos aversivos das práticas de avaliação da aprendizagem escolar. *Ensaio: aval. pol. públ. Educ.*, 17(62), 109-134.
- Lima, I. G. & Gandin, L. A. (2019). O contexto da consolidação das avaliações em larga escala no cenário brasileiro. *Educ. Soc.*, 40, 1-28.
- Luckesi, C. C. (2000). O que é mesmo o ato de avaliar a aprendizagem? *Pátio On-line*, 3(12), 1-7.
- Nogueira, E. L. (2018). *Práticas Influentes na Melhoria dos Índices educacionais: o caso de uma escola do Ensino Fundamental I em Manaus*. Dissertação (Mestrado Profissional em Gestão e Avaliação da Educação Pública) – Universidade Federal de Juiz de Fora, Juiz de Fora.
- Oliveira, J. M. & Medeiros, C. C. C. (2018). Categorias do juízo professoral: entre interpretações, julgamentos e implicações. *Estud. Aval. Educ.*, 29(72), 710-737.
- Oliveira, S. M. S. (2015). *O Sadeam e o Ensino de Língua Portuguesa em duas Escolas de Manaus: Implicações do fazer pedagógico para o Letramento*. Dissertação (Mestrado Profissional em Gestão e Avaliação da Educação Pública), Faculdade de Educação, Universidade Federal de Juiz de Fora, Juiz de Fora.
- Semed-Manaus. ADE, Manaus, ca 2019. Recuperado de <https://bit.ly/3ixDD1a>.
- Souza, A. M. (Org.). (2011). *Dimensões da Avaliação educacional*. Petrópolis: Vozes.
- Sousa, S. Z. (2014). Concepções de Qualidade da Educação Básica Forjadas por meio de Avaliações em Larga Escala. *Avaliação*, 19(2), 407-420.
- Souza, S. Z. & Oliveira, R. P. (2003). Políticas de Avaliação da Educação e Quase Mercado no Brasil. *Educ. Soc.*, 24(84), 873-895.

La evaluación para el ingreso al servicio educativo y su impacto en el rendimiento de la prueba PISA en México

Juan Rubén Compañ García 

Gobierno del estado de México

Resumen. El presente artículo tiene como propósito caracterizar y analizar la política de ingreso al servicio docente implementada por la administración de gobierno (2012-2018) y el impacto que pudo tener en los resultados de PISA 2018 en lectura, matemáticas y ciencias. Se utilizó a la evaluación para el ingreso al servicio educativo como un canal de comunicación, cuya finalidad fue establecer un nuevo sistema de regulación en el sistema educativo nacional, dirigidos a atraer a los mejores docentes y elevar la calidad educativa. El estudio advierte que se presentó una tendencia hacia mejorar los niveles de rendimiento en los docentes que ingresan al servicio educativo, lo cual pudo tener un efecto positivo en los logros de aprendizaje de los y las estudiantes, así como debilidades en el proceso de implantación del nuevo sistema de regulación de ingreso al servicio educativo. Se concluye en el texto, que se necesitan dispositivos de política educativa que se les dé seguimiento en términos de evaluarlos y focalizarlos. La política educativa tiene que gozar de cierta estabilidad, coherencia, coordinación y buscar el bienestar general, si de verdad se aspira a la calidad educativa.

Palabras clave: calidad educativa; evaluación docente; PISA; política educativa.

A avaliação para admissão ao serviço educacional e seu impacto no desempenho da prova PISA no México

Resumo. O objetivo deste artigo é caracterizar e analisar as prioridades das políticas educacionais implementadas pela administração do governo (2012-2018) e associá-las ao impacto que o PISA poderia ter nos últimos resultados, em termos de acesso, equidade e qualidade de a educação. A abordagem de prioridades é usada para estudar as seguintes prioridades de política educacional: 1) acesso à escola, 2) o Programa para Eliminar Atraso Educacional, o Programa de Escolas em Tempo Integral e o Programa de Inclusão e Alfabetização Digital e 3) a política de avaliação de professores. O estudo percebe, através das evidências, que algumas ações tiveram um efeito positivo nas realizações de aprendizagem dos alunos, bem como fragilidades consideráveis no processo de implementação de políticas educacionais. Há necessidade de dispositivos de políticas educacionais que sejam acompanhados em termos de avaliação e direcionamento deles. A política educacional deve gozar de certa estabilidade, coerência, coordenação e buscar o bem-estar geral, se realmente aspira à qualidade e à equidade da educação.

Palavras-chave: qualidade educacional; avaliação de professores; PISA; política educacional.

The assessment for the access to the education service and its impact on the performance of the PISA test in Mexico

Abstract. The purpose of this article is to characterize and analyze the educational policy priorities implemented by the government administration (2012-2018), and associate them with the impact that PISA could have on the latest results, in terms of access, equity and quality of The education. The priorities approach is used to study the following educational policy priorities: 1) access to school, 2) the Program to Eliminate Educational Lag, the Full-Time Schools Program and the Inclusion and Literacy Program Digital and 3) the teacher evaluation policy. The study realizes, through the evidence, that some actions had a positive effect on the learning achievements of the students, as well as considerable weaknesses in the process of implementing educational policies. There is a need for educational policy devices that are followed up in terms of evaluating and targeting them. The educational policy must enjoy a certain stability, coherence, coordination and seek general well-being, if it really aspires to educational quality and equity.

Keyword: educational quality, staff assessment, PISA, education policy.

1. Introducción

El Programa para la Evaluación Internacional de Estudiantes (PISA, por sus siglas en inglés), es una evaluación que se realiza con la participación de países miembros de la OCDE (Organización para la Cooperación y el Desarrollo Económico) y otros que no lo son –el número de participación de países varía en cada aplicación– a jóvenes de 15 años de edad cada tres años. La prueba evalúa tres áreas: lenguaje, matemáticas y ciencias. De acuerdo al rendimiento de aprendizaje que obtienen los y las estudiantes, es que se les clasifica a los países por no tener un desempeño favorable o resaltar los logros. En el caso de México, los resultados en los logros de aprendizaje en las pruebas (2000, 2003, 2006, 2009, 2012, 2015 y 2018) lo ubican con un desempeño no favorable, que está por debajo del promedio de la OCDE en las tres áreas evaluadas (Backhoff, Bouzas, Contreras, Hernández y García, 2007; Santiago, Mc Gregor, Nushe, Ravela y Toledo, 2012; Martínez, 2017; INEE, 2016).

El rendimiento en los logros de aprendizaje que tienen los países- los resultados no se dan por escuela ni por alumno, sino sobre el sistema educativo en su conjunto-, se utiliza como referente importante para diseñar e implementar acciones de política pública por los responsables -no tanto de prácticas de enseñanza y aprendizaje-, con aras a mejorar los resultados y a hacer un buen uso de los mismos (Martínez, 2012). Contextualizar los resultados permite, entre otras cuestiones, considerar aspectos no cognitivos como el contexto familiar del alumno, entornos de aprendizaje, factores demográficos, sociales y económicos (OCDE, 2017; Martínez, 2016) además de diversas acciones de política educativa que impulsan las diferentes administraciones de los gobiernos en relación a mejorar los aprendizajes (Rivas, 2015; Muñoz, 2005).

Existen diversas investigaciones que consideran diferentes aspectos asociados a los aprendizajes, con la finalidad de comprender con mayores elementos, el bajo nivel de rendimiento que los y las estudiantes logran en las pruebas. En estos estudios encuentran que factores como lo social, económico, político y educativo, están fuertemente asociados al rendimiento (Blanco, 2012; Muñoz Izquierdo, 2009; Backhoff *et al.*, 2007).

En ese tenor, el presente artículo tiene como propósito caracterizar y analizar la política de ingreso al servicio educativo implementada por la administración de gobierno (2012-2018) y el impacto que pudo tener en los últimos resultados de PISA. El artículo se integra por 5 apartados. El primero hace alusión al punto de referencia teórica, el cual se apoyará en los conceptos de canales de comunicación y dispositivos propuestos por Axel Rivas. El segundo habla, a modo de contextualizar, sobre algunos indicadores del sistema

educativo nacional. El tercer apartado presenta los resultados obtenidos en la prueba PISA en lectura, matemáticas y ciencias. El cuarto apartado aborda la política de evaluación para el ingreso al servicio educativo impulsada por la administración de gobierno (2012-2018) y su relación con los logros de aprendizaje de PISA. El quinto apartado cierra con los comentarios finales.

2. Punto de referencia teórica: canales de comunicación y dispositivos

La política educativa no se puede comprender de forma unidireccional, son prácticas de gobierno (Hunter, 1998; Rose, 1999) que interactúan de forma diversa y compleja a partir de sus actores e instituciones y que se modifican en sus diferentes niveles de gobierno (micro, meso y macro). Las prácticas de gobierno priorizan ciertas acciones y recursos, otras, no las consideran. Atienden a determinados sectores, otros, los excluyen (Rivas, 2015). Es importante evaluar las características de las políticas educativas más que su contenido formal (Scarstacini, Spiller y Stein, 2010; Rivas, 2015; Bernstein, 1998).

La política educativa utiliza diversos medios para materializarse en el conjunto del sistema educativo nacional, los canales de comunicación y los dispositivos, son un ejemplo. Los canales de comunicación “son mecanismos recurrentes con una lógica de intervención que establece niveles de coerción sobre actores e instituciones” (Rivas, 2015) que pueden ser impuestos sin mayor problema, o resistidos por diversos sectores (Pedró y Puig, 1998). Los canales de comunicación son sistemas que permiten establecer la comunicación entre el Estado y las instituciones educativas, con miras a atender la práctica educativa.

Rivas (2015) refiere que los canales de comunicación pueden ser de dos tipos: a) unos que funciona como arterias y b) otros que son más específicos. En el primero pueden ser: el diseño curricular; la distribución de materiales (en especial los libros de texto), la evaluación de los alumnos y los regímenes académicos, la supervisión escolar, la formación docente inicial y continua, la carrera docente y la normativa educativa. En el segundo: concursos de cargos específicos, planes de mejora, políticas compensatorias y programas institucionalizados específicos.

De acuerdo con Rivas (2015) los canales de comunicación pueden tener larga historia, como el acceso y la equidad en los sistemas educativos o pueden crearse de acuerdo a las necesidades de ese momento. El sistema

de evaluación de la carrera docente que se propuso durante la reforma educativa del 2013 fue un canal de comunicación nuevo que reguló el sistema educativo nacional y que tuvo un impacto sustancial.

Los canales de comunicación de la política educativa, se convierten en dispositivos cuando producen un efecto de encauzamiento de las acciones del sistema educativo, regulan prácticas y generan cierto control de los resultados. Los dispositivos “son la conversión de esos canales en motores de regulación de las prácticas con una orientación definida hacia resultados concretos” (Rivas, 2015, p. 321). La finalidad de los dispositivos es aprovechar los canales de comunicación y convertirlos en sistemas de regulación dirigidos a resultados específicos.

Los dispositivos al tener cierto control de los resultados y una orientación de los mismos, evalúa los resultados de las políticas que implementa para identificar que escuelas son las que necesitan los recursos o servicios, hace un seguimiento puntual de cómo son recibidos y utilizados los recursos y puede delegar responsabilidades a distintos actores.

240

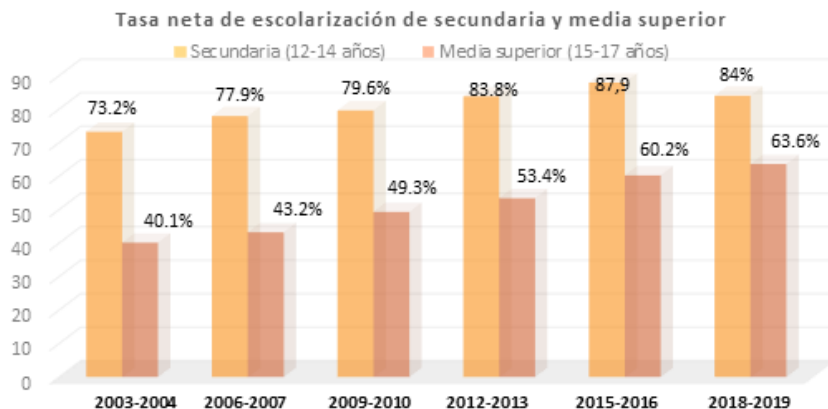
Los dispositivos pueden ser autoritarios o democráticos. Puede ser que el dispositivo sea muy centralizador, pero se base en las consultas sistémicas que reconozcan y redistribuyan las asimetrías sociales. O ser democrático y permita la participación de múltiples voces. Los dispositivos, pueden manipular conciencias o generar la capacidad de reflexión crítica y de compromiso social. Los dispositivos cambian la vida de las personas de forma masiva y sistemática (Rivas, 2015).

Una vez que se estableció el encuadre o el punto de referencia teórica, es momento de pasar al siguiente apartado, que abordará algunos indicadores del sistema educativo nacional, en relación a la población que atiende y a las oportunidades de acceso a la escuela en relación a la posición en la distribución del ingreso per cápita del hogar.

3. Contexto mexicano

El Sistema Educativo Mexicano, mantiene un crecimiento importante en la tasa neta de escolarización en la población de 12 a 17 años de edad, la cual estudia, la educación secundaria o media superior. Del año 2003 al año 2019, la mejora de los indicadores es notable (Gráfica 1). El sistema escolar pasó de tener una matrícula de 4 858 083 en secundaria (12-14 años) para 2003 a 5 634 455 en 2018, un incremento en la matrícula de 776 372. En media superior (15-17 años) se pasó de una matrícula de 2

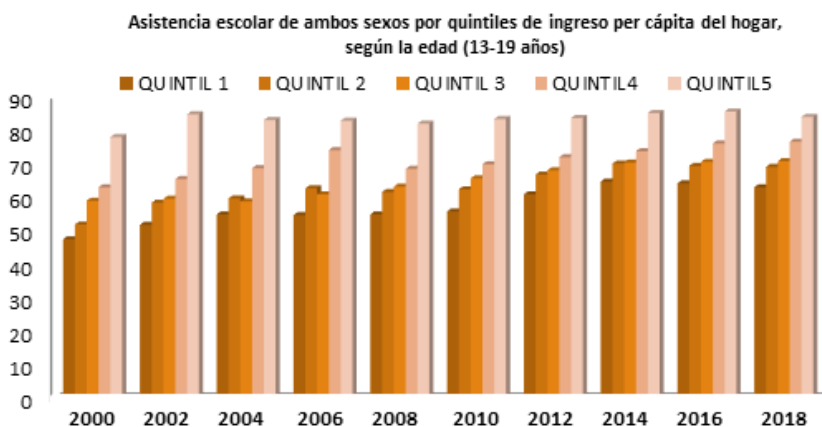
569 650 en 2003 a 4 233 240 en 2018, lo que implicó un crecimiento de 1 663 590 en la matrícula, es decir, México tuvo una mayor proporción de jóvenes escolarizados (Backhoff, 2019).



Gráfica 1. Tasa neta de escolarización en secundaria y media superior.
Fuente: SEP (2020).

Existe un importante avance en los indicadores en relación a la tasa neta de escolarización de los alumnos y alumnas en los niveles de secundaria y media superior, sin embargo, la igualdad de oportunidades educativas, no se desarrolla al mismo ritmo, pues las desigualdades aún son amplias. Aunque es importante referir, que hay avance en los indicadores (Gráfica 2).

241



Gráfica 2. Asistencia escolar de ambos sexos por quintiles de ingreso per cápita del hogar, según la edad.
Fuente: CEPALSTAT (2020).

La asistencia a la escuela de los sectores menos favorecidos, presenta un aumento importante. Para el año 2000 se tiene un 47.6% de asistencia escolar del quintil 1 (el menos favorecido) para incrementar a 63.1% en el año 2018, lo que equivale a un crecimiento de 15.5%. La mejora de los indicadores van a la alza, aunque la asimetría de los niños y niñas de los hogares más pobres con respecto a los hogares con condiciones más favorables, todavía son amplias. Se tiene para el 2018, una asistencia escolar del quintil 5 (sector más favorecido) del 84.1% contra 63.1% del sector menos favorecido (quintil 1), lo que señala una diferencia de 21%. Lo anterior advierte que mientras las oportunidades escolares para los niños y niñas de los hogares más pobres van en aumento, de manera paralela, la desigualdad entre sectores, continúa. Hay un acceso desigual a la escuela, que se relaciona de manera importante con la posición en la distribución de ingreso. Las oportunidades de acceso a la escuela no son equitativas, esto es lo que los datos muestran, a pesar de la mejora en los indicadores.

Una vez que se presentaron los indicadores en términos de la población que atiende el sistema educativo nacional y las oportunidades de acceso a la escuela en relación a la posición en la distribución del ingreso, es pertinente identificar los resultados de aprendizaje alcanzados por la población estudiantil, con la finalidad de tener una caracterización más amplia del sistema educativo nacional y, ello nos sirva como herramientas útiles para el posterior análisis.

4. Resultados de PISA en México

El Programa Internacional de Evaluación de Estudiantes (PISA, por sus siglas en inglés) mide la capacidad de los estudiantes de 15 años de edad para responder a los desafíos de la sociedad actual, que valora lo que las personas pueden hacer con lo que saben y no tanto lo que saben (OCDE, 2018). PISA examina si el alumno es capaz de aplicar su conocimiento a diversos contextos escolares y no escolares. Las evaluaciones de PISA se realizan cada tres años. Evalúan tres áreas de conocimiento: lectura, matemáticas y ciencias, basados en los marcos de PISA (OCDE, 2018) Cabe referir, que en cada aplicación de la prueba se enfatiza un área en específico. En el año 2000 (primera aplicación de la prueba) se concentró en la lectura; 2003, matemáticas y 2006 en ciencias, con ello se termina el primer ciclo. El segundo ciclo inicia en 2009 con lectura; 2012, matemáticas y 2015 ciencias. Para 2018, inicia el tercer ciclo de evaluación con el énfasis en lectura.

PISA utiliza ocho niveles para captar la progresión de la dificultad de la competencia lectora, el nivel más alto es el seis y el más bajo es el 1c, cada uno con sus características. Para matemáticas y ciencias son 7 niveles, el nivel más alto es el 6 y el nivel más bajo es el que está por debajo del nivel 1, como se muestra en la tabla 1.

Tabla 1. Descripción de los niveles de competencia de lenguaje, matemáticas y ciencias.

Nivel	Características Lectura	Características Matemáticas	Características Ciencias
6	Las tareas en este nivel generalmente requieren que el lector haga varias inferencias, comparaciones y contrastes que son a la vez detallados y precisos. Requieren la demostración de una comprensión completa y detallada de uno o más textos y pueden implicar la integración de la información de más de un texto. Las tareas pueden requerir que el lector se encuentre con ideas desconocidas, en presencia de información destacada de carácter similar, y que genere categorías abstractas de interpretaciones. Las tareas de reflexionar y evaluar pueden requerir que el lector emita hipótesis sobre o evalúe críticamente un texto complejo sobre un tema desconocido, teniendo en cuenta varios criterios o puntos de vista, y que aplique comprensiones sofisticadas más allá del texto. Una condición relevante para las tareas de acceder y obtener en este nivel es la precisión del análisis y la atención al detalle que es poco visible en los textos.	Los alumnos saben formar conceptos, generalizar y utilizar información basada en investigaciones y modelos de situaciones problemáticas complejas, así como usar sus conocimientos en contextos relativamente no habituales. Pueden relacionar diferentes fuentes de información y representaciones y traducirlas entre ellas de manera flexible. Los estudiantes de este nivel poseen un pensamiento y razonamiento matemático avanzado. Estos alumnos pueden aplicar su entendimiento y comprensión, así como su dominio de las operaciones y relaciones matemáticas simbólicas y formales, para desarrollar nuevos enfoques y estrategias para abordar situaciones nuevas. Los alumnos pertenecientes a este nivel pueden reflexionar sobre sus acciones y formular y comunicar con exactitud sus acciones y reflexiones relativas a sus descubrimientos, interpretaciones, argumentos y su adecuación a las situaciones originales.	Los estudiantes pueden recurrir a ideas y conceptos científicos interrelacionados de las ciencias físicas, de la vida y de la Tierra y el espacio y utilizar sus conocimientos procedimentales, epistémicos y de contenidos para presentar hipótesis explicativas de fenómenos, hechos y procesos científicos nuevos, o bien para hacer predicciones. Al interpretar datos y pruebas, son capaces de diferenciar la información relevante de la irrelevante y recurrir a conocimientos externos al programa educativo convencional. Pueden distinguir los argumentos que se basan en pruebas y teorías científicas de aquellos basados en otras consideraciones. Los estudiantes de nivel 6 pueden evaluar diseños enfrentados de experimentos complejos, estudios de campo o simulaciones, y justificar sus elecciones.

Nivel	Características Lectura	Características Matemáticas	Características Ciencias
5	Las tareas en este nivel que implican la recuperación de información requieren que el lector localice y organice varios fragmentos de información profundamente incrustada, deduciendo qué información del texto es relevante. Las tareas reflexivas requieren una evaluación crítica o una hipótesis, sobre la base de un conocimiento especializado. Las tareas de interpretación y las de reflexión requieren una comprensión completa y detallada de un texto cuyo contenido o forma es desconocido. Para todos los procesos de la lectura, las tareas en este nivel suelen implicar tratar con conceptos que son contrarios a las expectativas.	los alumnos saben desarrollar modelos y trabajar con ellos en situaciones complejas, identificando los condicionantes y especificando los supuestos. Pueden seleccionar, comparar y evaluar estrategias adecuadas de solución de problemas para abordar problemas complejos relativos a estos modelos. Los alumnos pertenecientes a este nivel pueden trabajar estratégicamente utilizando habilidades de pensamiento y razonamiento bien desarrolladas, así como representaciones adecuadamente relacionadas, caracterizaciones simbólicas y formales, e intuiciones relativas a estas situaciones. Pueden reflexionar sobre sus acciones y formular y comunicar sus interpretaciones y razonamientos	los estudiantes pueden utilizar ideas o conceptos científicos abstractos para explicar fenómenos, hechos y procesos más complejos que les son desconocidos y que incluyen numerosas relaciones causales. Son capaces de aplicar conocimientos epistémicos más sofisticados para evaluar diseños alternativos y experimentales, así como justificar sus elecciones y usar los conocimientos teóricos para interpretar la información o hacer predicciones. Los estudiantes de nivel 5 pueden evaluar los modos de explorar una cuestión dada de manera científica e identificar las limitaciones de la interpretación de los conjuntos de datos, incluyendo las fuentes y los efectos de la incertidumbre en los datos científicos.

Nivel	Características Lectura	Características Matemáticas	Características Ciencias
4	Las tareas en este nivel que implican recuperar información requieren que el lector localice y organice varios fragmentos de información incrustada. Algunas tareas en este nivel requieren interpretar el significado de matices del lenguaje en una sección de texto, teniendo en cuenta el texto en su conjunto. Otras tareas interpretativas requieren comprender y aplicar categorías en un contexto desconocido. Las tareas reflexivas en este nivel requieren que los lectores utilicen el conocimiento formal o público para emitir hipótesis sobre o evaluar críticamente un texto. Los lectores deben demostrar una comprensión exacta de los textos largos o complejos cuyo contenido o forma puede ser desconocido.	Los alumnos pueden trabajar con eficacia con modelos explícitos en situaciones complejas y concretas que pueden conllevar condicionantes o exigir la formulación de supuestos. Pueden seleccionar e integrar diferentes representaciones, incluidas las simbólicas, asociándolas directamente a situaciones del mundo real. Los alumnos de este nivel saben utilizar su gama limitada de habilidades y razonar con cierta perspicacia en contextos sencillos. Pueden elaborar y comunicar explicaciones y argumentos basados en sus interpretaciones, argumentos y acciones.	los estudiantes pueden utilizar conocimientos de contenidos más complejos o abstractos, que pueden dárseles o ser recordados, para elaborar explicaciones de hechos y procesos más complejos o menos conocidos. Pueden ejecutar experimentos que incluyan dos o más variables independientes en un contexto limitado. Son capaces de justificar un diseño experimental recurriendo a elementos del conocimiento procedimental y epistémico. Los estudiantes de nivel 4 pueden interpretar los datos obtenidos de un conjunto de datos moderadamente complejos o de un contexto menos familiar, extraer conclusiones apropiadas que vayan más allá de los datos y justificar sus elecciones.

Nivel	Características Lectura	Características Matemáticas	Características Ciencias
3	Las tareas en este nivel requieren que el lector localice, y en algunos casos reconozca la relación entre varios fragmentos de información que deben cumplir varias condiciones. Las tareas de interpretación de este nivel requieren que el lector integre varias partes de un texto con el fin de identificar una idea principal, comprender una relación o interpretar el significado de una palabra o frase. Tienen que tener en cuenta muchas características al comparar, contrastar o categorizar. A menudo, la información requerida no es prominente o hay mucha información de carácter similar; o bien hay otros obstáculos en el texto, como ideas que son contrarias a lo esperado o negativamente redactadas. Las tareas reflexivas en este nivel pueden requerir conexiones, comparaciones y explicaciones, o pueden requerir que el lector evalúe una característica del texto. Algunas tareas reflexivas requieren que los lectores demuestren una buena comprensión del texto en relación con el conocimiento familiar, de cada día. Otras tareas no requieren la comprensión de textos detallados, pero requieren que el lector recurra al conocimiento menos común.	Los alumnos saben ejecutar procedimientos descritos con claridad, incluyendo aquellos que requieren decisiones secuenciales. Sus interpretaciones son lo bastante sólidas para fundamentar la creación de un modelo sencillo o para seleccionar y aplicar estrategias de solución de problemas sencillos. Los alumnos de este nivel saben interpretar y utilizar representaciones basadas en diferentes fuentes de información y razonar directamente a partir de ellas. Muestran cierta capacidad para manejar porcentajes, fracciones y números decimales, así como para trabajar con relaciones proporcionales. Sus soluciones reflejan que pueden desarrollar una interpretación y un razonamiento básicos.	Los estudiantes pueden recurrir a conocimientos de contenido moderadamente complejos para identificar o elaborar explicaciones para fenómenos conocidos. En situaciones menos conocidas o más complejas, pueden elaborar explicaciones si se les da la base o la ayuda correspondiente. Pueden recurrir a elementos de sus conocimientos procedimentales o epistémicos para realizar un experimento simple en un contexto limitado. Los estudiantes de nivel 3 son capaces de distinguir las cuestiones científicas de las no científicas e identificar las pruebas que respaldan una afirmación científica.

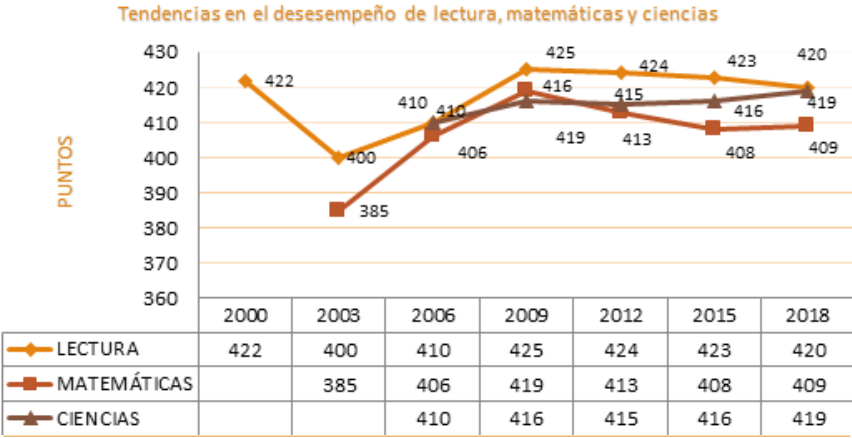
Nivel	Características Lectura	Características Matemáticas	Características Ciencias
2	Algunas tareas en este nivel requieren que el lector localice uno o más fragmentos de información, que pueden necesitar ser deducidos y pueden tener que cumplir una serie de condiciones. Otras requieren el reconocimiento de la idea principal de un texto, la comprensión de las relaciones, o interpretar su significado dentro de una parte limitada del texto cuando la información no es prominente y el lector debe hacer inferencias de bajo nivel. Las tareas en este nivel pueden incluir comparaciones o contrastes en base a una sola característica en el texto. Las tareas reflexivas típicas en este nivel requieren que los lectores hagan una comparación o varias conexiones entre el texto y el conocimiento exterior, y hagan uso de la experiencia y las actitudes personales.	Los alumnos saben interpretar y reconocer situaciones en contextos que solo requieren una inferencia directa. Los alumnos de este nivel pueden extraer información de una única fuente y usar un único modo de representación. Los estudiantes pueden utilizar algoritmos básicos, fórmulas, procedimientos o convenciones para resolver problemas que contengan números enteros. Son capaces de hacer interpretaciones literales de los resultados.	Los estudiantes son capaces de recurrir a conocimientos del día a día y a conocimientos procedimentales básicos para identificar una explicación científica adecuada, interpretar datos e identificar la cuestión de que se trata en un diseño experimental simple. Pueden utilizar conocimientos científicos básicos o del día a día para identificar una conclusión válida de un conjunto de datos simple. Los estudiantes de nivel 2 muestran conocimientos epistémicos básicos al poder identificar cuestiones que podrían investigarse científicamente.
1		Los estudiantes son capaces de contestar preguntas que impliquen contextos familiares donde toda la información relevante esté presente y las preguntas estén claramente definidas. Son capaces de identificar información y desarrollar procedimientos rutinarios conforme a instrucciones directas en situaciones explícitas. Pueden llevar a cabo acciones que sean obvias y seguirlas inmediatamente a partir de un estímulo.	Los estudiantes tienen un conocimiento científico tan limitado que sólo se puede aplicar a pocas situaciones que conocen. Dan explicaciones científicas obvias y parten de evidencia explícita.

Nivel	Características Lectura	Características Matemáticas	Características Ciencias
1a	Las tareas en este nivel requieren que el lector localice una o más piezas independientes de información explícita; para reconocer el tema principal o el propósito del autor en un texto sobre un tema conocido, o para hacer una conexión simple entre la información del texto y el conocimiento común, de todos los días. Normalmente, la información requerida en el texto es prominente y hay poca, o ninguna, información de carácter similar. El lector es explícitamente dirigido a considerar los factores relevantes en la tarea y en el texto.		
1b	Las tareas en este nivel requieren que el lector busque un único fragmento de información explícita en una posición prominente en un texto breve y sintácticamente simple con un contexto y tipo de texto familiar, como una narración o una simple lista. El texto normalmente proporciona apoyo al lector, como la repetición de la información, imágenes o símbolos conocidos. Hay poca información de carácter similar. En las tareas que requieren interpretación el lector puede tener que realizar conexiones simples entre piezas adyacentes de información.		

Nivel	Características Lectura	Características Matemáticas	Características Ciencias
1c	En este nivel, las tareas exigen al lector comprender el significado de palabras escritas individualmente y frases cortas. En las tareas se pide a los estudiantes que encuentren palabras o frases en una lista o un texto breves, que reconozcan la forma impresa de objetos y conceptos comunes o que extraigan el significado literal de frases sueltas y pasajes muy breves y sintácticamente sencillos en contextos familiares. Los textos ayudan a los estudiantes con indicadores explícitos, con repeticiones, imágenes o símbolos familiares y limitando los datos similares.		
Por debajo del nivel 1		Se trata de estudiantes que no son capaces de realizar las tareas de matemáticas más elementales que pide PISA.	Se trata de estudiantes que no son capaces de realizar las tareas de ciencias más elementales que pide PISA.

Fuente: OCDE (2018).

La prueba PISA no sólo evalúa los conocimientos y las habilidades de los estudiantes, también aplica una serie de cuestionarios a diferentes actores de la educación, con la finalidad de contextualizar los resultados de las pruebas y evitar interpretaciones reduccionistas (OCDE, 2017). Los aprendizajes alcanzados en PISA a lo largo de estos 18 años de aplicación en México en las tres áreas evaluadas (es importante referir que México ha participado en todas las pruebas), se muestran en la gráfica 3. Es importante referir que en este momento se presentan solo los puntaje- promedio alcanzados en cada una de las áreas, pero en líneas siguientes los resultados se contextualizaran tomando como elementos los canales de comunicación o los dispositivos que se utilizaron en la administración de gobierno (2012-2018) para la mejora de los aprendizajes.



Gráfica 3. Tendencias en el desempeño de lectura, matemáticas y ciencias.
Fuente: OCDE (2019).

250

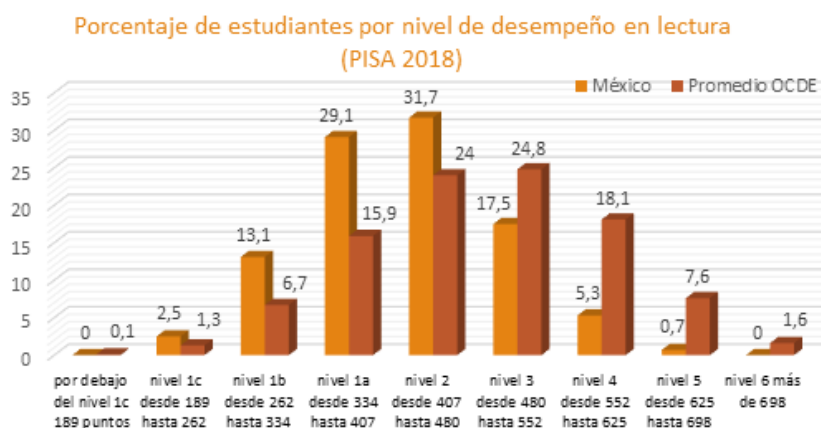
Hasta el momento, lo que se puede percibir de acuerdo a los indicadores, es que los resultados de aprendizaje en las tres áreas evaluadas por PISA (lectura, matemáticas y ciencias) no existen diferencias importantes ni para mejorar ni posicionarse mejor. La lectura que se tiene de México ante estos resultados, es que los logros de aprendizaje son bajos. Los indicadores muestran que los puntajes obtenidos en 2003 son significativamente inferiores (lectura y matemáticas) con respecto a los puntos obtenidos en 2018. Cabe señalar que en 2003 hubo una participación de 41 países. Con una tendencia al incremento, en 2018 participaron 79 países. La tendencia de los resultados en la aplicación de la prueba PISA en los otros años, no muestra una diferencia importante en relación a la aplicación de 2018 (OCDE, 2019). Sin embargo, las interpretaciones de los resultados PISA han sido diferentes a lo largo de los años; por ejemplo, los resultados PISA 2003 eran una oportunidad para las políticas públicas (Muñoz, 2005; Rodríguez, 2005), a transitar a: en México no hay avances ni en lectura, matemáticas y ciencias (Poy, 2016) hasta referir, con PISA 2018, que ninguna administración de gobierno es la responsable de los resultados (Backhoff, 2019).

De acuerdo a los resultados en las tres áreas evaluadas (lenguaje, matemáticas y ciencias) en PISA 2018 y en comparación con el promedio OCDE (OCDE, 2019), un porcentaje muy bajo alcanza los niveles más altos (5 y 6). En el caso de lectura se tiene 0.7% en el nivel 5 y nulo en el nivel 6 (véase gráfica 4). En matemáticas tenemos 0.5% en el nivel 5 y nulo en el nivel 6 (véase gráfica 5). Para ciencias hay 0.3 % en el nivel 5 y nulo en el nivel 6 (véase gráfica 6).

a) Resultados de PISA 2018 en Lectura

En cuanto a los niveles de desempeño más bajos, se puede observar que en el área de lectura, el 60% de estudiantes se concentró entre el nivel 1 y nivel 2 (gráfica 4). Con ello, se dice que los alumnos son capaces de identificar la idea principal de un texto, reconoce el propósito del autor en un texto, reflexiona sobre la forma de los textos cuando se les indica que lo hagan, tiene la capacidad de encontrar información en los textos. Y casi el 1% (nivel 5) puede discriminar entre un hecho y una opinión, diferenciar conceptos abstractos y contra intuitivos y comprender textos largos (OCDE, 2019).

En el nivel 3 y 4 de lectura concentra alrededor de 21% de estudiantes (gráfica 4). Lo anterior nos señala que los evaluados y evaluadas pueden ubicar fragmentos múltiples de información, vincular distintas partes de un texto, relacionar los textos con situaciones cotidianas, interpretar información y significados a partir del uso de lenguaje, evaluar críticamente un texto (OCDE, 2019).



Gráfica 4. Porcentaje de estudiantes por nivel de desempeño en lectura.

Fuente: OCDE (2019).

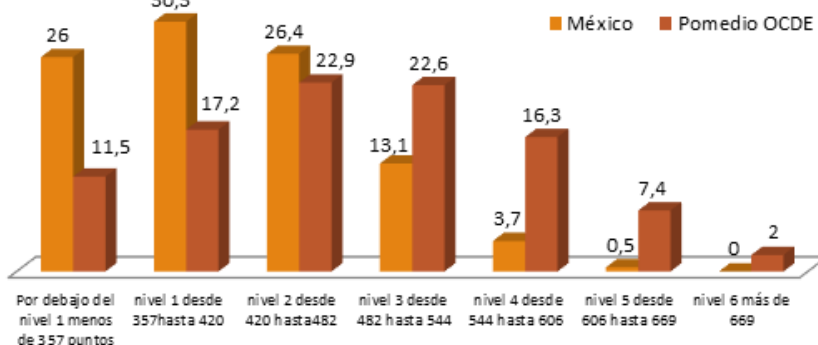
b) Resultados de PISA 2018 en matemáticas

En matemáticas 26% de los evaluados y evaluadas están por debajo del nivel 1 y el 56% se concentró entre el nivel 1 y nivel 2 (gráfica 5). Lo que implica que los y las alumnas (que alcanzan nivel 2) pueden reconocer e interpretar situaciones en contextos que requieren la inferencia directa, pueden identificar la información y realizar procedimientos rutinarios, hacen uso de un modo específico de elaborar las representaciones. El 0.5%

de estudiantes (nivel 5) puede seleccionar, comparar y evaluar estrategias pertinentes de solución para tratar un problema complejo, puede razonar sus acciones y comunicar sus argumentos (OCDE, 2019).

En matemáticas el nivel 3 y 4 concentra alrededor del 17% de estudiantes (gráfica 5). En estos niveles los estudiantes son capaces de ejecutar procedimientos descritos claramente, seleccionan y aplican estrategias simples para solucionar un problema, pueden interpretar y usar representaciones basadas en diferentes fuentes de información, pueden explicar y construir argumentos, pueden razonar flexiblemente con cierta comprensión (OCDE, 2019).

Porcentaje de estudiantes por nivel de desempeño en matemáticas (PISA 2018)



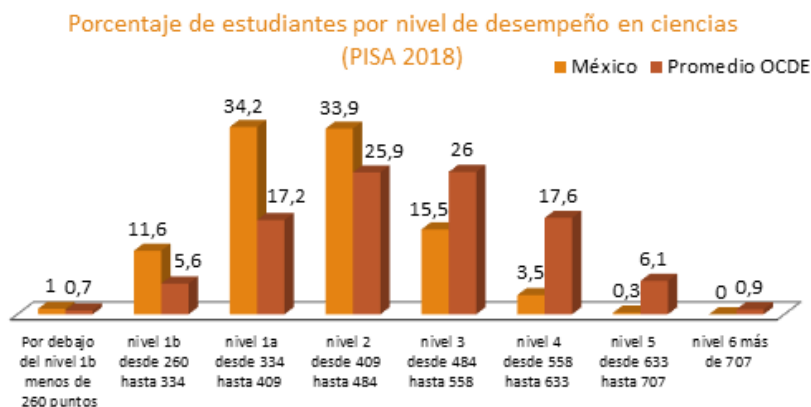
Gráfica 5. Porcentaje de estudiantes por nivel de desempeño en matemáticas. Fuente: OCDE, (2019).

c) Resultados de PISA 2018 en ciencias

En ciencias, alrededor del 79% de los estudiantes se concentró entre el nivel 1 y nivel 2 (gráfica 6), lo que implica que los evaluados y evaluadas pueden dar explicaciones científicas simples, partir de la evidencia para explicar situaciones, interpretar literalmente los resultados de una investigación científica, obtener conclusiones basadas en investigaciones sencillas. En contraste, solo el 0.3% (nivel 5) de los estudiantes puede aplicar los conceptos y el conocimiento de la ciencia a situaciones complejas, pueden comparar, evaluar y explicar la evidencia científica de forma creativa y autónoma, construyen explicaciones basadas en la evidencia científica (OCDE, 2019).

En ciencias el nivel 3 y 4 concentra al 19 % de los estudiantes (gráfica 6). En estos niveles los alumnos pueden seleccionar hechos y conocimientos para explicar fenómenos, hacen deducciones del papel de la ciencia y la tecnología, aplican estrategias sencillas de investigación,

comunican sus decisiones utilizando el conocimiento y la evidencia de la ciencia, seleccionan e integran a sus explicaciones la mirada de diferentes disciplinas (OCDE, 2019).



Gráfica 6. Porcentaje de estudiantes por nivel de desempeño en ciencias.
Fuente: OCDE, (2019).

Una vez que se tiene una caracterización de los resultados de la prueba PISA en lectura, matemáticas y ciencia, se pasará al siguiente apartado, el cual relacionan estos indicadores con la política de evaluación al ingreso al servicio educativo durante la administración de gobierno (2012-2018).

253

5. La política de ingreso al servicio educativo y su relación con el nivel de rendimiento en la prueba PISA

La reforma educativa de 2013 ponderó diversas acciones con miras hacia la calidad educativa, en la constitución política de los Estados Unidos Mexicanos así quedó plasmado:

El estado garantizará la calidad en la educación obligatoria de manera que los materiales y métodos educativos, la organización escolar, la infraestructura educativa y la idoneidad de los docentes y los directivos garanticen el máximo logro de aprendizaje de los educandos (DOF, 2013, p. 1).

En septiembre de ese mismo año se promulgó la Ley General del Servicio Profesional Docente (LGSPD), con la finalidad de garantizar la evaluación docente y cumplir con el reto de la calidad de la educación. La política de la evaluación se estableció como un canal de comunicación que pretendía cumplir con las necesidades de ese momento: la calidad de la

educación a partir de captar a los mejores docentes y mejorar el logro de los aprendizajes de los educandos. La LGSPD regula cuatro aspectos en el ámbito laboral de los docentes de educación básica: el ingreso, la promoción, la permanencia y el reconocimiento, lo que se le llama carrera docente (Murillo, 2007; Cuenca, 2015; Cordero y González 2016; Cuevas y Rangel 2019).

La política de ingreso a la docencia con la reforma educativa de 2013 dibujó, en apariencia, un ingreso al servicio educativo más sistemático y transparente¹. En este sentido el sistema de evaluación como canal de comunicación se convirtió en un sistema de regulación dirigido a resultados específicos. La evaluación se convirtió en el mecanismo para intentar captar a los mejores docentes que desean ingresar al sistema educativo nacional y terminar con la corrupción en la asignación de plazas docentes (Ornelas, 2012; INEE, 2016). La política de la meritocracia se consolidó como la única vía para ingresar a la docencia en México (así al menos lo referían las autoridades educativas y el ex presidente de México) y en otras regiones del mundo, dejando a un lado aquella que pondera la antigüedad en el servicio educativo y la acumulación de diplomas y certificaciones (Cuenca, 2015; Murillo 2007; Vaillant, 2009; y Cuevas y Rangel, 2019).

254

La LGSPD estableció en el Marco General de la Calidad de la Educación los perfiles, parámetros e indicadores que iban a servir como referentes en los concursos de oposición para la evaluación docente (LGSPD, 2013). Los perfiles son definiciones a priori de lo que se considera un buen desempeño docente (Cuenca, 2015).

El ingreso al Servicio Profesional Docente (SPD) fue un concurso de oposición abierta, mediante el cual se define si un docente es o no idóneo para cumplir la función docente. Al candidato se le aplicaban dos instrumentos de evaluación: 1) Examen nacional de conocimientos y habilidades para la práctica docente y 2) Examen nacional de habilidades intelectuales y responsabilidades ético-profesionales.

Es importante referir que la profesionalización de la carrera docente no sólo tiende a mejorar el perfil de los que se integran al servicio educativo, también incide en el aprendizaje de los alumnos. En ese sentido, un buen docente está asociado a las siguientes características: 1) el nivel de conocimientos, 2) la capacidad para transmitir esos conocimientos y 3) la motivación y compromiso para su labor (Darling-Hammond, 2001; De Hoyos y Estrada, 2018). En las siguientes líneas se va presentar una relación

¹ Antes de la reforma del 2013 el ingreso a la docencia era a través de: a) asignación de plaza inicial, b) la selección libre de personas y c) propuestas del sindicato (Montaño, 2005).

posible entre los efectos de la LGSPD sobre el perfil de los nuevos docentes de secundaria, para ello se tomará en cuenta el nivel de conocimientos en relación al lenguaje, matemáticas y ciencias. Se hace de esta manera porque la evidencia internacional identifica una relación importante entre los conocimientos de los docentes en las áreas señaladas y los aprendizajes de los alumnos en éstas áreas (Hanushek, Piopiunik y Wiederhold; 2018; De Hoyos y Estrada, 2018).

Los resultados de las evaluaciones del perfil de ingreso al servicio educativo se clasifican en tres niveles de desempeño (I, II, III) que refieren lo que los sustentantes son capaces de hacer y que los ubica en un grupo de desempeño en particular (A, B, C). Cabe señalar que los grupos de desempeño son para los sustentantes con evaluación idónea, los que tienen evaluación no idónea, sólo así son nombrados. La tabla 2 nos señala las descripciones de cada nivel y el del desempeño del grupo:

Tabla 2. Descripción de nivel y desempeño del grupo.

Nivel	Descripción	Grupo de desempeño	Descripción
I	Dominio insuficiente de los conocimientos y habilidades que se consideran indispensables para un adecuado desempeño docente	A	En los dos exámenes obtuvo el nivel de desempeño III.
II	Dominio suficiente y organizado de los conocimientos y habilidades que se juzgan indispensables para un adecuado desempeño docente	B	En un examen obtuvo el nivel de desempeño III y en el otro el nivel de desempeño II.
III	El sustentante, además de mostrar un dominio suficiente y organizado de los conocimientos y habilidades contemplados en el instrumento, demuestra una amplia capacidad para utilizarlos en una diversidad de situaciones didácticas.	C	En los dos exámenes obtuvo el nivel de desempeño II.

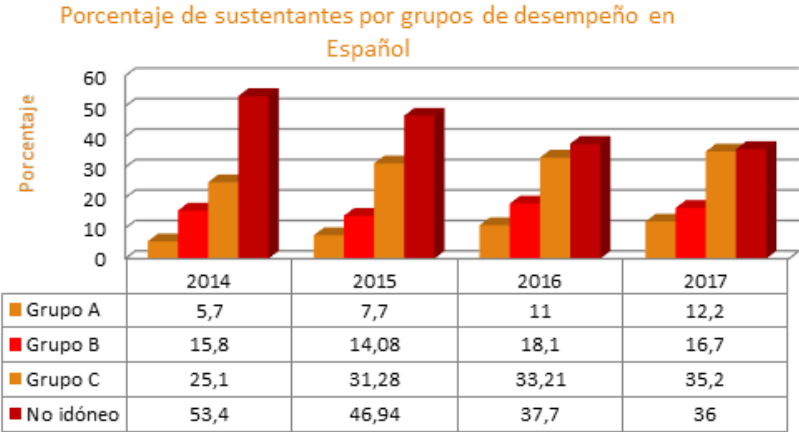
Fuente: SPD (2014).

Una vez que se tiene la descripción de los niveles y grupos de desempeño. Se presentan las siguientes gráficas que tiene que ver con los resultados de evaluación que obtuvieron los sustentantes a ingresar al servicio educativo en la función de docente a nivel secundaria en las asignaturas de la enseñanza del Español (gráfica 7), Matemáticas (gráfica 8) y Ciencias que considera Biología (B), Física (F) y Química (Q) (gráfica 9).

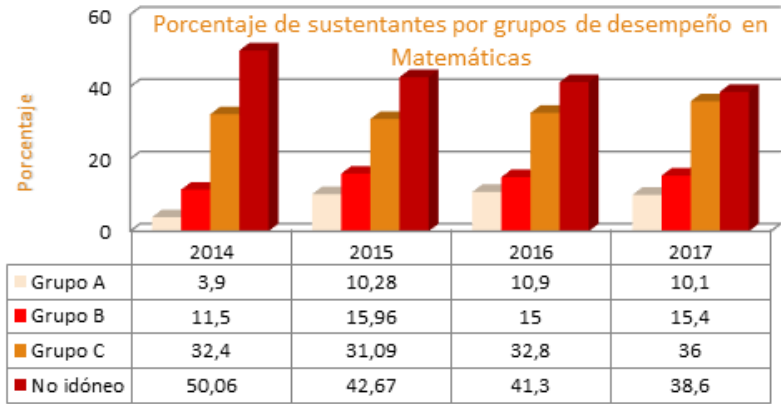
Lo que las gráficas presentan en relación al desempeño que lo sustentantes obtienen en las pruebas para ingresar al servicio educativo es una mejora gradual, quizá mínima, pero importante, en el sentido que el servicio profesional como dispositivo atrae a docentes con mejores conocimientos y habilidades para el ejercicio de la docencia, según las pruebas estandarizadas.

Por ejemplo, en español, el grupo A incrementó el porcentaje de sustentantes de 2014 a 2017 en un 6.5%. En matemáticas, el mismo grupo y en el mismo periodo incrementó 6.2%. En ciencias, en el caso de biología, se tuvo un incremento del 5.9%, en física un incremento del 7.4% y en química un incremento del 2.9%. Cabe señalar que en el área de ciencias hubo un decremento en el 2017 en el porcentaje de sustentantes en relación al constante incremento de los años anteriores.

256

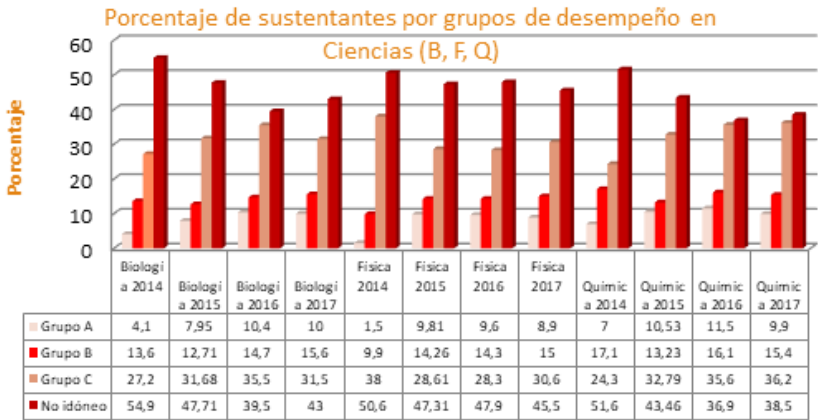


Gráfica 7. Porcentaje de sustentantes por grupos de desempeño en español.
Fuente: SPD (2014; 2015; 2016; 2017).



Gráfica 8. Porcentaje de sustentantes por grupos de desempeño en matemáticas.
Fuente: SPD (2014; 2015; 2016; 2017).

Los resultados descritos se pueden relacionar con los resultados de PISA 2015 y 2018, puesto que es el periodo cuando ingresan al servicio educativo los docentes evaluados. Por ejemplo, en el área de lectura, se presentó un crecimiento del 0.5% de estudiantes que logran ubicarse en el nivel 5 o superior. En matemáticas y ciencias hay un crecimiento del 0.2% (tabla 3).



Gráfica 9. Porcentaje de sustentantes por grupos de desempeño en ciencias.
Fuente: SPD (2014; 2015; 2016; 2017).

Tabla 3. Porcentaje de alumnos de alto rendimiento.

Competencia evaluada	Nivel 5 o superior PISA 2015 (%)	Nivel 5 o superior PISA 2018 (%)
Lectura	0.3	0.8
Matemáticas	0.3	0.5
Ciencias	0.1	0.3

Fuente: OCDE (2019).

Así como se presentó un incremento en el nivel de rendimiento más alto tanto en el desempeño de los docentes de nuevo ingreso al servicio educativo como de los alumnos que presentan la prueba PISA. También hay una disminución en porcentaje en el rendimiento más bajo, excepto en los resultados de lectura que se obtienen por los alumnos (tabla 4).

Por ejemplo, el porcentaje de alumnos que obtienen un nivel de rendimiento por debajo del nivel 2 en lectura incrementó en 3% en PISA 2018 en comparación con 2015. En la competencia de matemáticas se presentó

un decremento del 0.4% y en ciencias disminuyó en 1%. Es decir, un menor número de estudiantes se ubicó en los niveles más bajos de rendimiento en matemáticas y ciencias.

Tabla 4. Porcentaje de alumnos de bajo rendimiento.

Competencia evaluada	Por debajo del nivel 2 PISA 2015 (%)	Por debajo del nivel 2 PISA 2018 (%)
Lectura	41.7	44.7
Matemáticas	56.6	56.2
Ciencias	47.8	46.8

Fuente: OCDE (2019).

258

Si bien faltarían factores para poder determinar una relación más sistemática entre los resultados que los docentes de nuevo ingreso al servicio docente con los logros de aprendizaje que los alumnos alcanzan en los resultados PISA, es un hecho que el sistema educativo mexicano a partir de la evaluación docente, presentó una tendencia hacia la mejora en los resultados en las pruebas estandarizadas de docentes como de alumnos en las competencias de lectura, matemáticas y ciencias. Incluso, el incremento gradual mínimo en los logros de aprendizajes de docentes y alumnos, puede estar asociado, precisamente, con los criterios de ingreso que se establecen a partir de la creación del SPD. Lo que se puede advertir es que el SDP fue un dispositivo con impacto en la tendencia a la mejora de los resultados y que posiblemente podía ser un camino importante en aras a la calidad educativa. La política de la evaluación que se implementó en México presenta una tendencia positiva en la calidad educativa de la nación (Estrada, 2019).

A pesar de que la política de evaluación fue una iniciativa importante para atraer y captar a los mejores docentes, hubo ciertas cuestiones que probablemente influyeron en que la evaluación docente como canal de comunicación y el SPD como dispositivo no se desarrollara en las mejores condiciones:

1. El ingreso al servicio docente siguió operando a través de usos y costumbres diferentes al mérito (Backhoff, 2017). Es importante mencionar que el sindicato, aun con acotamiento de poder y maniobra, sigue en estas prácticas. El hecho de que el gobierno haya centralizado ciertas decisiones, no implica que la actuación del sindicato se haya aniquilado; posiblemente, tomó otro lugar y se valió de diferentes mecanismos para hacerse presente.
2. La falta de información acerca de los docentes del país y las condiciones en las que trabajan no se conocen con claridad. El Sistema de Información y de Gestión Educativa (SIGED) no cuenta de la información clara y sistemática de los docentes,

no se logró captar la información del 15.7% del personal (México Evalúa, 2014). Aunque la tarea se le encomendó no pudo recabar los datos, todavía hay incertidumbre y un sistema de información que no se acaba de completar, que tiene problemas, o que por lo menos, ello parece manifestar. Al no existir datos transparentes en torno a los docentes, era muy probable y factible que el número de plazas u horas que ofrecían las diversas entidades para los aspirantes tuvieran ciertas variaciones. Es decir, las autoridades pueden ofertar cierto número de plazas u horas que no corresponde con las vacantes que existen en los recintos escolares; habría que revirar a algunas convocatorias que se emiten de las entidades, a través de la página del Servicio Profesional Docente (SPD). Hay infinidad de centros escolares que no cuentan con la plantilla docente completa a pesar de que hay una lista de docentes que, de acuerdo al mérito, están en derecho de estarlo. Al no haber información sistemática, precisa y transparente, se posibilita el flujo de los mecanismos discrecionales (Fernández, 2018).

3. A pesar de que la LGSPD establece el sistema de tutorías para los docentes de nuevo ingreso durante sus primeros dos años de docencia, con la finalidad de que el asesor pudiera apoyar al docente con sus fortalezas y debilidades en el ejercicio de la docencia. Sin embargo, la cobertura del sistema de tutores fue insuficiente. Según el INEE sólo el 58% de los docentes de nuevo ingreso habían recibido tutorías en el ciclo escolar 2014-2015 (INEE, 2017). El problema de la asignación de tutores estuvo relacionado a falta de presupuesto, comunicación y coordinación entre autoridades federales y estatales (Hernández, 2015) y porque las evaluaciones de los docentes de nuevo ingreso llegaban a destiempo, lo que dificultó la planeación de las tutorías (Fernández, 2018.)
4. Con la reforma educativa que se impulsó en 2013, la precarización salarial es notable, y es más “visible” por el poco número de horas que se les asigna a los docentes normalistas de nuevo ingreso, lo que alimenta las condiciones laborales inciertas (Compañ, 2018). Los docentes reciben un número de horas inferior a lo que se les otorgaba anteriormente. En la década de los noventa se les otorgaba a los egresados de las escuelas normales un mínimo de 19 horas, lo que les permitía tener un salario más elevado. Actualmente, hay docentes con 5 horas y un salario insuficiente. El número de horas que se les otorga a los docentes es variable; depende de la especialidad y el número

de grupos existentes en el centro escolar, lo que repercute en su ejercicio docente (Compañ, 2018)

La política de evaluación docente fue un acierto importante de la administración de gobierno, que se asoció, según la evidencia, con los logros de aprendizaje de los estudiantes. Sin embargo, las condiciones laborales de los docentes no mejoraron. Se captó a los mejores docentes, pero no se asoció ello con incentivos salariales o prestaciones favorables. Lo que se tradujo en contenido importante pero un proceso opaco.

6. Comentarios finales

Si bien, la política de evaluación al ingreso al servicio educativo como canal de comunicación y la creación del SPD como dispositivo en la administración de gobierno anterior, no resolvió los rezagos en materia educativa, sí sentó las bases que posibilitaran una transformación en el sistema educativo nacional en aras a la calidad educativa.

260

La política de la evaluación docente no sólo se estableció como un mecanismo para acabar con la corrupción en la asignación de plazas docentes, incluso, no se logró en su totalidad, sino como un mecanismo de regulación que pudiera captar a los mejores docentes para incorporarlos al servicio educativo y de esa manera elevar el nivel de rendimiento de los alumnos. A partir de la creación del SPD se percibe un incremento gradual mínimo en el rendimiento de los docentes de educación secundaria, que ingresan al sistema educativo nacional en el área de español, matemáticas y ciencias y, un avance en los logros de aprendizaje en las competencias evaluadas en PISA.

El hecho de que un canal de comunicación con impacto sustantivo, como la evaluación al ingreso a la docencia, no haya tenido mayor éxito y quizá más impacto en los logros de aprendizaje, se debe a que el dispositivo, en este caso el SPD, no pudo tener un control sistemático de ciertos procesos y resultados, cuestión importante en la implementación de un dispositivo (Rivas, 2015). La falta de coordinación y comunicación entre autoridades federales estatales, las limitaciones del presupuesto para la asignación de tutores que coadyuvaran a la formación de los docentes de nuevo ingreso, la falta de planeación para el trabajo con los tutorados y la participación de actores como la CNTE, que se resistieron y se opusieron a los procesos de evaluación, fueron elementos que no ayudaron a que la evaluación al ingreso a la docencia se pudiera sentar sobre condiciones más favorables y sólidas.

Asegurar la calidad de la educación en términos de captar a los mejores docentes y mejorar los logros de aprendizaje de los alumnos requiere: 1) un presupuesto destinado para la formación de los docentes de nuevo ingreso, 2) realizar estudios más sistemáticos y con mayores elementos que permitan establecer la relación entre los logros de aprendizaje de los alumnos y la evaluación docente, 3) mejorar las condiciones laborales de los docentes, 4) mayor coordinación entre gobierno federal y gobierno estatal, 5) tener la información clara y precisa de los docentes. Se necesitan dispositivos de política educativa que se les dé seguimiento en términos de evaluarlos y focalizarlos. La política educativa tiene que gozar de cierta estabilidad, coherencia, coordinación y buscar el bienestar general, si de verdad se aspira a la calidad educativa. De lo contrario, la política educativa presentará contenidos interesantes pero se desarrollará bajo procesos opacos, irregulares e improvisados.

Lamentablemente, todo lo anterior, parece no estar en la agenda de la administración de gobierno actual. Por ejemplo, hasta ahora, la evaluación al ingreso a la docencia no se lleva a cabo, lo que puede implicar que la probabilidad de que los logros de aprendizaje mejoren se va más distante, además que lo que sí se mira próximo es el regreso a las prácticas clientelares de asignación de plazas docentes.

Referencias

- Backhoff, E. (8 de agosto de 2017). Terminó la injusta y opaca asignación de plazas para maestros. *La jornada*. Recuperado de <https://bit.ly/30iPqdo>
- Backhoff, E. (9 de diciembre de 2019). México en PISA 2018, *Nexos*. Recuperado de <https://bit.ly/2GfryR3>
- Backhoff, E; Bouzas, A. Contreras, C. Hernández, E. y García, M. (2007). *Factores escolares y aprendizaje en México*. México: INEE.
- Bernstein, B. (1998). *Pedagogía, control simbólico e identidad*. Morata: Madrid.
- Blanco, E. (2012). El reciclaje de la desigualdad: exclusiones educativas en América Latina. En Puchet, M. et al.(Eds.). *América Latina en los albores del siglo XX: 2. Aspectos sociales y políticos*. (pp. 63-94). México: FLACSO.
- CEPALSTAT (2020). Asistencia escolar de ambos sexos por quintiles de ingreso per cápita del hogar, según grupos de edad y área geográfica. *Estadísticas e indicadores sociales, CEPAL*. Recuperado en <https://bit.ly/3kTtTzQ>
- Compañ, J. (22 de agosto de 2018). La política de ingreso a la docencia: aciertos y desafíos. *Nexos*. Recuperado de <https://bit.ly/2GrvduL>

- Cordero, G. y González, C. (2016). Análisis del modelo de evaluación del desempeño docente en el marco de la reforma educativa Mexicana. *Archivos Analíticos de Políticas Educativas*. (24) pp.1-26.
- Cuenca, R. (2015). *Las carreras docentes en América Latina: la acción meritocrática para el desarrollo profesional*. Chile: OREALC-UNESCO.
- Cuevas, Y. y Rangel, K. (2019). Análisis de la carrera docente en la educación primaria en México: Entre el credencialismo y la meritocracia. *Archivos Analíticos de Políticas Educativas*. 27(44). <http://dx.doi.org/10.14507/epaa.27.4323>.
- Darling-Hammond, L. (2001). *El derecho de aprender. Crear buenas escuelas para todos*. España: Ariel.
- De Hoyos, R. y Estrada, R. (2018). ¿Los docentes mejoraron? ¡Sí!. *Nexos*. Recuperado de <https://bit.ly/30I9WKo>
- Diario Oficial de la Federación (DOF) (2013). *Decreto por el que se declara deformado el artículo 3 de la Constitución Política de los Estados Unidos Mexicanos*. Recuperado de <http://www.dof.gob.mx>
- Estrada, R. (2009) Reglas versus discreción en el servicio público: contratación de docentes en México. *Diario de economía laboral*, (3)7 (2), 545-579.
- Fernández, M. (2018). *Construir una educación inclusiva, equitativa y de calidad*. Ciudad de México: México Evalúa.
- Hanushek, Piopiunik y Wiederhold (2018). *The value of smarter teachers: international evidence on teacher cognitive skills and student performance*. NBER Working Paper Series 20727. Cambridge: MA.
- Hernández, L. (5 de septiembre de 2015). Destina SEP sólo 200 pesos a capacitación docente de maestros. *Excélsior*. Recuperado de <https://bit.ly/2Sanx2C>
- Hunter, I. (1998). *Repensar la escuela. Subjetividad, burocracia y crítica*. Barcelona: Pomares.
- Instituto Nacional para la Evaluación Educativa (INEE) (2016). *México en PISA 2015*. México: INEE.
- Instituto Nacional para la Evaluación de la Educación (INEE) (2017). *Las tutorías como estrategia de inserción a la docencia. La educación obligatoria en México*. Ciudad de México: INEE.
- Ley General para el Servicio Profesional Docente (2013). México: Gobierno de México.
- Martínez, F. (2012). El futuro de la educación, *Sinéctica*, 40, 1-11. Recuperado de <https://bit.ly/3ib114o>
- Martínez, F. (2016). Impacto de las pruebas en gran escala en contextos de débil tradición técnica: Experiencia de México y el Grupo Iberoamericano de PISA. *Revista Electrónica de Investigación y Evaluación Educativa*, 22(1). Recuperado de <https://bit.ly/3l2NM0q>
- Martínez, F. (1 de marzo de 2017). Hay realmente una catástrofe educativa en México, *Nexos*. Recuperado de <https://bit.ly/3cJ6u14>
- México Evalúa (2014). *Censo Educativo. Radiografía del dispendio presupuestal*. Ciudad de México: México Evalúa.

- Montaño, L. (2005). *La formación del docente actual de las escuelas secundarias generales en el Distrito Federal*. Caminos para la docencia. Tesis de Maestría en Desarrollo Educativo. México. Universidad Pedagógica Nacional.
- Muñoz, C. (2005). Análisis de los resultados de México en el PISA 2003: una oportunidad para las políticas públicas, *Perfiles latinoamericanos*, 13(26), 83-107.
- Muñoz, C. (2009). Construcción del conocimiento sobre la etiología del rezago educativo y sus implicaciones para la orientación de las políticas públicas: la experiencia de México. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación* (7) (4) 28-45.
- Murillo, J. (2007). *Evaluación del desempeño y carrera profesional docente. Un estudio comparado entre 50 países de América y Europa*. Chile: OREALC-UNESCO.
- Organización para la Cooperación y el Desarrollo Económico (OCDE) (2017). *El programa PISA de la OCDE. Qué es y para qué sirve*. París: OCDE. Recuperado de <https://bit.ly/3igNyYT>
- Organización para la Cooperación y el Desarrollo Económico (OCDE) (2018). Marco de Evaluación y de Análisis de PISA para el desarrollo. París: OCDE.
- Organización para la Cooperación y el Desarrollo Económico (OCDE) (2019). *Programa para la Evaluación Internacional de Alumnos PISA 2018 - Resultados*. París: OCDE.
- Ornelas, C. (2012). *Educación, colonización y rebeldía: La herencia del pacto Calderón-Gordillo*. México: Siglo XXI.
- Pedró, F. y Puig, I. (Coords) (1998). ¿Sirve para algo la Política Educativa? *En Las reformas educativas. Una perspectiva política y comparada*. Barcelona: Paidós.
- Poy, Laura (6 de diciembre de 2016). México, sin avances en matemáticas, lectura y ciencia en una década, *La Jornada*. Recuperado de <https://bit.ly/36laKTh>
- Rivas, A. (2015). *América Latina después de PISA: Lecciones aprendidas de la educación en siete países (2000-2015)*. Buenos Aires: CIPPEC-Natura-Instituto Natura.
- Rodríguez, R. (2005). "México en los resultados PISA 2003. Una interpretación no catastrófica". *Revista Mexicana de Investigación Educativa*, 24(10), 255-266.
- Rose, N. (1999). *Powers of Freedom: Reframing Political Thought*. Cambridge: Cambridge University Press.
- Santiago, P., McGregor, I., Nusche, D., Ravela, P. y Toledo, D., (2012) *Revisiones de la OCDE sobre la Evaluación en Educación México OCDE*. México: INEE SEP.
- Scarstacini, C. Spiller, P. y Stein, E. (2010). *El juego político en América Latina ¿Cómo se deciden las políticas educativas?* Colombia: BID.
- Secretaría de Educación Pública (2020). Reporte de indicadores educativos. *Indicadores y pronósticos educativos*. Recuperado de <https://bit.ly/3jguA5P>
- Vaillant, D. (2009). Políticas de inserción a la docencia en América Latina. La deuda pendiente. *Revista del currículum y formación del profesorado*, 1(13), 27-41.

