

# La idiotez y los debates sobre las evaluaciones internacionales a gran escala en las reformas educativas nacionales<sup>1</sup>

*Gustavo E. Fischman\**  
*Amelia Marcetti Topper\*\**  
*Iveta Silova\*\*\**

## Resumen

Nosotros realizamos una revisión exploratoria de la literatura de investigación y de políticas públicas relativa a las evaluaciones internacionales a gran escala (ILSA en inglés) e hicimos dos encuestas con expertos, creadores de políticas públicas y educadores para entender si, hasta qué punto y cómo tales evaluaciones influenciaron la producción de políticas públicas en el contexto nacional. Nuestra revisión de la literatura confirma que ILSA, con sus usos múltiples y ambiguos, funcionan cada vez más como soluciones idiotas en busca del problema cierto - o sea, parecen ser usadas como herramientas para legitimar reformas educativas. Con "idiotas" no nos referimos al uso común de la palabra, que denota a alguien que no es muy inteligente, sino al significado original de la palabra en griego antiguo, es decir, particular en oposición a perspectivas cívicas o públicas. Los resultados de nuestra encuesta mostraron una percepción creciente entre investigadores, creadores de políticas públicas y educadores, además de otros involucrados, de que las ILSA están teniendo un efecto en las políticas educativas nacionales, visto que más de un tercio (38%) de los entrevistados afirmaron que las ILSA fueron, en general, mal empleadas en los contextos de las políticas públicas nacionales. Sin embargo, mientras que la literatura sobre la ILSA indica que esas evaluaciones están teniendo alguna influencia, hay poca evidencia de que exista una relación causal positiva o negativa entre la participación de la ILSA y la implementación de reformas educativas. Tal vez el cambio más significativo asociado al uso de las ILSA en la literatura que revisamos sea la manera en que nuevas condiciones para la comparación educativa a nivel nacional, regional y global se hicieran posibles.

**Palabras-chave:** Evaluaciones internacionales a gran escala; Reforma educativa nacional; Medición global de aprendizaje.

\* Profesor del centro de formación de profesores Mary Lou Fulton, Arizona State University, Tempe, Arizona, Estados Unidos.

\*\* PhD in Education and Evaluation Policy, Arizona State University, Tempe, Arizona, Estados Unidos.

\*\*\* Profesora y directora del Centro de Estudios Avanzados en Educación Global del centro de formación de profesores Mary Lou Fulton, Arizona State University, Tempe, Arizona, Estados Unidos.

## Idiocy and debates about large-scale international assessments in national education reforms

### Abstract

We conducted an exploratory review of the research and policy literature on international large-scale assessments (ILSAs) and administered two surveys to experts, policymakers, and educators to understand whether, to what extent, and how have ILSAs influenced education policymaking at the national level. Our review of the literature confirms that ILSAs with their multiple and ambiguous uses, increasingly function as idiotic solutions in search for the right problem – that is, they appear to be used as tools to legitimate educational reforms. By “idiotic,” we are not referring to the common usage of somebody who is not very clever, but in the original meaning of the word in Ancient Greek as selfish and in opposition to civic or public oriented perspectives. Our survey results pointed to a growing perception among researchers, policymakers, educators, and other stakeholders that ILSAs are having an effect on national educational policies, with over one-third (38%) of respondents stating that ILSAs were generally misused in national policy contexts. However, while the ILSA literature indicates that ILSAs are having some influence, there is little evidence that any positive or negative causal relationship exists between ILSA participation and the implementation of education reforms. Perhaps the most significant change associated with the use of ILSAs in the literature we reviewed is the way in which new conditions for educational comparison at the national, regional and global level have been made possible.

**Keywords:** International large-scale assessments; National school reform; Global learning metrics.

## A idiotia e os debates sobre as avaliações internacionais em grande escala nas reformas educacionais nacionais

### Resumo

Nós realizamos uma revisão exploratória da literatura de pesquisa e de políticas públicas relativas às avaliações internacionais em grande escala (ILSAs em inglês) e fizemos duas enquetes com especialistas, criadores de políticas públicas e educadores para entender se, até que ponto e como tais avaliações influenciaram a produção de políticas públicas no contexto nacional. Nossa revisão da literatura confirma que as ILSAs, com seus usos múltiplos e ambíguos, funcionam cada vez mais como soluções idióticas em busca do problema certo – ou seja, elas parecem ser usadas como ferramentas para legitimar reformas educacionais. Com “idióticas” não estamos nos refe-

rindo ao uso comum da palavra, que denota alguém que não é muito inteligente, mas ao significado original da palavra em grego antigo, ou seja, particular em oposição a perspectivas cívicas ou públicas. Os resultados de nossa enquête mostraram uma percepção crescente entre pesquisadores, criadores de políticas públicas e educadores, além de outros envolvidos, de que as ILSAs estão tendo um efeito nas políticas educacionais nacionais, visto que mais de um terço (38%) dos entrevistados afirmaram que as ILSAs foram, de modo geral, mal aplicadas nos contextos das políticas públicas nacionais. Contudo, ao passo que a literatura sobre a ILSA indica que essas avaliações estão tendo alguma influência, há pouca evidência de que exista qualquer relação causal positiva ou negativa entre a participação da ILSA e a implementação de reformas educacionais. Talvez a mudança mais significativa associada com o uso das ILSAs na literatura que revisamos seja a maneira pela qual novas condições para a comparação educacional em nível nacional, regional e global se tornaram possíveis.

**Palavras-chaves:** Avaliações internacionais em grande escala; Reforma educacional nacional; Medição global de aprendizagem.

## Introducción

El uso de evaluaciones internacionales a gran escala (ILSA, por sus siglas en inglés) para medir el acceso al sistema educativo y el aprendizaje por parte de los estudiantes es un fenómeno mundial relativamente reciente (KIRSCH, LENNON, von DAVIER, GONZALEZ & YAMAMOTO, 2013; United Nations Educational, Cultural and Scientific Organization [UNESCO], 2016). Desde la evaluación piloto de doce países por parte del Instituto de la UNESCO para la Educación (IUE) en 1958, estas mediciones del aprendizaje estudiantil comparativas entre países han crecido en cantidad, del mismo modo que ha crecido el número de estudiantes incluidos, las habilidades evaluadas y los países participantes (KIRSCH et al., 2013). Con base en la intensa cobertura mediática que se da a la publicación de cada nueva serie de clasificaciones, las ILSA tienen el potencial de desempeñar un papel cada vez más decisivo en la conducción, apoyo y configuración de políticas educativas y de reformas promovidas por gobiernos de diferentes países a nivel nacional, regional y municipal (BAIRD et al., 2016; MEYER & BENAVIDOT, 2013).

Entre las ILSA más conocidas e influyentes se encuentran el Programa para la Evaluación Internacional de Alumnos (PISA) y la Encuesta de Escuelas Primarias del proyecto de Indicadores Mundiales de Educación (WEI-SPS) de la Organización para la Cooperación y el Desarrollo Económicos (OCDE) y el Estudio Internacional de Progreso en Comprensión Lectora (PIRLS), el Estudio Internacional de Tendencias en Matemáticas y Ciencias (TIMSS) y el Estudio Internacional de Educación Cívica y Ciudadana (ICCS) de la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA). Existen también ILSA para regiones específicas, tales como: el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), el Consorcio del África Meridional y Oriental para la Supervisión de la Calidad de la Educación (SAQMEC), el Programa de Análisis de los Sistemas Educativos de la Conferencia de Ministros de Educación de los Estados y Gobiernos de la Francofonía (PASEC), el Informe Anual del Estado de la Educación de la India

(ASER) o la Prueba Diagnóstica EGRA de Lectura Inicial del Instituto RTI International para los países de bajos ingresos. Consideradas en su conjunto, estas ILSA se han realizado, en diferentes momentos, en casi todos los países del mundo<sup>2</sup>, muchos de los cuales participan en varias evaluaciones a la vez<sup>3</sup> y muchos de los cuales también han ido aumentando su participación a lo largo de las dos últimas décadas.<sup>4,5</sup>

Para comprender de forma más exhaustiva y comparada el aprendizaje estudiantil, las organizaciones que realizan las evaluaciones a gran escala, junto a otras organizaciones que actúan en el escenario educativo global y nacional, han empezado a trabajar en el ambicioso objetivo de desarrollar métricas globales del aprendizaje (GLM, sigla que usaremos para referirnos a ellas a lo largo del texto) relacionadas a evaluaciones existentes. Las GLM proporcionan resultados de aprendizaje estandarizados, generalmente en las áreas de alfabetización y matemáticas, que pueden servir como parámetros para que los países midan su progreso a lo largo del tiempo. Algunas de las organizaciones con participación activa en la elaboración de las GLM son: el Instituto de Estadística de la UNESCO (UIS), el consorcio SAQMEC, el IEA, la OCDE, el Banco Mundial, la CONFEMEN (Conferencia de los Ministros de Educación de los Estados y Gobiernos de la Francofonía), el Centro ASER<sup>6</sup> y la Alianza Mundial por la Educación (GPE)<sup>7</sup>. La OCDE, por auspiciar el programa PISA, tal vez sea la organización más involucrada en el desarrollo de las GLM en la actualidad, ya que el programa PISA actualmente se lleva a cabo en 72 países del mundo, y también a nivel infranacional en países como Argentina, Canadá, China, la India y los Estados Unidos (OECD, n.d.).

En general, los esfuerzos existentes para desarrollar GLM han tendido a centrarse en el logro por parte de los estudiantes de ciertas metas, como la comprensión lectora en el tercer año de escolarización. Aunque las recomendaciones varían según las iniciativas y los grupos de trabajo, se están haciendo esfuerzos para fomentar la colaboración entre instituciones interesadas en el desarrollo de una lista más estandarizada de GLM (ver, entre otros, WINTHROP, ANDERSON & CRUZALEGUI, 2015). Por ejemplo, la Comisión Especial sobre Métricas del Aprendizaje (LMTEF, una iniciativa conjunta del UIS y el Centro para la Educación Universal de la Brookings Institution creada expresamente para ese objetivo) viene recomendando la medición del aprendizaje estudiantil en torno de siete dimensiones que van desde materias escolares al uso (como educación matemática, ciencias y tecnología) a participación cívica y bienestar general (como arte y cultura, bienestar social y emocional o bienestar físico).

El uso de las GLM — definidas como medidas para el “acceso más aprendizaje” (LMTEF, 2013, p. 10, énfasis en el original; véase también PERLMAN ROBINSON, 2011) — representa un alejamiento estratégico del foco tradicional de las organizaciones nacionales e internacionales en documentar las matrículas escolares. Este cambio de foco se cimentó a partir de la aprobación por parte de las Naciones Unidas del Marco para la Educación 2030, en mayo de 2015, y de los Objetivos de Desarrollo Sostenible, en septiembre del mismo año. Entre esos objetivos se incluía garantizar

una educación de calidad, es decir, velar por una enseñanza primaria y secundaria “que ha de ser gratuita, equitativa y de calidad y producir resultados escolares pertinentes y eficaces” (UN, 2016).

Mientras que la necesidad del acceso universal a la educación es consensual, la forma en que el aprendizaje debe medirse y la necesidad de establecer métricas globales del aprendizaje son objeto de calurosos debates.<sup>8</sup>

Es común que los debates sobre las ILSA, se conviertan en discusiones acaloradas, que provocan reacciones viscerales fuertes, que disminuyen las chances de poder escuchar análisis cuidadosos, examinar las evidencias existentes y de considerar las perspectivas contrastantes. Nuestro punto de partida es que no deberíamos sorprendernos por la polarización de tales debates ya que políticos, investigadores, docentes, administradores, estudiantes y sus familias suelen tener opiniones y perspectivas muy firmes sobre las escuelas, lo que debe arreglarse y cuál debería ser la "solución".

Uno de los argumentos más comunes en esos debates suele formularse como "Eres un idiota; todos están de acuerdo con mi idea" que es complementado con un comentario desdenoso, “su idea carece de evidencia, e incluso si tiene alguna, no es tan fuerte como la mía”. De hecho, parece que cuando se discute sobre cuestiones educativas, la tendencia a ser idiota es bastante común, y en muchos casos orgullosamente idiotas. Por "idiota", no nos estamos refiriendo al uso común de alguien que no es muy inteligente, sino en el significado original de la palabra en griego antiguo, como Walter C. Parker (2005) explica:

[...] la idiotez comparte con la palabra idiosincrasia y la raíz “idio”, que significa privado, separado, egocéntrico, egoísta. "Idiota" era en el contexto griego, un término de reproche. Cuando el comportamiento de una persona se volvía "Idiota", preocupado miopemente por sus beneficios privados y sin tener en cuenta los bienes en común, entonces se creía que la persona era como un barco sin timón, sin consecuencias salvo por el peligro que representaba para otros. Este significado de idiotez logra su fuerza cuando se contrasta con *politēs* (ciudadano) o público. Aquí tenemos una poderosa oposición: el individuo privado versus el ciudadano público (p. 344).

A pesar de la intensidad y del carácter frecuentemente polémico de los debates sobre el propósito de las ILSA<sup>9</sup> y el desarrollo de las GLM, existe poca investigación que examine hasta qué punto los ministros de educación, los legisladores y otros actores sociales y políticos en contextos nacionales le dan valor a este tipo de mediciones del rendimiento estudiantil. Tampoco existe una literatura afianzada que analice hasta qué punto, o cómo, estos agentes educativos incorporan las evaluaciones a gran escala o las métricas globales a su trabajo a nivel nacional.<sup>10</sup>

Para ofrecer respuesta a estas cuestiones, hemos llevado a cabo una revisión de literatura sobre la investigación y las políticas relacionadas con las evaluaciones a gran escala y las métricas globales del aprendizaje y hemos realizado dos encuestas con involucrados y/o interesados en ese ámbito para entender cómo estas evaluaciones y métricas están influyendo en las políticas de educación primaria y secundaria a nivel nacional. Más específicamente, pretendemos entender:

-¿Hasta qué punto las ILSA/GLM han influido en las políticas nacionales de educación primaria y secundaria?

-¿Qué cambios en políticas nacionales de educación primaria y secundaria se han producido en países donde hay indicios de la influencia de las ILSA/GLM?

Nuestro objetivo en última instancia es ofrecer a la comunidad internacional una mejor comprensión de la forma en que (si de hecho lo hacen) las ILSA/GLM influyen en el trabajo de los agentes educativos a nivel nacional (considerando, por ejemplo, el “shock PISA” y fenómenos similares). Nuestra investigación también pretende servir como base para futuros estudios sobre los efectos de las ILSA/GLM sobre la toma de decisiones y las reformas en educación a nivel nacional e infranacional.

## Metodología

Como descrito anteriormente, llevamos a cabo una revisión de literatura sobre ILSA/GLM y realizamos dos encuestas sobre la percepción de las mismas por parte de expertos en política educativa y otros investigadores, legisladores, representantes de sindicatos y educadores. En las próximas secciones se detalla nuestro abordaje metodológico para recolectar y analizar esas fuentes de datos.

## Revisión de literatura

### Recogida de datos

Recopilamos material de relevancia publicado entre 2000 y 2016 en inglés, español y portugués. Dado que la mayoría de ILSA dirigidas a la enseñanza primaria y secundaria comenzaron después del año 2000 (ver glosario), nuestra revisión abarcó parte sustancial de la literatura sobre ILSA/GLM disponible hasta la fecha.

Realizamos búsquedas en cinco bases de datos electrónicas de reconocido prestigio (Google Scholar, EBSCO-Education Full Text, ERIC, JSTOR y SciELO) por medio de palabras clave (como, por ejemplo, evaluaciones educativas a gran escala, métricas globales del aprendizaje, modelo de reforma educativa global e impacto de políticas, uso de políticas, influencia de políticas). Nuestro mayor interés era obtener informes oficiales y notas sobre políticas producidos por organizaciones nacionales e internacionales (como la OCDE o la UNESCO) y artículos académicos publicados en revistas revisadas por pares que ofreciesen análisis basados en investigaciones empíricas. En otras palabras, no incluimos comentarios o perspectivas basadas en opiniones, aunque estuvieran publicadas en revistas revisadas por pares. También incluimos capítulos de libros relevantes, pero no libros enteros, dados los límites temporales de nuestro estudio. Al buscar artículos periodísticos relevantes, nuestro objetivo fue centrarnos en publicaciones empíricas en vez de recurrir a informes que no fuesen de primera mano. No obstante, los artículos que encontramos no documentaban directamente la influencia de las ILSA/GLM en políticas educativas.

Aunque existen ILSA y GLM orientadas específicamente a la educación superior y de adultos — como el Estudio Internacional sobre la Formación Inicial en

Matemáticas de los Maestros (TEDS-M) del IEA, la Evaluación de los Resultados del Aprendizaje en la Educación Superior (AHELO) y el Programa para la Evaluación Internacional de las Competencias de Adultos (PIAAC) de la OCDE, o el Informe Mundial sobre el Aprendizaje y la Educación de Adultos (GRALE) de la UNESCO — nos hemos limitado en este estudio inicial a las ILSA que se dirigen a estudiantes de educación primaria y secundaria. La próxima fase de nuestro análisis consistirá en expandir nuestra investigación e incluirlas todas, independientemente de la población a la que se dirigen.

#### Muestra de literatura

En nuestra búsqueda inicial, identificamos un total de 238 publicaciones, 83 de las cuales específicamente sobre cómo las ILSA/GLM han influido en políticas educativas a nivel nacional o infranacional. Resulta interesante que casi todas esas publicaciones son en inglés, siendo muy pocos los artículos publicados en español o portugués basados en investigación empírica con foco específicamente en la utilización de las ILSA/GLM a nivel nacional o en su impacto en políticas educativas. Algunos de los artículos en español y portugués se trataban de traducciones de artículos publicados en inglés.

Más de dos tercios (69%) del conjunto final de 83 textos provenían de revistas revisadas por pares. Las revistas más citadas fueron las siguientes: *European Journal of Education* (9)<sup>11</sup>, *Journal of Education Policy* (7), *Comparative Education* (5) y *European Educational Research Journal* (5). La mayoría de los artículos (80%) se publicaron después de 2009, lo cual ilustra el interés creciente en documentar la influencia de las ILSA/GLM en los esfuerzos políticos por reformas educativas a nivel nacional e infranacional. El conjunto final de publicaciones abarca el perfil de las políticas de 47 países, así como un comentario general sobre los países de la Unión Europea y de la OCDE. Tal como observó Lockheed (2015), los países de altos ingresos del Norte Global son los casos más estudiados en la literatura, siendo Alemania (11) el país cuyo perfil es abordado con mayor frecuencia, seguido de Francia (8), Inglaterra/ Reino Unido (8), Estados Unidos (7) y Canadá (7).

El foco de la mayoría (82%) de publicaciones analizadas es la respuesta al programa PISA a nivel nacional, y en algunos casos a nivel infranacional (ENGEL & FRIZZELL, 2015). Cuatro estudios se concentran en el impacto en reformas educativas del TIMSS, tres en el de la EGRA y dos en el del PIRLS y el SACMEQ. No encontramos ninguna publicación que examinase la influencia de las otras ILSA más relevantes (ICCS, LAMP, LLECE y WEI-SPS) en la política educativa, es decir, a pesar de haber publicaciones que documentan sus resultados, la literatura revisada no se dedica expresamente a estudiar sus impactos en los esfuerzos de reforma de políticas educativas nacionales. Este hecho puede estar asociado a la prominencia mundial del programa PISA y a las recomendaciones de la OCDE a los países que aspiran a mejorar sus resultados en el mismo, así como al extraordinariamente publicitado “shock PISA” (véase HOPFENBECK et al., 2017; OECD, 2014; WALDOW, 2009; YOUNG, 2015) que han sufrido muchos países.

## Análisis de datos

Tras concluir el corpus de artículos, dimos comienzo a nuestra revisión de literatura seleccionando un grupo aleatorio de ocho artículos para calibrar nuestro proceso de codificación. Garantizar la concordancia entre códigos se considera una buena práctica cuando en el proyecto están involucrados diversos investigadores, pues ayuda a aumentar la fiabilidad del análisis de datos estableciendo un conjunto común de códigos (CRESWELL, 2007). Los miembros del equipo de investigación leímos y codificamos por separado los ocho artículos y después nos reunimos para discutir nuestras consideraciones. Durante esta discusión inicial, alteramos la matriz de revisión para que reflejara los factores más significativos en los que nos pusimos de acuerdo (como, por ejemplo, contexto del país, influencia en las políticas o indicios). Entonces procedimos a usar la matriz para los artículos restantes.

## Posibles limitaciones

Nuestra revisión de literatura se limitó a publicaciones en inglés, español y portugués y de esta forma excluyó la literatura relevante publicada en otras lenguas, como el libro en alemán de Tillmann, Dederling, Kneuper, Kuhlmann y Nessel (2008) sobre el impacto del programa PISA en la política educativa alemana. No obstante, el inglés es la lengua en que la inmensa mayoría de revistas académicas se publica (van WEIJEN, 2012). Por otro lado, mientras el marco temporal de nuestra selección de literatura coincide con la realización de las ILSA más importantes e implementadas a nivel mundial, puede que haya literatura relevante publicada antes del año 2000. Nuestra decisión de excluir libros también puede haber conllevado pasar por alto algunas publicaciones relevantes, pero la inclusión de capítulos de libros contribuyó a mitigar esta posible limitación.

## Encuesta con profesionales con intereses en las ILSA

### Instrumentos de encuesta

Para complementar nuestra revisión de literatura sobre las ILSA, elaboramos y realizamos dos encuestas para investigar la percepción que tienen del uso que se hace de las mismas por parte de investigadores legisladores, representantes de sindicatos y educadores (ver apéndice). Ambas encuestas se diseñaron para recoger impresiones sobre qué ILSA se utilizan y cómo y hasta qué punto facilitan o dificultan el emprendimiento de reformas globales y nacionales, y consistieron en preguntas abiertas y cerradas. La segunda encuesta, sin embargo, incluyó cuestiones demográficas adicionales y abordó el uso de las ILSA por parte de los encuestados y su involucración en ese ámbito. Se desarrollaron encuestas piloto con cinco expertos.

### Recogida de datos

Nuestras encuestas siguieron los métodos de muestreo de bola de nieve y de conveniencia. Realizamos ambos tipos de encuestas entre noviembre y diciembre de 2016.



Muestra de la encuesta a expertos. Diseñamos la primera encuesta para recoger las impresiones de investigadores académicos y del tercer sector reconocidos nacional e internacionalmente por su trabajo en el ámbito de las ILSA/GLM y de profesionales de las propias organizaciones que realizan las ILSA o están involucradas en el desarrollo de las GLM. Con base en nuestro trabajo sobre las ILSA, identificamos una muestra de conveniencia de 50 expertos que sabíamos tienen una vasta experiencia en la investigación y trabajo en este ámbito y les mandamos un enlace a la encuesta por e-mail. La encuesta incluía un pedido de recomendaciones de compañeros a los que tal vez también les gustaría compartir sus perspectivas sobre el uso e influencia de las ILSA/GLM. Este muestreo de bola de nieve nos llevó a incluir 34 profesionales más para un índice de respuesta final del 29% (n=24).

Muestra de la encuesta a profesionales con intereses en las ILSA. Diseñamos la segunda encuesta para recoger las opiniones de profesionales con intereses en cuestiones relacionadas con las ILSA/GLM, como educadores y legisladores. Los participantes fueron los inscritos en el Simposio Inaugural de la Sociedad de la Educación Internacional y Comparada (Inaugural Symposium of the Comparative and International Education Society), realizado en la Arizona State University los días 10 y 11 de noviembre de 2016. El simposio se centró específicamente en la conveniencia y posibilidades de las ILSA/GLM y a él asistieron un grupo diverso de investigadores, legisladores y otros interesados en la cuestión. La encuesta fue enviada por e-mail a la lista completa de inscritos en el simposio (n=132) y tuvo un índice de respuesta del 50% (n=66).

#### Análisis de la encuesta

Ambas encuestas se efectuaron a través de Qualtrics y los ítems cuantitativos los analizamos también a través de ese software. Exportamos las respuestas abiertas en inglés y español, dos de los miembros de nuestro equipo de investigación las codificaron por separado en Microsoft Excel y entonces comparamos los códigos para garantizar la fiabilidad interevaluadores. Había cinco preguntas formuladas de forma idéntica en los dos tipos de encuestas. Agrupamos las respuestas a esos ítems para analizar los resultados de forma independiente según el estatus de los participantes, como expertos o como otros profesionales con intereses en las ILSA.

En las próximas secciones presentamos, en primer lugar, nuestra revisión de literatura sobre las ILSA/GLM y, en segundo, nuestro análisis de las dos encuestas de profesionales incumbidos por ellas. Nos gustaría ya hacer constar que para hablar de la implementación y uso de las ILSA/GLM nos referimos intencionalmente a gobiernos y agencias gubernamentales en lugar de hacerlo de forma más general (como es también más habitual) a naciones o países, ya que la contribución a esfuerzos de reforma o su obstaculización dependen de los intereses políticos de gobiernos y legisladores.

## Revisión de literatura: las ILSA/GLM como herramientas de legitimación

La atención mediática que las ILSA vienen recibiendo, particularmente tras el lanzamiento de los informes PISA y los rankings internacionales (BAIRD et al., 2016; CARVALHO, COSTA & GONÇALVES, 2017; FERRER, 2017; HOPFENBECK & GÖRGEN, 2017; MEYER & BENAVIDOT, 2013) indica que la participación en las ILSA está teniendo un impacto considerable en los discursos sobre educación y políticas a nivel mundial (BREAKSPEAR, 2012; HOPFENBECK et al., 2017). Nuestra revisión de literatura sugiere, no obstante, que no es fácil identificar vínculos causales directos entre la participación en las ILSA y reformas educativas, una vez que algunos gobiernos parecen haber implementado cambios en sus políticas a consecuencia de los resultados de esas evaluaciones mientras que otros parecen haber usado esos resultados para justificar y/o acelerar reformas educativas ya en marcha. En cualquier caso, las ILSA están siendo efectivamente usadas para legitimar esas diversas agendas de reformas en las políticas educativas. Además, el arduo esfuerzo de los gobiernos por subir en los rankings ha contribuido a que los países emulen políticas entre ellos y a una convergencia generalizada entre las mismas (BAIRD et al., 2011; BREAKSPEAR, 2014; LAWN & GREK, 2012; MOWAT, 2017; VERGER, 2014). Ha habido también una cantidad creciente de estudios sobre la participación y resultados en las ILSA en revistas académicas (DOMINGUEZ, VIEIRA & VIDAL, 2012; LENKEIT, CHAN, HOPFENBECK & BAIRD, 2015), lo que ha contribuido a una percepción más amplia de que la participación en esas evaluaciones ha tenido un impacto significativo en las políticas educativas a nivel mundial. En este sentido, los gobiernos de los países participantes están blandiendo las ILSA (y, por extensión, las GLM) como herramientas de legitimación para promover o justificar la creación, mantenimiento o transformación de políticas educativas y esfuerzos de reforma (ver también AD-DEY, SELLAR, STEINER-KHAMSI, LINGARD & VERGER, 2017).<sup>12</sup>

Uno de los principales desafíos a la hora de determinar hasta qué punto hay pruebas concluyentes de la relación entre la participación en las ILSA y el rumbo de las reformas educativas nacionales es que en países con resultados semejantes en las evaluaciones las respuestas gubernamentales a las políticas de reforma educativa han sido diferentes (MARTENS & NIEMANN, 2013). Dos ejemplos llamativos se dieron a partir de la divulgación de los resultados del informe PISA en 2001 y 2004, momentos en los que un grupo de países sufrieron lo que se ha denominado el “shock PISA” (OECD, 2014; WALDOW, 2009; YOUNG, 2015). Los legisladores de Alemania, Japón y Dinamarca vieron cómo otros países alrededor del mundo les superaban considerable y sorprendentemente y, bajo un escrutinio público y mediático intenso, reaccionaron rápidamente y empezaron a desarrollar e implementar reformas educativas a gran escala (BREAKSPEAR, 2012, 2014; BULLE, 2011; TASAKI, 2017). El gobierno alemán creó una nueva serie de estándares nacionales acordes con el informe PISA que le permitió monitorizar la trayectoria del rendimiento de los estudiantes a lo largo del tiempo y desarrolló nuevos sistemas de apoyo a estudiantes desfavorecidos e inmigrantes, cuyos resultados fueron particularmente bajos (BREAKSPEAR, 2014; ERTLE, 2006). Así también hizo el gobierno danés.

El vigoroso debate público y político que se siguió a la caída de Japón en el ranking de las evaluaciones del programa PISA entre 2000 y 2003, sin embargo, llevó a sus legisladores a invertir en evaluaciones nacionales y elevar los niveles asociados a las mismas (BREAKSPEAR, 2014; TAKAYAMA, 2008; TASAKI, 2017).

Este problema también ha sido bien descrito por Baird et al. (2016) en su estudio sobre las reacciones políticas y mediáticas entre los resultados de 2009 y 2012 en Canadá, China (Shanghái), Inglaterra, Francia, Noruega y Suiza. Los autores observaron que los gobiernos hacen uso de sus resultados en el informe PISA como un “hechizo” o “varita mágica” para justificar cualquier reforma en sus políticas, independientemente de los resultados obtenidos (véase también FERRER, 2017; HOPFENBECK & GÖRGEN, 2017). Dentro de un abordaje diferente y más positivo, el análisis de Lockheed y Wagemaker (2013) de la utilidad de las ILSA como herramientas políticas las describe como “termómetros” para medir progresos y a la vez como “látigos” que instigan a la acción política. Otros estudiosos las entienden como herramientas de “regulación del conocimiento” (CARVALHO, COSTA & AFONSO, 2009) en los espacios de decisión política, legitimación de discursos y construcción de capacidad práctica que han de servir como agentes de transformación y no apenas como parámetros de acceso o progreso (ver CARVALHO, 2012; CARVALHO & COSTA, 2015; LASCOURMES & LE GALES, 2007). Al margen de cómo la literatura ha descrito el posicionamiento en las ILSA, resulta evidente que las autoridades educativas y políticas y los grupos políticos de interés de cada país están haciendo uso de las ILSA para legitimar políticas educativas de formas que reflejan el contexto político en que se encuentran.

A través de la literatura revisada, encontramos indicios de que la participación nacional en las ILSA y los resultados que de ella desprenden se han utilizado para legitimar y acelerar el emprendimiento de reformas ya existentes o ya planificadas. Por ejemplo, según Baird et al. (2011), el gobierno francés exageró el desempeño relativamente insuficiente en las evaluaciones TIMSS, PIRLS y PISA para justificar reformas gubernamentales sistémicas que ya estaban en marcha. Estas reformas incluyeron volver a centrarse en habilidades básicas a través de la revisión curricular de la educación primaria con un foco especial en la alfabetización y las ciencias y descentralizar la financiación para que las escuelas tengan control sobre sus presupuestos. Según Dobbins y Martens (2012), los resultados de Francia en el PISA, que fueron precedidos por un mal desempeño en la Encuesta Internacional sobre Alfabetización de Adultos (IALS) de la OCDE en 1997, resultó en un descontento público que le dio a los dirigentes gubernamentales la oportunidad de avanzar rápidamente en las reformas estructurales pretendidas dirigidas a mejorar el desempeño de los estudiantes, la rendición de cuentas y la autonomía tanto de las escuelas como de los profesores (ver también Pons, 2016). Esta última medida (darle más autonomía financiera a escuelas y profesores) refleja la forma en la que el gobierno francés buscó emular las que entendieron como principales causas del “milagro” finlandés en el ranking PISA (DOBBINS & MARTENS, 2012; ver también MICHEL, 2017).

De la misma forma, la participación de Suiza en el PISA 2000 y en los programas PIRLS y TIMSS de la IEA contribuyó a que el gobierno acelerara medidas pre-existentes para unificar en mayor medida los planes de estudio y estándares en todo el país, así como para desarrollar sistemas más eficientes para controlar el progreso educativo entre regiones y dentro de cada una de ellas (BAIRD et al., 2011; BIEBER & MARTENS, 2011). El nuevo gobierno laborista del Reino Unido también ha hecho uso de los resultados del informe PISA en el país para promover una narrativa sobre un declive en los estándares educativos que sirvió para justificar reformas en los objetivos del Certificado General de la Educación Secundaria (GCSE) (SELLAR & LINGARD, 2013). En Japón y Portugal, los legisladores han hecho uso del informe PISA como un aliado retórico más sutil para re-contextualizar y orientar las reformas gubernamentales propuestas. Los mismos han reinterpretado el éxito finlandés para justificar reformas curriculares y la implementación de evaluaciones estandarizadas (TAKAYAMA, 2008, 2010). Mientras que en Portugal es más difícil aislar la influencia del informe PISA en reformas educativas de otros factores (como la agenda educativa de la Unión Europea), las preocupaciones de los legisladores portugueses en relación a los resultados del informe PISA han coincidido con la implementación de evaluaciones nacionales en lengua portuguesa y matemáticas y con la reconsideración de los currículos en términos de competencias (TEODORO & ESTRELA, 2010; ver también CARVALHO et al., 2017).

Muchas de las reformas implementadas por los gobiernos se han centrado en desarrollar o fortalecer sus programas y marcos de evaluación nacional como resultado de su participación en el PISA. Engel (2015), por ejemplo, señala que el gobierno español ha venido usando progresivamente el informe PISA como “una referencia para identificar las áreas prioritarias de la política educativa española” (p. 110; ver también FERRER, 2017). Los responsables de las políticas educativas finlandesas, habiendo sufrido también el “shock PISA”, implementaron un sistema nacional de valoración de la calidad y evaluación de las escuelas, así como proyectos sobre currículo nacional y evaluación (BAIRD et al., 2011). Mientras que hay otros gobiernos nacionales que han implementado nuevos sistemas de exámenes y evaluación durante este periodo de tiempo, como Rumanía, Suecia, Macedonia, Malasia, Noruega e Israel (BAIRD et al., 2011; CHOI & JERRIM, 2016; ELLEY, 2002; ENGEL, 2015; FEINGER et al., 2012; FORESTIER & CROSSLEY, 2014), el informe PISA no puede ser señalado como la única causa determinante para tal ya que la tendencia educativa mundial ha sido la de aumentar la cuantificación y la evaluación y monitorización del desempeño estudiantil.

Asimismo, algunos gobiernos nacionales han introducido nuevos estándares y/o modificado su currículo nacional. El gobierno de Shanghái, por ejemplo, ha hecho cambios significativos en el currículo, tales como el paso de pedagogías con foco en la “transmisión de contenido” a pedagogías con foco en problemas “de la vida real” (SELLAR & LINGARD, 2013). Los responsables de las políticas educativas de Macedonia, Malasia, Portugal, Rumanía, España y Suiza también desarrollaron estándares y/o currículos nacionales nuevos (BAIRD et al., 2011; CARVALHO et al., 2017; CHOI & JERRIM, 2016; ELLEY, 2002; ENGEL, 2015; TEODORO &

ESTRELA, 2010). Los legisladores españoles, por ejemplo, usaron los informes PISA, PIRLS y TIMSS para justificar la simplificación del currículo nacional centrándolo en competencias, lenguas extranjeras y tecnologías de la información y la comunicación (CHOI & JERRIM, 2016; ver también GIL, BELTRÁN & REDONDO, 2016). Los legisladores macedonios, tras la primera participación del país en el TIMSS-R en 1999 (su primera participación en una evaluación internacional) empezaron a trabajar en un mejor alineamiento del currículo de matemáticas y ciencias con las evaluaciones internacionales, revisaron sus evaluaciones nacionales y pusieron en marcha un programa de formación continua de profesores (ELLEY, 2002).

Otra reacción común a las ILSA/GLM ha sido proponer la modificación de las políticas de calidad docente, particularmente en las áreas de formación de profesores, evaluación, desarrollo docente y requisitos docentes. Se ha visto que este ha sido el caso en España, Estonia, Polonia, Brasil, Colombia, Japón, Portugal e Israel (AFONSO & COSTA, 2009; BIAŁECKI, JAKUBOWSKI & WIŚNIEWSKI, 2017; CHOI & JERRIM, 2016; ENGEL, 2015). Paine, Bloemeke y Aydarova (2016) defienden que los cambios en políticas en el ámbito docente son habituales porque se trata de un área en la que los legisladores pueden influir con facilidad. Postulan que la forma más eficaz de influir en los resultados del desempeño de los alumnos en las ILSA es centrarse en cómo se contribuye a ellos. Mientras factores complejos como clase social son los que tal vez mejor ayuden a prever los logros de los estudiantes, Paine y sus colegas defienden que centrarse en la “calidad docente” está “al alcance” de los legisladores, lo que resulta en medidas para rendición de cuentas que han hecho que el control de los profesores y de los programas de formación de profesores se haya tornado más estricto. Cabe señalar que es difícil, si no imposible, atribuir exclusiva y explícitamente tales cambios en las políticas educativas a la participación en las ILSA. Como defienden Afonso y Costa (2009), “el programa PISA se adapta a los contextos, los gobiernos se lo apropian y se re-politiza según una agenda preexistente” (p. 61).

Algunos estudiosos señalan el discurso creado por la cobertura mediática de los resultados de las ILSA como la justificación que los gobiernos usan para el emprendimiento de reformas educativas (BAROUTSIS & LINGARD, 2016; BIAŁECKI, JAKUBOWSKI & WIŚNIEWSKI, 2017; HOPFENBECK & GÖRGEN, 2017; SELLAR & LINGARD, 2013; TAKAYAMA, 2008). Se podría decir, tal como se señaló anteriormente, que las ILSA han tenido un mayor efecto retórico que directamente político (BAIRD et al., 2011; PAINE et al., 2016; TAKAYAMA, 2010; TEODORO & ESTRELA, 2010). En este sentido, las ILSA han llegado a servir como una lógica política que contribuye a legitimar reformas. Berényi & Neumann (2009) dieron parte de cómo, durante la elección húngara de 2002, el partido liberal hizo cada vez más referencia a los resultados del informe PISA como una razón para fundamentar su agenda educativa (véase también BAJOMI, BERÉNYI, NEUMANN & VIDA, 2009). El éxito de Hong Kong en el informe PISA ha llamado la atención de las autoridades británicas, que también se han fijado en las mejores prácticas de Singapur, Suecia y Finlandia (FORESTIER & CROSSLEY, 2014). El entonces secretario de Estado de Educación, Michael Gove, frecuentemente se refirió al informe PISA y a los países con los mejores resultados para justificar sus medidas

en educación (BITTLINGMAYER, BOUTIUC, HEINEMANN & KOTTHOFF, 2016; FORESTIER & CROSSLEY, 2014; ver también BAIRD et al., 2011). El uso de comparaciones y rankings cuasi-educativos no ha sido excepcional entre los países, ya que muchos han atestiguado un giro discursivo a partir de que los medios de comunicación hayan empezado a informar sobre los resultados del informe PISA. Algunos países con sistemas educativos tan diferentes como Argentina, Australia, Francia, Alemania, Japón, Portugal, España, Turquía y Estados Unidos han atestiguado efectos retóricos semejantes (AFONSO & COSTA, 2009; BAROUTSIS & LINGARD, 2016; CARVALHO, COSTA & GONÇALVES, 2017; FROESE-GERMAIN, 2010; GORUR, 2011; NIEMANN, MARTENS & TELTEMANN, 2017; RAUTALIN & ALASUUTARI, 2008; TAKAYAMA, 2008; GIL et al., 2016). Estos países también han implementado varias políticas educativas que se han asociado a la conectividad y competitividad crecientes (posibilitadas por la participación en las ILSA), tales como la introducción en Australia de estándares de enseñanza nacionales (BREAKSPEAR, 2014; FIGAZZOLO, 2009; GORUR, 2011; SELLARS & LINGARD, 2013), el uso por parte de los Estados Unidos de los resultados del desempeño de los alumnos para medir la calidad docente (BIEBER & MARTENS, 2011; ENGEL & FRIZZELL, 2015; FIGAZZOLO, 2009; MARTENS & NIEMANN, 2013; PLISKO, 2013; SELLAR & LINGARD, 2013) y la reducción por parte de España de su currículo para centrarlo en materias evaluadas (CHOI & JERRIM, 2016) y en competencias (GIL et al., 2016).

Nuestra revisión ha encontrado indicios de que las ILSA también se han usado como herramientas de legitimación a nivel infranacional o local. Morgan (2015) identificó una rápida adopción infranacional alrededor de las reformas en Canadá acordes con el programa PISA. De la misma forma, Engel y Frizzell (2015) documentan la forma en que las provincias canadienses han hecho uso de los resultados del informe PISA para otorgar validez a la forma en que las políticas existentes han contribuido al buen desempeño de sus alumnos (como, por ejemplo, en Alberta) o a las reformas para combatir el mal desempeño de los mismos (como en Saskatchewan). Según Breakspear (2014), seis países de la OCDE han empezado a fijar nuevos objetivos para los rankings internacionales y a emular prácticas y políticas provenientes de otros lugares (prestando especial atención a las “mejores prácticas” de los países con mejores clasificaciones), lo que ha tenido en sí muy poco impacto práctico real.

Figura 1– Quantidade e motivos de exclusão de publicações durante o processo de seleção de artigos

País	Objetivo
Australia	Ser clasificado entre los 5 mejores países del PISA por 2025.
Brasil	Llegar al promedio de la OECD del rendimiento del PISA por 2021, y todas las escuelas y regiones sean clasificadas en el Índice de Desarrollo de Educación Básica, que es medido por el PISA.
Dinamarca	Ser clasificado entre los 5 mejores países del PISA.
México	Llegar a los 435 en matemáticas y lectura por 2012.
Tailandia	Llegar al promedio de la OECD del rendimiento del PISA por 2021.
Gales	Ser clasificado entre los 20 mejores países del PISA por 2015.

Fuente: Breakspear, 2014.

Son pocos los indicios documentados de la influencia de las ILSA en las reformas educativas en los Estados Unidos, en parte posiblemente por la compleja relación de este país con esas evaluaciones. Aunque los responsables por las políticas educativas estadounidenses presionaron a la OCDE para que desarrollara comparaciones entre el desempeño escolar de los países a principios de los noventa (LINGARD & LEWIS, 2016), el país siempre ha mostrado un desinterés histórico en involucrarse en las ILSA a nivel nacional (BIEBER & MARTENS, 2011)<sup>13</sup>. Hay algunos indicios de que legisladores de los estados han empezado a hacer uso de los resultados en el informe PISA como prueba de la eficacia de sus sistemas educativos (ENGEL & FRIZZELL, 2015; LINGARD & LEWIS, 2016). En Massachusetts, por ejemplo, recientemente se citaron los resultados del informe PISA de 2012 como prueba de la competitividad nacional e internacional de ese estado, y en Florida se incorporaron los resultados del informe PISA a la propuesta de becas del programa Race to the Top (literalmente “correr/carrera hasta la cima”) (ENGEL & FRIZZELL, 2015). Resulta interesante que, aunque casi no se han observado conexiones entre las ILSA y los esfuerzos de reforma educativa estadounidenses, Martens & Niemann (2013) documentan el paralelismo entre los objetivos establecidos por el programa PISA y los objetivos del programa Race to the Top conforme establecidos en la ley estadounidense de recuperación y reinversión American Recovery and Reinvestment Act, de 2009, y la ley de educación A Blueprint for Reforms (“un proyecto de reformas”) re-aprobada en 2010. Según Froese-Germain (2010), el informe PISA se está usando en los Estados Unidos como un sutil aliado para justificar reformas educativas dirigidas a aumentar el rendimiento y la competitividad (y posiblemente en otros países también).

Un resultado de la creciente influencia de las ILSA en los debates sobre reforma educativa es su contribución a la creación de agencias nacionales de evaluación de la educación y el desarrollo paralelo de la capacidad técnica de aplicar, monitorizar e interpretar resultados (ADDEY ET AL., 2017; GOVE & CVELICH, 2011; GREGER, 2012; LEWIS & LINGARD, 2015; STRIETHOLT & SCHERER, 2017). Gove y Cvelich (2011) dan parte de las consecuencias del EGRA en varios países participantes señalando los esfuerzos para mejorar la posición en el ranking por parte de países de bajos ingresos que han tenido históricamente una capacidad de evaluación limitada. De forma semejante, antes del informe PISA, Alemania y la República Checa no implementaban evaluaciones nacionales y el “shock” resultante dio lugar a que los gobiernos de ambos países desarrollasen métodos más formales para evaluar el progreso de sus estudiantes (GREGER, 2012; NIEMANN, MARTENS & TELTEMANN, 2017). Pons (2017) le atribuye al informe PISA particularmente la creación del área de política educativa global y la consecuente necesidad de nuevos marcos teóricos y herramientas metodológicas para entender, medir e interpretar los resultados de las ILSA.

Más allá de fijar objetivos en relación a los resultados y rankings del informe PISA, y a pesar de las obvias diferencias entre los diversos contextos, las ILSA ofrecen información sobre las competencias de los estudiantes de los países y, en última instancia, ofrecen indicios sobre lo que se puede considerar estándares nacionales de calidad de los sistemas educativos. Además, el examen de Carvalho y Costa (2015) de

la recepción política del informe PISA en seis países europeos sugiere que ese informe en particular “invita al optimismo sobre la posibilidad de reformas y crea confianza en los agentes de la política nacional” y legitima “problemas y soluciones en las políticas con la aprobación de un conocimiento supuestamente universal, independiente y especializado” (p. 644). Esta alteración de condiciones de posibilidad ha influido de diversas formas en todo el mundo, lo que ha conllevado la generación de un robusto mercado educativo para las ILSA, el desarrollo de sectores enteros dedicados a su implementación y procesamiento (como, por ejemplo, en la IEA, la OCDE, Pearson, la UNESCO, etc.), el desarrollo de la capacidad de evaluación a través de unidades de evaluación y sistemas de recolección de datos nacionales y la emergencia de un nuevo sector profesional de asesoría sobre las ILSA.

### Encuestas a profesionales: La percepción de las ILSA/GLM como instrumentales para la reforma de políticas educativas

Un total de 90 profesionales respondieron a las encuestas, entre los cuales se incluyen 24 expertos que trabajan directamente con ILSA/GLM, 40 investigadores y legisladores y 13 estudiantes de posgrado que desarrollan investigación sobre el tema, así como otros profesionales involucrados en este ámbito (profesores, gestores universitarios y personal de fundaciones y organizaciones contribuyentes). La encuesta A se dirigió específicamente a expertos con experiencia directa trabajando en las ILSA/GLM dentro y fuera de sus países. La mayoría de los encuestados con intereses en las ILSA/GLM (72%) (como educadores y legisladores) afirmaron tener también experiencia en el trabajo con esas evaluaciones junto a otros profesionales con intereses en ese ámbito (como políticos, otros legisladores, sindicatos, profesores, medios de comunicación, etc.) en diferentes países. El 23% (n= 15) de ellos afirma trabajar frecuentemente con esos grupos de profesionales, el 49% (n= 32) ocasionalmente y el 28% (n= 18) nunca. La experiencia de los participantes de ambas encuestas con las ILSA/GLM abarca una gran variedad de regiones, que van desde África a América Latina, Asia, Europa Occidental y del Este, Oriente Medio, América del Norte y Australia.

Al ser interrogados sobre su experiencia en el trabajo con ILSA/GLM, de los 90 encuestados, 58% (n=52) respondieron que los grupos con intereses en el ámbito educativo están preocupados en primer lugar con los estudios internacionales sobre rendimiento estudiantil como PISA (40%), TIMSS (24%), PIRLS (12%) y PIAAC (9%). Los encuestados también mencionaron otros instrumentos de evaluación que pueden desempeñar un papel importante en contextos de políticas educativas nacionales, como el EGRA de la USAID, la Evaluación del Aprendizaje para el Desarrollo Internacional (IDELA) de Save the Children, el Tercer Estudio Regional Comparativo y Explicativo (TERCE) de la UNESCO o evaluaciones estudiantiles nacionales, regionales y municipales (como las brasileñas SAEB y ENEM). Las próximas secciones dan parte detallado de las perspectivas de los encuestados sobre el uso e impacto de las ILSA/GLM en contextos de políticas nacionales y sobre si las ILSA/GLM facilitan o dificultan el emprendimiento de reformas educativas.



## Las ILSA/GLM en contextos de política nacional

Aunque la literatura existente tal vez no sea fructífera en lo que se refiere a la relación causal entre los resultados de las ILSA/GLM y su uso en los contextos de política de educación nacional, tanto las encuestas a expertos como a los otros profesionales con intereses en las ILSA resultan significativas en relación a las perspectivas sobre este asunto. De hecho, la inmensa mayoría de los encuestados (96%) están de acuerdo en que se hace uso de las ILSA/GLM en los contextos educativos nacionales (Tabla 1). Sin embargo, las opiniones de los encuestados se dividen en lo que respecta a cómo se hace uso de estos instrumentos en los diferentes contextos de formulación de políticas. En particular, el 58% de todos los encuestados esbozan diferentes formas en que las ILSA/GLM se usan instrumentalmente en los contextos educativos nacionales, mientras que el 38% de los mismos muestran críticas en relación al uso inadecuado de las ILSA/GLM. Sólo el 4% de todos los encuestados afirma que los instrumentos no se usan en absoluto en los procesos nacionales de elaboración de políticas educativas. Al focalizar el uso instrumental de las ILSA/GLM en diferentes contextos nacionales, los encuestados explican que los instrumentos se usan generalmente para establecer objetivos, metas y parámetros para las reformas educativas nacionales, especialmente en términos de equidad educacional y agenda sobre calidad, así como para supervisar la implementación de esas reformas. Tal como afirmó un investigador, por ejemplo: “No se puede hablar sobre mejoras en la educación si no hay parámetros”. Tres encuestados señalaron que las ILSA/GLM también se usan para alinear la evaluación de los currículos y los instrumentos de evaluación estudiantil a estándares internacionales. Además, los encuestados afirmaron que las ILSA/GLM se usan más comúnmente como “desencadenantes/niveladores de políticas” para justificar reformas existentes, para otorgar validez y legitimidad a los cambios en ciertas políticas o para demostrar la necesidad de financiación, especialmente en contextos de desarrollo internacional, lo cual es coherente con lo señalado en nuestra revisión de literatura. El porcentaje de encuestados que discutieron las ILSA/GLM en términos de su valor instrumental fue más alto entre los profesionales con intereses en el ámbito de las ILSA que entre los expertos (63% y 52%, respectivamente)<sup>14</sup>. En otras palabras, una mayoría más significativa de encuestados con intereses en las ILSA/GLM, pero, quizás, con menos experiencia con estos instrumentos respondieron favorablemente en relación al uso instrumental de las ILSA/GLM.

Más de un tercio (38%) del total de encuestados opinaron que las ILSA/GLM se usan generalmente de forma inadecuada en contextos de política nacional. Entre los que confirman ese uso inadecuado, el mayor porcentaje corresponde a los expertos y el menor a los otros encuestados con intereses en el ámbito de las ILSA (43% y 33%, respectivamente). A pesar de reconocer la circulación de estos instrumentos a nivel local y mundial, según los dos grupos de encuestados, los legisladores conocen las ILSA/GLM, pero no las usan adecuadamente por varios motivos. Según los encuestados, algunos legisladores muestran reservas de peso en relación al uso de las ILSA/GLM porque creen que tales instrumentos son, tal como lo describe uno de los expertos encuestados, “instantáneas imperfectas que con más frecuencia distorsionan que dilucidan”. En segundo lugar, las ILSA/GLM son demasiado amplias y abstractas

para un uso provechoso. En tercer lugar, los encuestados señalaron que los legisladores tal vez sencillamente carezcan de las habilidades necesarias para analizar e interpretar los datos de las ILSA/GLM para tornarlos útiles en sus contextos nacionales.

Aunque puede que los legisladores consulten los datos comparativos generados por las ILSA/GLM, las respuestas a las encuestas indican que, en general, los mismos no usan esos datos de forma significativa. Más que a orientar reformas, un encuestado indicó que las ILSA/GLM llevan a la construcción de narrativas de “crisis” y “necesidad”, tal como fue ampliamente documentado en nuestra revisión de literatura. Estas narrativas conducen las reformas educativas nacionales en direcciones específicas, con frecuencia aumentando la rendición de cuentas de profesores y estudiantes y el foco en competencias de aprendizaje. De forma coherente también con nuestra revisión de literatura, cuatro de los expertos encuestados observaron que es más probable que los legisladores en algunos países hagan referencias políticas o simbólicas a las ILSA/GLM como estrategia (para hacer referencia a nombres o culpables) o para comparar la posición de sus países en los rankings mundiales, como se refleja en este comentario de uno de los encuestados: “los legisladores no usan las evaluaciones, sino que simplemente están preocupados con los rankings.” De la misma forma que la literatura sobre las ILSA/GLM indica que estas evaluaciones y medidas se están usando como herramientas para la legitimación de políticas (ver también ADDEY et al., 2017), la mayoría de los expertos encuestados y los otros encuestados con intereses en las ILSA atestiguan esta realidad a nivel nacional.

Tabla 2– El uso de las ILSA/GLM en los contextos de las políticas nacionales de educación

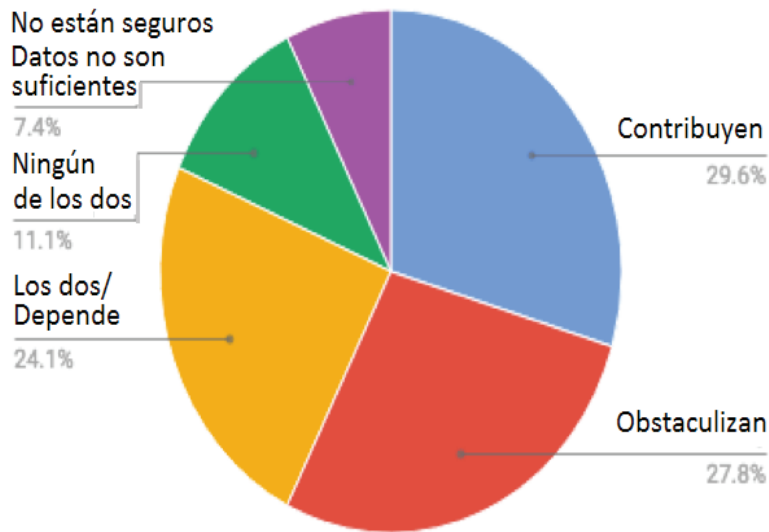
Uso de las ILSA/GLM	Encuesta con los Expertos	Encuesta General	Respuestas combinadas
<i>Usadas para propósitos instrumentales</i> -Para establecer objetivos, metas e medidas para reformas e medir el progreso de las reformas -Validar el cambio de las políticas o justificar las reformas existentes -Demonstrar la necesidad de inversión -Adaptar y/o desarrollar los testes nacionales	52%	63%	58%

<i>Mal usadas</i>	43%	33%	38%
-Los creadores de políticas tienen poco entendimiento de las ILSA/GLM			
-Las ILSA/GLM son descontextualizadas y por lo tanto no son útiles			
-Las ILSA/GLM son muy amplias y abstractas para ser usadas significativamente			
-Para propósitos de “nombre, vergüenza y culpa”			
-Para efectos ceremoniales			
<i>No son usadas</i>	5%	4%	4%
<i>Número de encuestados</i>	21	24	45

#### Percepción del impacto de las ILSA/GLM en la formulación de políticas nacionales

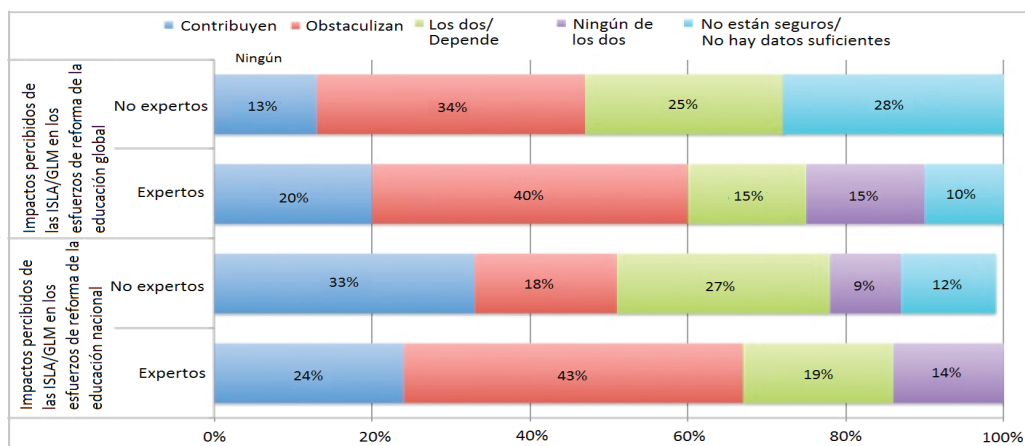
Aunque la gran mayoría del total de encuestados (96%) señalan que las ILSA/GLM se están utilizando en los contextos de elaboración de políticas nacionales, sus opiniones se dividen en relación a si las ILSA/GLM facilitan o dificultan los procesos de reforma educativa (figura 1). En particular, el 30% del total de encuestados cree que las ILSA/GLM contribuyen para las reformas nacionales, mientras que el 28% opina que las dificultan (n=54). Aproximadamente un cuarto (24%) del total de encuestados opinó que las ILSA/GLM tanto pueden facilitar como dificultar los procesos de reformas educativas a nivel nacional, dependiendo de las circunstancias particulares y de las formas en que son utilizadas. Finalmente, el 11% de los encuestados opinó que las ILSA/GLM ni facilitan ni dificultan los procesos de reforma educativa a nivel nacional y el 7% afirmó que no hay indicios o información suficiente para determinar el efecto que las ILSA/GLM tienen en los procesos de elaboración de políticas nacionales.

Figura 1 – La percepción de los encuestados sobre la contribución y/o dificultades de las ILSA/GLM para los esfuerzos de reforma de la educación nacional <sup>15</sup>



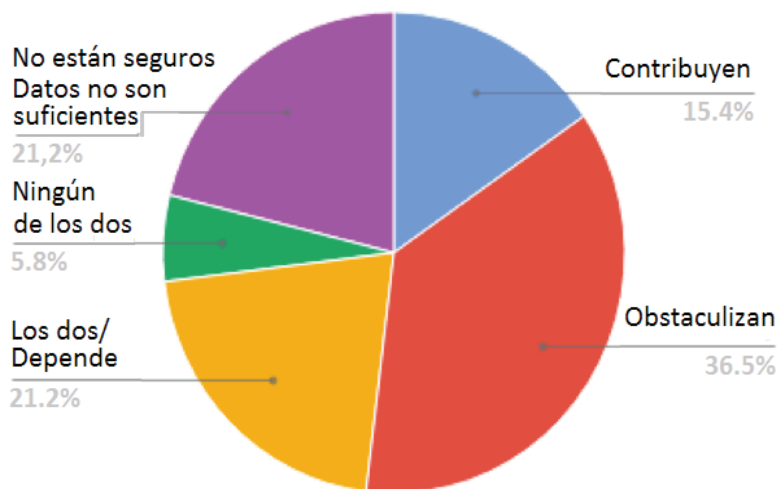
Una vez más, la visión de los expertos encuestados y de los otros encuestados con intereses en el ámbito de las ILSA/GLM es divergente en lo que se refiere a si las ILSA/GLM constituyen una ayuda o un obstáculo para las medidas nacionales de reforma educativa (figura 2). Los expertos se mostraron más críticos hacia las ILSA/GLM. El 43% de los mismos respondieron que las ILSA/GLM obstaculizan los procesos de reforma, frente a tan solo el 18% de los otros encuestados con intereses en las ILSA/GLM. Es posible que su inmersión en la literatura teórica y de investigación que conduce las ILSA/GLM en diversos contextos nacionales fundamente las opiniones de los expertos en lo que respecta a las controvertidas formas en que las evaluaciones internacionales tienen efectos sobre las políticas nacionales.

Figura 2 – La percepción de los encuestados sobre el impacto de las ILSA/GLM en las políticas de educación nacional y global



Observamos un cambio en la percepción de los encuestados a respecto de si las ILSA/GLM facilitan o dificultan los procesos globales de reforma educativa en comparación con sus respuestas en relación a los procesos nacionales (figura 3). En el ámbito de las reformas globales, más de un tercio (37%) del total de encuestados opinaron que las ILSA/GLM dificultan la elaboración de políticas educativas (n=52). Tan solo el 15% del total de encuestados creen que las ILSA/GLM contribuyen a las reformas globales (n=52). Comparando las respuestas de las dos encuestas, el 33% (n=33) de los encuestados del grupo con intereses en las ILSA opinó que las ILSA/GLM ayudan en las reformas a nivel nacional, pero tan solo el 13% (n=32) siente que estén contribuyendo a reformas educativas globales. De forma similar, solo según el 18% (n=33) de ellos las ILSA/GLM obstaculizan los esfuerzos de reforma a escala global. Esa visión es coherente con la de los expertos encuestados. Por ejemplo, el 24% (n=21) de los expertos respondió que las ILSA/GLM facilitan las medidas para reformas nacionales y el 20% (n=20) respondió que las ILSA/GLM facilitan las reformas globales. De forma semejante, también hubo un pequeño cambio en relación a la percepción de los expertos sobre si las ILSA/GLM dificultan los esfuerzos de reforma nacionales y globales. Según el 43% (n=21) de los expertos encuestados, las ILSA/GLM obstaculizan las reformas nacionales mientras que el 40% (n=20) respondió que las ILSA/GLM dificultan los esfuerzos de reformas globales.

Figura 3 – Las percepciones de los encuestados sobre la contribución y/o dificultades de las ILSA/GLM para los esfuerzos de reforma global <sup>16</sup>



#### La contribución de las ILSA/GLM en los esfuerzos de reforma educativa nacionales/globales

Según los encuestados que opinaron que las ILSA/GLM sirven de ayuda a los esfuerzos de reforma educativa nacionales (30% n=54), las ILSA/GLM permiten que los gobiernos alineen sus sistemas educativos a estándares internacionales/globales, alineamiento que incluye también los sistemas de evaluación y examinación y el currículo. Como señaló uno de los expertos encuestados, las ILSA/GLM se usan efectivamente para desencadenar avances en agendas particulares de reforma educativa, pero sin ofrecer detalles sobre tales agendas. No sorprende que este sea exactamente el principal aspecto criticado por aquellos expertos en políticas/investigación que reflexionan sobre el ascenso de las ILSA/GLM en todo el mundo (BIEBER & MARTENS, 2011; ENGEL & FRIZZELL, 2015; FIGAZZOLO, 2009; MARTENS & NEIMANN, 2013; PLISKO, 2013; SELLAR & LINGARD, 2013), lo cual, por otro lado, es coherente con las conclusiones de nuestra revisión de literatura.

#### Las ILSA/GLM como obstáculos para esfuerzos de reforma educativa nacionales/globales

Varios de los encuestados (13%, n=24) del grupo de aquellos con intereses en las ILSA señalaron que las ILSA/GLM constituyen una “distracción” que distorsiona el propósito de la educación reduciéndola simplemente a medidas orientadas a resul-

tados y lleva el centro de atención del aprendizaje de los estudiantes a la obtención de una buena posición en los rankings de las ILSA. Uno de los expertos encuestados afirmó: “los instrumentos globales de medición están obstaculizando los esfuerzos de reforma porque los países parecen haber desarrollado una fijación con su clasificación en la liga en vez de con una discusión sólida en torno a qué es la calidad educativa.” Haciéndose eco de ese mismo argumento, otro encuestado señaló:

Los instrumentos de evaluación global tienden a anquilosar la conversación sobre lo que es importante. No digo que no sean valiosos, pero no deberían dominar los debates sobre la eficacia de la educación, pues están dando mucho “juego” y hay muchas otras dimensiones importantes del desarrollo humano a las que la educación puede afectar.

Y lo que es más importante, algunos expertos señalaron que la dependencia de las ILSA/GLM tiende a imponer mundialmente una agenda de reforma educativa neoliberal, que foca en la *performativity*<sup>17</sup> o capacidad de desempeño y en resultados en lugar de en un proceso de educación y en si ese proceso es justo o no. Mientras contribuyen al avance de la agenda de reformas neoliberales, las ILSA/GLM contribuyen simultáneamente a eliminar opciones de políticas alternativas. Como afirmó uno de los encuestados, las ILSA/GLM “evitan que las naciones desarrollen su creatividad y valores cooperativos asociados al cuidado de la infancia... [en vez de eso] la infancia se está volviendo globalmente cada vez menos democrática y abierta, lo que está convirtiendo a los niños en adultos emprendedores.” En otras palabras, los críticos de las ILSA/GLM enfatizan que las mismas reducen las perspectivas y discusiones sobre las políticas educativas a un conjunto predeterminado de objetivos y resultados, lo que alienta una agenda neoliberal de reformas educativas.

#### Las ILSA/GLM como ventaja y como obstáculo

Es destacable que casi un cuarto de todos los encuestados (24%, n=54) opinó que las ILSA/GLM pueden constituir al mismo tiempo tanto una ventaja como un obstáculo para los esfuerzos nacionales de reforma educativa. Conforme uno de los encuestados afirmó, las ILSA/GLM pueden “ser demasiado imperfectos para fundamentar firmes recomendaciones sobre las políticas, pero pueden ser de algún uso para señalar posibles áreas de aciertos y vulnerabilidades”. De forma similar, otros expertos señalaron que las ILSA/GLM son útiles en su función de establecer parámetros, pero no determinaciones en políticas locales, tal como afirmó un encuestado: “son un importante parámetro externo, pero no ofrecen mucha información sobre las necesidades locales y las perspectivas y valores de los educadores locales en relación a las reformas educativas.” Está implícita en varias de esas respuestas la crítica a cómo las ILSA/GLM se usan a nivel nacional, la cual señala que el potencial de las ILSA/GLM se desbarata por la falta de capacidad por parte de los legisladores nacionales de usar estos instrumentos de evaluación de forma adecuada. En particular, diversos expertos encuestados señalaron explícitamente que una de las razones para el uso inadecuado de las ILSA/GLM es la falta de una “alfabetización en el informe PISA” entre los legisladores. En otras palabras, esos expertos concluyen que el problema central es la limitada capacidad técnica institucional a nivel nacional y local, antes que el impulso político.

En un tono más crítico, algunos expertos opinaron que la discusión sobre las ILSA/GLM no se puede ver reducida a cómo hacer y comparar evaluaciones. Al contrario, aquellos con intereses en el ámbito de la educación que hacen uso de las ILSA/GLM deben considerar continuamente las dimensiones políticas porque “la distribución de la educación está basada en el poder.” Asimismo, otro experto encuestado señaló que uno de los principales efectos de las ILSA/GLM parece ser “el fortalecimiento de una distribución desigual de poder y prestigio educativo y un reordenamiento del mapa cultural mundial con el ascenso de valores confucianos.” Sin embargo, otros encuestados señalaron que las ILSA/GLM podrían quizás ser útiles si se usaran como una de otras muchas formas de medida para entender el desarrollo educativo. A pesar de que una diversificación de evaluaciones y medidas basada en dimensiones de valor geopolítico y educacional no significaría un sistema de parámetros “pulcro”, podría abrir posibilidades para un uso de estas evaluaciones más innovador pedagógicamente y más contextualmente relevante en el ámbito nacional y global de esfuerzos de reforma educativa.

## Discusión

La participación en las ILSA/GLM ha generado la percepción de que esas evaluaciones tienen un impacto profundo en las políticas educativas, pero es difícil establecer esa relación causal directa (PAINE et al., 2016). Nuestra investigación confirma que algunos países tienen los mismos resultados en las ILSA, pero implementan políticas diferentes, y viceversa. No obstante, no se pueda afirmar definitivamente que estas convergencias y reacciones políticas estén explícitamente vinculadas a la participación y desempeño en las ILSA/GLM. Nuestra investigación ha pretendido documentar hasta qué punto el uso de las ILSA/GLM ha influido en las políticas para la educación primaria/secundaria y, hemos verificado que la forma en que la comunidad educativa piensa sobre el uso de estas evaluaciones difiere significativamente en función de los diversos contextos.

Todos los países participantes tienen una larga historia de reformas en política educativa que se extiende más allá de su participación en las ILSA/GLM. En los Estados Unidos, por ejemplo, el cambio del gobierno hacia políticas que enfatizan la rendición de cuentas de profesores, escuelas y Estado basada en el rendimiento de los estudiantes en exámenes estandarizados semejantes a las ILSA se ha atribuido históricamente a la publicación del informe *A Nation at Risk* (“Una nación en riesgo”) en 1983. Sin embargo, la evaluación nacional bianual del progreso educativo (NAEP) del Ministerio de Educación estadounidense antecedió a tal informe en 20 años. Con la importancia adquirida por las ILSA a lo largo de las últimas dos décadas (por el informe PISA principalmente), el desempeño nacional en estas evaluaciones se ha convertido en el nuevo indicador para justificar reformas como la de las políticas *No Child Left Behind* y *Race to the Top*. Cuando se observa la política educativa nacional históricamente, se convierte en un desafío determinar qué reformas o políticas se pueden atribuir a qué acontecimientos. Cada país participante en las ILSA tiene una historia compleja, y la participación en las ILSA/GLM se convierte en un hilo más



en la intrincada red de movimientos políticos preexistentes y emergentes. En suma, a pesar de la firme retórica sobre influencia, todas estas políticas son contingentes contextualmente y tal vez sean mejor estudiadas a través de estudios de caso detallados por país.

Dicho esto, nuestra revisión de literatura y los resultados de nuestras encuestas indican que las ILSA/GLM han hecho posible que los gobiernos conciban nuevos modelos de política educacional y los legitimen (BREAKSPEAR, 2014; LAWN & GREK, 2012; VERGER, 2014). Según Baird et al. (2011), las ILSA/GLM han contribuido menos a diferenciar políticas distintas que a fomentar la convergencia política. El uso extendido de rankings y comparaciones internacionales, aparejado a la cobertura mediática de los informes sobre los resultados de las ILSA/GLM, constituye una dinámica clave en los debates nacionales sobre sistema educativo (HAAS & FISCHMAN, 2010). Nuestra revisión identifica una tendencia clara según la cual las autoridades educativas a nivel nacional se ven presionadas a mostrar mejoras en los resultados de las ILSA y en los rankings de las GLM y los gobiernos han empezado a buscar estrategias para subir hasta posiciones más altas (BIEBER & MARTENS, 2011; ENGEL, 2015; MEYER & BENAVIDES, 2013). A pesar de los factores contextuales que afectan inexorablemente cómo los estudiantes se desempeñan en las diversas ILSAs, la competición internacional ha incentivado un sentido de mentalidad descontextualizada de “lo que funciona” (LEWIS, 2016) con la que los gobiernos intentan emular políticas y prácticas de los países con los mejores resultados (ENGEL, 2015; LINGARD, RAWOLLE & TAYLOR, 2005; PHILLIPS & OCHS, 2003). Sin embargo, como defienden Lewis and Lingard (2015), estas “evaluaciones del informe PISA ‘supuestamente’ objetivas y descontextualizadas se interpretan de formas que reflejan los contextos políticos y culturales locales en los cuales se reciben esos datos, lo que subraya la forma contingente en la que se entiende el conocimiento producido por el informe PISA dentro de los sistemas educativos nacionales, a pesar de su circulación global” (p. 627).

Nuestra investigación presenta evidencias robustas sobre las grandes limitaciones de los análisis y perspectivas que establecen un vínculo directo entre los resultados de las ILSA (o los rankings de las GLM) y los cambios en las políticas educativas. Pero lo que podemos reconocer es que las ILSA/GLM posibilitan un campo de comparación pretendidamente objetivo que puede, como de hecho lo hace, transformar radicalmente las condiciones y contextos globales que afectan a las políticas educativas. Como resultado de ello, los gobiernos enfrentan nuevas presiones para conseguir mejores clasificaciones y, consecuentemente, para garantizar que sus políticas se adapten a ese objetivo. Pese a quien pese, cada vez es más evidente que las ILSA/GLM tienen una profunda influencia en la forma como los involucrados en la educación en todo el mundo están tomando decisiones sobre políticas educativas y agenda de reformas, sirviendo para identificar los resultados que se encajan en narrativas previamente establecidas sobre reformas educativas adecuadas.

Quizás el hallazgo más significativo asociado con el uso de ILSAs en la literatura que revisamos es la forma en que las nuevas condiciones para la comparación

educativa han sido posibles a nivel nacional, regional y global. Por un lado, las ILSAs tienen el potencial de proporcionar a los gobiernos y diferentes actores educativos información comparable que supuestamente permiten la evaluación de los logros educativos tanto dentro de las ciudades, estados y regiones, como entre países. Por otro lado, analizar los resultados de ILSAs desde perspectivas “idiotas”, sin considerar la fuerte influencia de las desigualdades en las oportunidades educativas o reconocer agendas políticas o económicas más amplias que impulsan la producción y el uso de ILSA en educación- es bastante peligroso.

Cuando los gobiernos y/o diferentes actores educativos se esfuerzan más en generar narraciones simplistas para legitimizar políticas usando ILSAs, ignorando las diferencias contextos escolares y los múltiples obstáculos asociados a esos contextos, mostrando una falta de preocupación por las desigualdades en oportunidades educativas y los derechos de millones de estudiantes de familias pobres, y enfocando todas las energías en justificar sus propias opiniones, mientras descartan rápidamente cualquier evidencia contraria para legitimar intereses y beneficios, es una forma genuina de idiotez educativa.

¿La mejor defensa contra la idiotez educativa? No existe un antídoto que funcione perfectamente, pero proponemos que para reducir la idiotez educativa precisamos de dos estrategias complementarias y un principio fundamental. La primera estrategia, consiste en discutir, experimentar, evaluar y considerar las múltiples evidencias existentes y generar otras formas de evidencias que reconozcan la heterogeneidad de los contextos educativos. Segundo, el sector educativo tiene que evitar la dependencia exclusiva de medidas cuantitativas simplistas para determinar los resultados educativos, desviar la atención de las estrategias a corto plazo diseñadas para escalar rápidamente las clasificaciones de ILSA, implementar estrategias comprobadas para reducir las desigualdades en las oportunidades de mejorar los resultados a largo plazo.

En cuanto al principio, debemos evitar que las reformas educativas sean justificadas y legitimadas en una sola medida de la supuesta calidad educativa. Debemos considerar múltiples tipos y fuentes de datos, necesitamos explorar formas más significativas de informar datos comparativos, debemos reconocer la importancia de los propósitos cívicos y públicos de la educación, y debemos involucrar a nuestras diversas comunidades – las familias, docentes, sindicatos de educación, administradores, líderes comunitarios y estudiantes - en un diálogo público sobre qué es y qué debe ser la educación. La superación de la idiotez de la educación implicaría, por lo tanto, examinar las diferentes realidades educativas, atendiendo la diversidad de las poblaciones y sobre todo con una mirada que tome en cuenta las desigualdades, así como las múltiples relaciones entre los sistemas educativos, factores ambientales, económicos, sociales y culturales, que, por supuesto no deberían ignorar los resultados de las evaluaciones internacionales a gran escala, pero evitando la idiotez en sus usos.

## Referências

- APPLE, M. Can critical pedagogies interrupt rightist policies? **Educational Theory**, 50, 2003, p. 229-254. doi: 10.1111/j.1741-5446.2000.00229.x.
- BALL, S. J. The teacher's soul and the terrors of performativity. **Journal of Education Policy**, 18, 2003, p. 215-228. doi: 10.1080/0268093022000043065.
- BIESTA, G. J. J. Against learning. Reclaiming a language for education in an age of learning. **Nordisk Pedagogik**, 24, 2004, p. 70-82.
- BIESTA, G. J. J. The learning democracy? Adult learning and the condition of democratic citizenship. **British Journal of Sociology of Education**, 26, 2005, p. 693-709. doi: 10.1080/01425690500293751.
- BIESTA, G. J. J. **Beyond learning: Democratic education for a human future**. Boulder, CO: Paradigm Publishers, 2006.
- BIESTA, G. J. J. Good education in an age of measurement. **Educational Assessment, Evaluation and Accountability**, 21, 2009a, p. 33-46. doi: 10.1007/s11092-008-9064-9.
- BIESTA, G. J. J. What kind of citizenship for European Higher Education? Beyond the competent active citizen. **European Educational Research Journal**, 8, 2009b, p. 146-157. doi: 10.2304/eej.2009.8.2.146.
- BIESTA, G. J. J. **Good education in an age of measurement: Ethics, politics, democracy**. Boulder, CO: Paradigm, 2010.
- BIESTA, G. J. J. How useful should the university be? On the rise of the global university and the crisis in higher education. **Qui Parle: Critical Humanities and Social Sciences**, 20(1), 2011, p. 35-47. doi: 10.1353/qui.2011.0040
- BIESTA, G. J. J. Interrupting the politics of learning. **Power and Education** 5(1), 2013, p. 4-15. doi: 10.2304/power.2013.5.1.4
- BIESTA, G. J. J. **The beautiful risk of education**. Boulder, CO: Paradigm Publishers, 2014.
- BOGOTCH, I., Mirón, L., & Biesta, G. Effective for what; Effective for whom? Two questions SESI should not ignore. En T. Townsend (Ed.), **International handbook of school effectiveness and school improvement**, Dordrecht/Boston: Springer, 2007, p. 93-110.
- CHARLTON, B. G. The ideology of 'accountability'. **Journal of the Royal College of Physicians of London**, 33, 1999, p. 33-35.
- CHARLTON, B. G. Audit, accountability, quality and all that: The growth of managerial technologies in UK Universities. En S. Prickett, & P. Erskine-Hill (Eds.), **Education! Education! Education! - Managerial ethics and the law of unintended consequences**. Exeter: Imprint Academic, 2002, p. 13-28.
- DAVIS, A., & White, J. Accountability and school inspection: In defence of audited self-review. **Journal of Philosophy of Education**, 35, 2001, p. 667-681. doi: 10.1111/1467-9752.00252
- EPSTEIN, D. Defining accountability in education. **British Educational Research Journal**, 19, 1993. 243257.
- GEWIRTZ, S. **The managerial school**. Post-welfarism and social justice in education. Londres/Nueva York: Routledge, 2002.
- GRANGER, D. No Child Left Behind and the spectacle of failing schools: The mythology of contemporary school reform. **Educational Studies**, 43, 2008, p. 206-228. doi: 10.1080/00131940802117654
- GREEN, A. **Education and state formation**. Londres: Macmillan, 1990.
- HENRY, G. T. Choosing criteria to judge program success: A values inquiry. **Evaluation**, 8, 2002. 182204.
- HESS, F. M. Accountability without angst? Public opinion and No Child Left Behind. **Harvard Educational Review**, 76, 2006, p. 587-610.
- HOPMANN, S. No child, no school, no state left behind: Schooling in the age of accountability. **Journal of Curriculum Studies**, 40, 2008, p. 417-456. doi: 10.1080/00220270801989818
- HOUSE, E. R., & Howe, K. R. **Values in evaluation and social research**. Thousand Oaks, CA: SAGE. 1999.
- KUMASHIRO, K. K. **Bad teacher! How blaming teachers distorts the bigger picture**. Nueva York: Teachers College Press, 2012.

- MARQUAND, D. **Decline of the public**. Cambridge: Polity Press, 2004.
- MARTENS, K., Nagel, A.-K., Windzio, M., & Weymann, A. (Eds.). **The transformation of education policy**. Nueva York: Palgrave/Macmillan. 2010.
- MICHIE, J. **The handbook of globalisation**. Cheltenham: Edward Elgar. (2nd ed.), 2011.
- NICOLAIDOU, M., & Ainscow, M. Understanding failing schools: Perspectives from the inside. **School Effectiveness and School Improvement**, 16, 2005, p. 229-248. doi: 10.1080/09243450500113647
- O'NEILL, O. **A question of trust**. Recuperado de <http://www.bbc.co.uk/radio4/reith2002>. BBC Reith lectures, 2002.
- PHILLIPS, D. Policy borrowing in education: Frameworks for analysis. En J. Zajda et al. **International handbook on globalisation, education and policy research**: Part 1, Eds, Dordrecht: Springer, 2005, p. 23-34.
- PHILLIPS, D., & Ochs, K. **Educational policy borrowing**: Historical perspectives. Didcot: Symposium. Eds., 2004.
- POULSON, L. Accountability: a key-word in the discourse of educational reform. **Journal of Education Policy**, 11, 1996, p. 579-592.
- POULSON, L. Accountability, teacher professionalism and education reform in England. **Teacher Development**, 2, 1998, p. 419-432.
- POWER, M. **The audit explosion**. Londres: Demos, 1994.
- POWER, M. **The audit society**: Rituals of verification. Oxford: Oxford University Press. 1997.
- RIZVI, F., & Lingard, B. **Globalizing education policy**. Londres/Nueva York: Routledge. 2009.
- SCHWANDT, T., & Dahler-Larsen, P. When evaluation meets the 'rough' ground' in communities. **Evaluation**, 12, 2006, p. 496-505. doi: 10.1177/1356389006073745
- TOMLINSON, S. Sociological perspectives on failing schools. *International Studies in Sociology of Education*, 7, 1997, p. 81-98.
- TOWNSEND, T. **International handbook of school effectiveness and school improvement**. Dordrecht/Boston: Springer, Ed. 2007.
- TRÖHLER, D. Harmonizing the educational globe. World polity, cultural features, and the challenges to educational research. **Studies in Philosophy and Education** 29(1), 2010, p. 5-17. doi: 10.1007/s11217-009-9155-1
- USHER, R. Lyotard's performance. **Studies in Philosophy and Education**, 25, 2006, p. 279-288. doi: 10.1007/s11217-006-9009-z
- WALDOW, F. What PISA did and did not do: Germany after the 'PISA-shock' (Review Essay). **European Educational Research Journal**, 8(3), 2009, p. 476-483. doi: 10.2304/eej.2009.8.3.476
- WEBB, P. T. The evolution of accountability. **Journal of Education Policy**, 26(6), 2011, p. 735-356. doi: 10.1080/02680939.2011.58753

## Notas

<sup>1</sup> Traducción al español de Carlos Maroto Guerola.

<sup>2</sup> PISA – a cada 3 años; PIRLS – a cada 5 años; TIMSS – a cada 4 años; LLECE – con periodicidad indeterminada; ASER - anualmente (en el ambiente familiar); EGRA/EGMA – con periodicidad indeterminada/varía en función del país

<sup>3</sup> Por ejemplo, los Estados Unidos participan en las evaluaciones PISA, PIRLS y TIMSS, mientras que Filipinas sólo ha realizado la evaluación TIMSS, particularmente en 1999 y 2003 (IEA, 2017; OECD, n.d.).

<sup>4</sup> 43 países participaron en la evaluación PISA en 2000, por ejemplo, en comparación con los 76 que participaron en 2015 (OECD, n.d.).

<sup>5</sup> Cabe señalar que los estudios de la IEA se han usado principalmente para investigación por parte de los especialistas de cada país mientras que la OCDE desempeña un agresivo papel como asesor de políticas educativas e intérprete de sus propios datos.

<sup>6</sup> Como parte de la administración responsable del informe ASER en la India

<sup>7</sup> A través de su trabajo con el UIS y con los países participantes de la alianza.

<sup>8</sup> Hasta la fecha, no se han aceptado estándares consensuales de desempeño ni hay exámenes pactados que garanticen que las métricas y las medidas de los países participantes sean comparables entre sí ni a lo largo del tiempo. El marco de indicadores de los objetivos de desarrollo sostenible recomiendan tres parámetros específicos para medir el progreso global de los resultados de aprendizaje: “Porcentaje de niños y jóvenes a) en los grados 2º/3º, b) al final de la enseñanza primaria, y c) al final del primer ciclo de la enseñanza secundaria que hayan alcanzado al menos un nivel mínimo de competencia en I) lectura y II) matemáticas”.

<sup>9</sup> Los que están a favor de las GLM las ven como una forma de mejorar el desarrollo educativo de los estudiantes (LMTF, 2013; SCHLEICHER, 2009; WINTHROP & SIMONS, 2013), mientras que sus críticos argumentan que ese tipo de medidas suelen resultar limitadas para obtenerse una comprensión provechosa del progreso educativo de un país (ARY, JACOBS, SORENSEN & RAZAVICH, 2006; Education International, 2011). Además, entre los críticos hay otros cuya mayor preocupación está relacionada a los sistemas de rendición de cuentas y a la comercialización de la educación pública (GORUR, 2016; VERGER, 2008; VERGER et al., 2012).

<sup>10</sup> Cabe señalar que en nuestra revisión de literatura no han aparecido referencias que traten de la resistencia de gobiernos nacionales a las evaluaciones a gran escala.

<sup>11</sup> Ocho de los artículos fueron publicados en 2017 como parte de la edición especial sobre la influencia del programa PISA a nivel nacional e infranacional del *European Journal of Education*.

<sup>12</sup> La perspectiva para entender la base discursivo-ideológica de la legitimación la estamos tomando del Análisis Crítico del Discurso (ACD) (FAIRCLOUGH, 1989; VAN LEEUWEN, 2008; WODAK, 2004). Dentro de esa fundamentación, entendemos la idea de legitimación como la creación de un sentido de acción positivo, beneficioso, ético, comprensible, necesario o de algún otro modo aceptable en un contexto específico, mientras que deslegitimación significa establecer un sentido de acción o estado general de cosas negativo, moralmente censurable o de alguna otra forma inaceptable (VAARA, 2014, p. 503). Véase también el trabajo de Gil et al. (2016) sobre el uso del informe PISA como una herramienta política de legitimación en España.

<sup>13</sup> Como se discute en Lingard y Lewis (2016), el federalismo en los Estados Unidos es complejo e idiosincrático, una vez que los estados y municipios siempre tuvieron un control sustancial sobre la estructura y los contenidos de sus sistemas en los niveles educativos preuniversitarios hasta que en los noventa la ley No Child Left Behind (literalmente “ningún niño se queda atrás”) y, más actualmente, la ley Every Student Succeeds (“todos los estudiantes tienen éxito”) amplió la supervisión e influencia federal.

<sup>14</sup> Véase la tabla 2 para una lista de puntos que se incluyen entre los propósitos instrumentales.

<sup>15</sup> Incluye respuestas de encuestas tanto con los expertos cuanto con los no expertos.

<sup>16</sup> Incluye respuestas de encuestas tanto de los expertos cuanto de los no expertos.

<sup>17</sup> Hacemos uso de la comprensión de Gorur (2016) de “performativity” en el sentido que se usa en la teoría actor-red y en estudios de ciencia y tecnología de forma más general (ver, por ejemplo, LAW, 2009; KNORR-CETINA, 1999; PICKERING, 1993), que postulan que ejercicios como mediciones “no son tan solo representaciones o descripciones de realidades existentes sino que participan en su producción.” (p. 205).

## Agradecimientos

En este artículo se presentan los resultados de un proyecto de investigación con apoyo financiero de la Open Society Foundation. Queremos agradecer la colaboración de Jessica L. Holloway y Janna Goebel en el proyecto de investigación inicial. Le estamos muy agradecidos también a los revisores anónimos del JEP y al Dr. David C. Berliner, Dr. Pasi Sahlberg y Dra. Stephanie McBride-Schreiner por sus valiosos comentarios, que tanto nos ayudaron a refinar nuestros argumentos y análisis.

## Correspondência

**Gustavo E. Fischman** – Arizona State University, Tempe, AZ 85281, EUA.

*E-mail:* fischman@asu.edu.

Recebido em 4 de março de 2018

Aprovado em 25 de julho de 2018



This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0)