

Revista Interamericana de Educación para la Democracia

RIED  **IJED**

Interamerican Journal of Education for Democracy



**El arrastre de PISA:
Incertidumbre, influencia e
ignorancia**

Vol 3, No. 1

Abril, 2010

Documento disponible en:
www.ried-ijed.org

ISSN: 1941-1799

El arrastre de PISA: Incertidumbre, influencia e ignorancia

Sharon Murphy
York University

Resumen:

Si las tasas de participación son un indicador, el Programa para la Evaluación Internacional de Alumnos (PISA, por sus siglas en inglés) patrocinado por la Organización para la Cooperación y el Desarrollo Económicos (OCDE) está aumentando su influencia y la OCDE planea mayores expansiones en el alcance de PISA. Tras presentar una discusión sobre algunos problemas de validez relacionados con PISA, se postula un número de razones que justifican la popularidad o “el arrastre” de PISA, entre las que se encuentran: la transferencia desde otras jurisdicciones ante las incertidumbres que plantea la globalización, el apalancamiento político que pueden obtener los estados-nación al poder compararse con otros tanto en términos de PISA como económicamente, y el discurso compartido que rodea a PISA, el cual es producto de la profesionalización de la evaluación educativa. Se argumenta que la influencia de PISA puede poner en riesgo la democratización de la política educativa en la medida en que permite que las élites avancen sus propias agendas con mínimo aporte del público.

En el año 2000 la Organización para la Cooperación y el Desarrollo Económicos lanzó la primera ronda de PISA –el Programa para la Evaluación Internacional de Alumnos (PISA, por sus siglas en inglés) (OECD, s.f.). A diferencia de otras evaluaciones internacionales de gran escala tales como el Estudio Internacional de Tendencias en Matemáticas y Ciencias (TIMSS por su sigla en inglés), PISA se distingue por su conexión con la OCDE, lo cual le ofrece a esta evaluación la maquinaria de la OCDE y, como tal, garantiza mayor protagonismo para esta prueba en el escenario mundial. A medida que su influencia crece, las señales indican que la OCDE está contemplando extender el alcance de PISA a poblaciones más jóvenes y, posiblemente, a otros aspectos de la educación (Governing Board, 2005), diferenciando así las evaluaciones que se han engendrado en el marco de PISA de cualquier otra evaluación internacional de gran escala.¹ Así mismo, al incrementar rápidamente el número de países participantes en un plazo

relativamente corto, se está construyendo un mensaje implícito de credibilidad en el instrumento de evaluación. La consecuencia de esta credibilidad es que la “ventaja comparativa” o ‘desventaja comparativa’ de cada sistema puede ser determinada y ser utilizada política y económicamente” (Steiner-Khamsi, 2004, p. 207).

Los 28 países que conformaban entonces la OCDE, incluyendo los Estados Unidos, Canadá y México, participaron en la primera administración de PISA. Desde aquél momento, la participación se ha más que duplicado y ahora incluye cerca de 60 países. En las Américas, Estados Unidos, Canadá y México, las naciones miembro de la OCDE, continúan participando, pero desde el año 2000 la participación se ha expandido para incluir a Brasil (Krawchuk y Rust, 2002), Uruguay (OECD, 2005), la República Dominicana (OECD, 2007), Argentina, Chile, Colombia, Panamá y Perú (OECD s.f.; OECD, 2009b). En breve, en



sólo en unos cuantos ciclos de aplicación, PISA ha conquistado bastante territorio en las Américas en tal medida que “naciones asociadas” que no son de la OCDE –naciones que no necesitan participar—se sienten impulsadas a hacerlo. ¿Cuál es el arrastre de PISA y cuáles son las consecuencias de participar en la evaluación?

En este artículo planteo que las motivaciones para participar en PISA son complejas y a veces contradictorias y que la naturaleza y las consecuencias de dicha participación están soportadas por la formación no democrática de las políticas educativas. Inicio mi argumentación discutiendo primero la credibilidad de PISA como instrumento de evaluación, dado que otros planteamientos se basan en la premisa de la “metodología meticulosa” de PISA (Steiner-Khamsi, 2004, p. 207). Sustentaré que PISA es un “instrumento muy maniobrado” que en últimas privilegia ciertos tipos de conocimiento y que, a pesar de tener una fachada de invitar a la participación en su diseño, efectivamente la limita. Después examinaré el atractivo de PISA como una solución de política pública para los estados-nación utilizando teorías sobre cómo reaccionan las organizaciones ante condiciones ambientales similares, así como sobre la manera en que se dan y toman en préstamo las políticas educativas. A manera de conclusión destacaré los elementos comunes que unen el diseño de PISA y los usos de PISA por los estados nación en términos de la erosión de principios democráticos básicos.

Una introducción a PISA

Aunque PISA podría ser considerado uno más en una lista creciente de mecanismos internacionales de evaluación como el Estudio Internacional de Tendencias en Matemáticas y Ciencias (TIMSS), o el Estudio Internacional de Progreso en Comprensión Lectora (Progress in

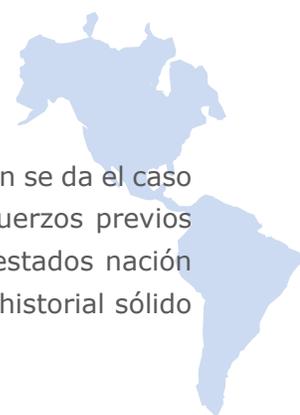
International Reading Literacy – PIRLS), PISA se distingue por ser patrocinado por la OCDE, una organización que define su misión así:

Agrupar los gobiernos de países comprometidos con la democracia y la economía de mercado alrededor del mundo para:

- Apoyar el desarrollo económico sostenible
- Incrementar el empleo
- Elevar los niveles de vida
- Mantener la estabilidad financiera
- Apoyar el desarrollo económico de otros países
- Contribuir al crecimiento del comercio mundial (OECD, 2009a).

La OCDE (2009c) describe el propósito de PISA de la siguiente manera:

¿Están preparados los estudiantes para los desafíos del futuro? ¿Pueden analizar, razonar y comunicarse efectivamente? ¿Tienen la capacidad para continuar aprendiendo a lo largo de sus vidas? El Programa para la Evaluación Internacional de Alumnos (PISA) de la OCDE da respuesta a éstas y otras preguntas mediante su estudio de jóvenes de 15 años en los principales países industrializados. Cada tres años, el estudio evalúa qué tanto han adquirido los estudiantes que se acercan a finalizar su educación obligatoria de ciertos conocimientos y habilidades esenciales para participar plenamente en la sociedad. (p. 1)



Las características de PISA

La descripción de PISA está diseñada para preocupar y tranquilizar, y hace referencia a las ansiedades que enfrentan los individuos y los estados nación sobre la vida en la sociedad contemporánea. La promesa de PISA se cumple en forma de pruebas de papel y lápiz (sobre comprensión matemática, lectora y científica y, recientemente, solución de problemas) con el formato de preguntas de selección múltiple y respuesta construida (respuesta corta).

Se ha desarrollado un total de 7 horas de material para pruebas para PISA; sin embargo, cada estudiante completa sólo 2 horas de prueba. Pueden hacerse comparaciones entre estudiantes que han respondido preguntas distintas mediante un procedimiento estadístico en el que se asume que un rasgo específico da cuenta de las respuestas a los ítems de la prueba (Adams, 2002; McQueen y Mendelovits, 2003). Encuestas a los estudiantes y a los directores escolares proporcionan un componente de PISA que busca ayudar a contextualizar los resultados. En 2006 se añadió una encuesta para padres. Las encuestas se enfocan en las características, percepciones y antecedentes de los estudiantes, las características de la escuela, y los antecedentes y percepciones de los padres de familia.

Si bien los datos recogidos por la OCDE para otros propósitos desde 1960 no son un componente de PISA como tal, sí son parte del panorama en el que a menudo se discuten sus resultados. A pesar de que el alcance y la naturaleza de las estadísticas han cambiado recientemente y que ahora se hace énfasis en “la teoría del capital humano, que involucra un vínculo claro entre las competencias y calificaciones de un individuo en el nivel micro y el crecimiento económico en el nivel macro”

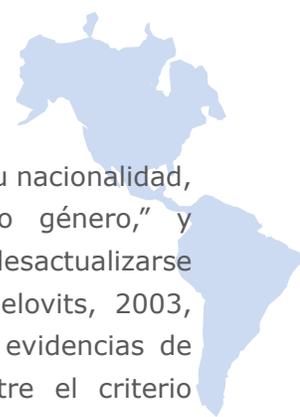
(Ioannidou, 2007, p. 337), también se da el caso de que, sobre la base de sus esfuerzos previos agregando información, muchos estados nación consideren que la OCED tiene un historial sólido de recolección de datos.

La calidad de PISA como herramienta de evaluación

Si PISA fuera una herramienta de evaluación extraordinaria de gran calidad en el contexto del diseño y desarrollo de evaluaciones, ser “asignado un lugar” o ser ordenado por rango como resultado de dicha evaluación ameritaría una seria consideración. Sin embargo, un examen cercano de PISA hecho por académicos independientes de su desarrollo revela, no sólo que la validez de PISA ha sido cuestionada seriamente por varias razones, sino también que algunos de estos problemas de validez generan interrogantes sobre quiénes tienen la oportunidad de participar auténticamente en la formación de las políticas educativas.

Participando en la creación de PISA

Desde la perspectiva de la validez, varias de las preguntas en la prueba de comprensión de PISA son idénticas a la Encuesta Internacional de Alfabetización de Adultos (International Adult Literacy Survey – IALS), lo cual genera la pregunta sobre si PISA es una prueba para adultos o adolescentes. El hecho de que quienes desarrollaron IALS hayan ganado la licitación para el desarrollo de PISA (Rochex, 2006) podría explicar las similitudes generales entre PISA e IALS. De hecho, los desarrolladores se interesaron en comparar los desempeños en IALS y PISA. Sin embargo, el haber insertado en PISA “un subconjunto de las tareas de comprensión de la escala de comprensión de prosa de IALS” (Yamamoto, 2002, n.p; ver también Grisay, 2002, p. 59) hace que uno se pregunte sobre la



idoneidad de los ítems inicialmente desarrollados para los participantes de 16 a 65 años de edad (como en IALS), en lugar de los participantes de 15 años de PISA. Aunque mucho se habla en los documentos de PISA sobre la cantidad de aportes que presuntamente hicieron los estados nación en el desarrollo de los ítems de la prueba, la inclusión de estos ítems de IALS recibe apenas una mención breve en los Manuales Técnicos de PISA (por ejemplo Wu, 2000, p.21) y no se ofrece ninguna evidencia de si los estados-nación objetaron frente a estos ítems. Si uno asume que el que aquellos que se vean afectados por una decisión deban tener la oportunidad de participar en esa decisión es un principio democrático básico, entonces este principio parece ser “maniobrado” e incluso frustrado cuando los estados nación envían sus representantes a participar en las reuniones sobre el diseño de PISA.

El contenido de PISA

El conocimiento que es valorado dentro de PISA es un tema que genera varios problemas de validez y varios problemas democráticos fundamentales. Tres aspectos distintos son de interés en cuanto a la validez de PISA: a) las áreas del conocimiento que se piensa que requieren los jóvenes de 15 años para “participar plenamente en la sociedad,” b) las formas en que el lenguaje representa ese conocimiento, y c) las formas en que ciertos aspectos de las culturas son escogidos o no para ser representados como la base de las preguntas de la prueba PISA.

El conocimiento requerido por los jóvenes de 15 años. Varios problemas emergen al pensar las maneras en que el conocimiento que se presume que necesitan los jóvenes de 15 años aparece en PISA. En primer lugar, se afirma que los materiales de las pruebas de PISA han sido diseñados para ser “idóneos y de interés para

jóvenes de 15 años sin importar su nacionalidad, cultura, nivel socioeconómico o género,” y que se supone que no deben “desactualizarse rápidamente” (McQueen y Mendelovits, 2003, p. 214). Rochex (2006) reportó evidencias de que había un desencuentro entre el criterio de la “idoneidad e interés” y el de la “no desactualización rápida,” al observar que los estudiantes se mostraban “desconcertados” por la “naturaleza irrelevante o contra-intuitiva de ciertos puntos de vista” (p.189) en una tarea de lectura. Si, como lo dijeron quienes lo desarrollaron, una de las motivaciones de PISA es ser “de avanzada” (Adams, 2002, p. 15), basar parte de la prueba para jóvenes de 15 años en preguntas pasadas hechas para gente entre los 16 y los 65 años resulta contradictorio. Se añade a esta complejidad el que la visión del mundo de los jóvenes de 15 años puede estar en contradicción no sólo con la de la población del IALS sino también con la de quienes desarrollaron la prueba. Adicionalmente, esta visión parece sugerir que poco desarrollo ocurre durante la vida. Este tema es discutido mínimamente en documentos de PISA. Aún así, problemas como éste sí parecen hacer una diferencia. Por ejemplo, cuando se entrevistó a una muestra de participantes en PISA sobre preguntas individuales en la prueba, tuvieron dificultades para suspender sus propios puntos de vista a favor de aquellos que se habían incorporado en la prueba (Rochex, 2006).

El conocimiento de los jóvenes de 15 años se circunscribe de otra manera por el criterio de la matrícula en la educación obligatoria. Aún así, es desconcertante que quienes desarrollaron la prueba comenten repetidamente que no les interesa “qué tanto estos estudiantes dominen un currículo escolar específico,” mas sí “lo que los estudiantes pueden hacer con lo que aprenden en la escuela” (Adams, 2002, p. 15). Sin discutir qué tipos de situaciones pueden

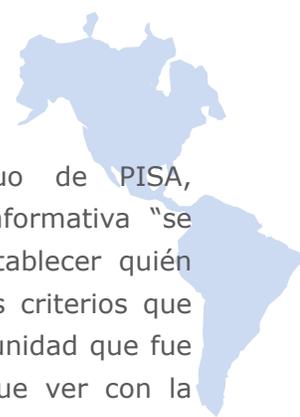


encontrar los jóvenes de 15 años en el mundo, ni las bases sobre las que se seleccionaron estas situaciones, resulta razonable cuestionar qué es exactamente lo que evalúa PISA. De hecho, en el comentario de Dohn (2007) de que “PISA evalúa, con cierto grado de confiabilidad, conocimiento y habilidades para PISA. Nada más, nada menos” (p. 10) está implicada la sugerencia de que se crea un “mundo PISA” a través de estas pruebas.

Algunos podrán decir que aún si se crea un mundo PISA, el desempeño en las pruebas PISA provee indicadores comparables de desempeño estudiantil dentro de ese mundo artificial. Sin embargo, acá debemos mirar el procedimiento estadístico que soporta algunas de las inferencias sobre los alcances estudiantiles. PISA utiliza un procedimiento común en el desarrollo de pruebas conocido como la Teoría de Respuesta por Ítems (item response theory – IRT). La IRT se basa en el supuesto de que “un sólo rasgo latente da cuenta de las respuestas a los ítems de una prueba... [basado en] un modelo de cómo los examinados con distintos niveles de habilidad para ese rasgo deberían responder a un ítem” (Crocker y Algina, 1986, p. 339). De esta manera puede comprarse el desempeño de los estudiantes que han tomado distintas versiones de la prueba, como es el caso de PISA. Sin embargo, cuando Rochex (2006) examinó las respuestas de los estudiantes para ver si habían relaciones predecibles basadas en la similitud del desempeño de los estudiantes por ítems, subpruebas, o incluso partes de las pruebas, estas relaciones fueron “difíciles de predecir” (p. 185). Esta falta de relaciones predecibles sugiere que hay más de un rasgo en operación en los ítems en PISA, y hace cuestionar el supuesto de que los estudiantes que toman distintas versiones están tomando “la misma” prueba.

No sólo se asume que formas diferentes de la prueba son equivalentes utilizando IRT, sino que también se han hecho preguntas sustantivas –y no equivalentes- sobre si la prueba evalúa lo que pretende evaluar. Por ejemplo, cuando Rochex (2006) entrevistó a estudiantes sobre su desempeño, encontró que estos hablaban de poder contestar algunas preguntas sin leer los pasajes en la prueba. Estas posibilidades efectivamente remueven la “lectura” de una prueba de lectura. Cualquier inferencia sobre la lectura de pasajes para este tipo de ítems es sospechosa y la prueba de lectura se convierte en una prueba de conocimientos contextuales y experiencia.

El lenguaje y la representación del conocimiento. La documentación de PISA es generalmente autocomplaciente en lo que se refiere a los obstáculos superados en relación con el tamaño de la tarea de la traducción de ítems. Consideremos, por ejemplo, que al final, se utilizaron 47 versiones lingüísticas distintas, entre ellas 7 variantes de Inglés, 4 de Francés, 4 de Alemán, y 5 de Español (Grisay y Monseur, 2007). Los modismos y los matices se evitaron. Algunas palabras o fraseos fueron reemplazados por los locales (por ejemplo, moneda, nombres de lugares, nombres de personas, la sintaxis, la ortografía), y se les pidió a los países participantes que calificaran los ítems con base en una serie de características. Las pruebas fueron revisadas de acuerdo con estas valoraciones. Los desarrolladores de PISA enfatizan los “procedimientos de doble traducción” en cuanto a que son más precisos que otras alternativas, pero también reconocen el impacto que los problemas de lenguaje pueden tener el desempeño (Grisay, 2002). Por ejemplo, las pruebas PISA son cronometradas y se basan en una lectura de la versión en Inglés a razón de 3500 palabras por hora (McQueen y Mendelovits, 2003). El manual técnico de PISA 2000 señala



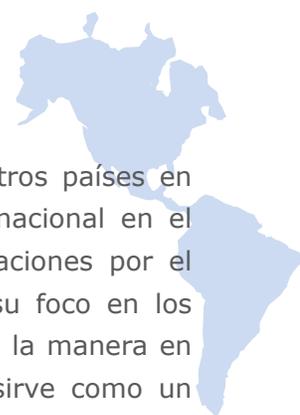
que, en algunos casos, los textos se extendieron en longitud como resultado de la traducción y si bien la carga adicional que se impuso sobre quienes tomaban la prueba para terminarla a tiempo no “parecía sustancial... la hipótesis de que esto tuvo un efecto en el desempeño de los estudiantes no puede descartarse” (Grisay, 2002, p. 66).

Si, como defiende Rochex (2006), la traducción es clave para poder hacer inferencias comparativas razonables sobre PISA, ¿qué impacto tiene la traducción sobre los resultados de PISA? Un estudio comparativo de idiomas indo-europeos y no indo-europeos, por ejemplo, arroja que “muchas preguntas de selección múltiple no funcionan de maneras equivalentes en sus versiones [para países asiáticos], comparado con las versiones de la prueba en idiomas occidentales” (Grisay y Monseur, 2007, p. 77). Estos hallazgos sugieren que, la clasificación de PISA de los países que difieren sobre la base del lenguaje bien puede ser un ejercicio sin sentido.

La cultura y la representación del conocimiento. Si el idioma es problemático para los ítems de prueba de PISA, también lo es el tema menos discutido de la representación cultural. A pesar del propósito de representar la diversidad cultural más que la neutralidad cultural en los componentes de la prueba, ésta meta fue difícil de alcanzar. Durante la etapa de desarrollo de PISA, se les pidió a los representantes de los países involucrados en la evaluación PISA que comentaran sobre la diversidad cultural con respecto a la prueba. El resultado fue que “no todas las unidades que se mantuvieron en el conjunto de preguntas eran favorecidas por todos los países, pero cada preocupación que se expresó fue explorada cuidadosamente y se tomó una decisión sobre la base de toda la información disponible” (McQueen y Mendelovits, 2003, p. 216).

Ese acercamiento oblicuo de PISA, caracterizado por la frase no-informativa “se tomó una decisión,” dificulta establecer quién hizo el juicio y cuáles fueron los criterios que lo soportan. Un ejemplo de una unidad que fue considerada inapropiada tenía que ver con la cantante alternativa independiente Ani DiFranco. El que esta unidad haya sido considerado inapropiada porque “el feminismo, el objeto de la letra de la canción, era un concepto desconocido para estudiantes de ciertos países” (McQueen y Mendelovits, 2003, p. 216), hace que uno se pregunte cómo se imaginó la diversidad cultural en general. Con sólo unos pocos ejemplos de las preguntas que fueron eliminadas de la prueba PISA, poco sentido se le puede encontrar a las preocupaciones de países individuales frente a ítems específicos o frente a la prueba como conjunto. Aún así, la escasa información provista por los desarrolladores de PISA frente a este aspecto del diseño de la prueba genera más preguntas de las que responde sobre los factores culturales que influyeron en él. Ante este tipo de problemas, los resultados de la prueba PISA deberían considerarse “evidencia frágil” (Murphy, Shannon, Johnston y Hansen, 1998), en el mejor de los casos, o debieron sufrir de una “muerte por mil cortadas” pues las múltiples limitaciones diferentes asociadas con la prueba, al agregarse, deberían implicar que la prueba no puede defenderse como unidad. Sin embargo, PISA ha logrado prevalecer a pesar de sus limitaciones. PISA, con su:

diseño borroso... trata los vínculos entre estudiante, escuela, y el logro nacional, como algo que es evidente por sí mismo; permitiendo así que se de una aproximación a la escolaridad como una caja negra, donde la coincidencia de resultados y factores se transforma



en correlaciones y causalidades, sin establecer cómo se materializa esta linealidad. (Hopmann, 2008, p. 444)

En el mejor de los casos, en las discusiones sobre PISA, las fallas técnicas son de poco interés para muchos, pero en el peor son ignoradas por la mayoría. No obstante, en el núcleo de estos problemas de validez residen preguntas fundamentales sobre si el proceso de toma de decisiones responde a principios democráticos. Si las pruebas de PISA tienen contenido basado en supuestos erróneos sobre grupos generacionales, tienen posturas contradictorias sobre el papel de la escuela, similitudes insuficientes entre preguntas que supuestamente miden el mismo rasgo, tareas que pueden no requerir que los participantes hagan uso de la principal destreza que está siendo evaluada, variaciones en el idioma en el que se imprime la prueba, y elementos culturales altamente selectivos, entonces se están tomando decisiones que privilegian a algunos y marginalizan o excluyen a otros. Así mismo, la estructura y la complejidad del proyecto PISA es tal que los participantes no parecen tener canales de acción cuando enfrentan una serie de decisiones que, vistas en conjunto, van a representar a los jóvenes de 15 años de su país de maneras infundadas. ¿Por qué participan entonces los estados-nación en PISA cuando no están obligados a hacerlo? Una respuesta puede tener su origen en las preocupaciones que tienen los estados-nación sobre su futuro.

La participación en PISA como una respuesta a la incertidumbre

Además de alabar el sistema de evaluación mismo, el sitio web de PISA explora preocupaciones, incluyendo preocupaciones sobre el futuro y la preparación adecuada de los estudiantes, preocupaciones implícitas sobre

la preparación en relación con otros países en el mundo (por el contexto internacional en el que opera la OCDE), y preocupaciones por el desarrollo económico mediante su foco en los países industrializados. Como tal, la manera en que la OCDE posiciona a PISA sirve como un ejemplo de isomorfismo mimético, en el que “la incertidumbre promueve la imitación” (DiMaggio y Powell, 1983, p. 151). Las fuerzas implícitas de la incertidumbre en los propósitos descritos de PISA son la globalización y la economía.

El término “globalización,” en su acepción más simple, hace referencia a los “cambios políticos y culturales... [los cuales] afectan de maneras comunes a grandes segmentos de la población mundial” (Spring, 2008, p. 331). Las fuerzas globalizantes siempre han existido. Sin embargo, de acuerdo con Dale (2003), estas fuerzas son preocupantes en contextos contemporáneos por varias razones: a) anteriormente, mediante el colonialismo y el imperialismo, la globalización se daba un país a la vez, pero ahora la globalización “no puede reducirse a las intenciones de una nación” (p. 94); b) los intereses del estado-nación ya no son supremos; por el contrario “la supervivencia del sistema es más importante que los intereses de naciones individuales” (p. 94); y c) “el estado sólo puede dar curso a sus responsabilidades nacionales priorizando sus responsabilidades extra-nacionales” (p. 95). Los estados-nación han sido puestos en la posición poco envidiable de perder el control estando aparentemente en control, una situación que sin dudas gesta incertidumbre. Ante la incertidumbre, una es las respuestas de las organizaciones es modelar o tomar prestadas prácticas de otras jurisdicciones (DiMaggio y Powell, 1983, p. 151; Steiner-Khamsi, 2004), lo cual, a su vez, podría explicar la participación de países que no son miembros de la OCDE en PISA.



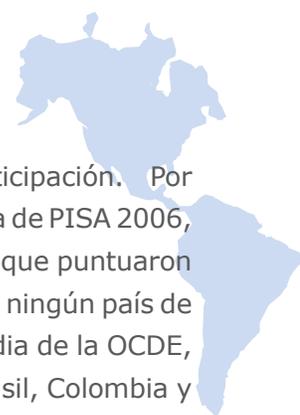
Al involucrarse en estos préstamos, los estados participantes ceden parte de su soberanía interna (Robertson, 2009), creando así “una forma constreñida de democracia en la que las decisiones ya no son tomadas por los representantes electos” (Kelly, 2009, p. 52). Las restricciones sobre la toma de decisiones se hacen crudamente visibles en las descripciones del consejo gobernador de PISA y en las reuniones de proyectos nacionales. Los estados-nación participantes envían representantes a estas reuniones fundamentalmente para que participen en las decisiones sobre PISA (incluyendo decisiones sobre el diseño general). Si bien no existen reportes publicados para participantes de las Américas, entrevistas conducidas por Grek, Lawn y Ozga (2009) con participantes escoceses e ingleses en estas reuniones revelan candidamente cómo son manejados los países por la OCDE a través de PISA. Citamos extensamente a Grek y otros (2009) porque sus descripciones son bastante persuasivas:

[En] las reuniones del Consejo... los miembros de cada país a menudo aparecen para representar sus ‘estereotipos’ nacionales y para promover reconocimiento ‘nacional’... [L]a tradición y los valores educativos son los estereotipos que se viven más repetidamente en las reuniones de PISA de la OCDE... De una manera, las reuniones del Consejo eran descritas como el lugar en donde se “alisaban” todas las diferencias y las tradiciones nacionales para poder alcanzar un consenso. No obstante, también se anotaba muy interesantemente que las ideas que se presentaban eran aquellas que tenían más posibilidad de conducir al acuerdo entre los miembros. Por ende, las reuniones eran descritas como

muy manejadas y controladas por el Secretariado de la OCDE y el mismo Andreas Schleicher.

En términos de los problemas técnicos discutidos en las reuniones, otro entrevistado comentó que los asuntos técnicos casi nunca se discuten –por el contrario, los expertos de ACER a menudo ofrecen presentaciones técnicas que los miembros del Consejo nunca cuestionan sino en las que ‘confían’. El proceso parece traducir la participación política (por ejemplo, mediante representación de los valores y las prácticas del sistema) en procesos técnicos mediante el ritual de la puesta en escena del estereotipo y de la ‘confianza’ en la experticia técnica para reducir la divergencia y la diferencia a sistemas manejables y comparables. (p. 8)

De hecho, esa “construcción de consenso” considerada un hito de muchos esquemas educativos transnacionales (Ioannidou, 2007) es poco más que un engaño en el contexto de las reuniones descritas por Grek y otros (2009), y las descripciones generan preguntas aún más profundas sobre quiénes están tomando realmente las decisiones sobre PISA. Cuando el seudo proceso de construcción de consenso se superpone al hecho de que las reuniones de PISA se sostuvieron en inglés o francés, los participantes de naciones-estados no anglo-parlantes resultan en peor desventaja, lo cual fue discutido también por los participantes ingleses y escoceses de Grek y otros (2009). Quizás la falla de los representantes de los estados-nación para ser más explícitos sobre la naturaleza de estos encuentros es análoga a la historia del traje nuevo del emperador, en la que el pueblo puede ver que el emperador está desnudo pero



cada individuo teme hablar y desenmascarar su ignorancia aparente. Alternativamente, pueden ofrecerse otras razones para la participación continua de los estados-nación en PISA.

La participación en PISA como una fuente de influencia y legitimidad

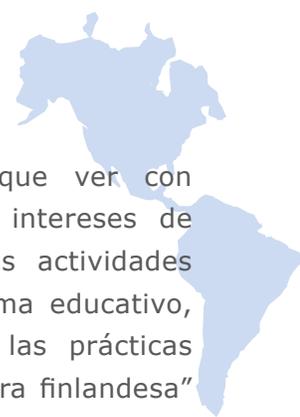
Las preocupaciones por la economía están incrustadas en las preocupaciones por la globalización. Esta relación nace del hecho de que algunos vean que la globalización contemporánea tiene “sus orígenes en factores económicos, sobre todo financieros y de producción” (Dale, 2000, p. 95). Para PISA, las presiones económicas formales e informales de la globalización llevan a una especie de isomorfismo coercitivo (DiMaggio y Powell, 1983, p. 150) para los países no miembros de la OCDE, que se sienten presionados a participar en PISA. La pérdida de autonomía del estado-nación como resultado de la respuesta a esta presión puede tener como efecto la erosión de la soberanía de la nación. Los entrevistados de Grek y otros (2009) confirman que la hipótesis de que la participación en PISA está motivada por las presiones económicas de la globalización tiene mérito sustancial, ya que el vínculo entre crecimiento económico y el rendimiento escolar era exactamente la razón por la que los entrevistados vieron la participación en PISA como más valiosa que la participación en TIMMS u otras evaluaciones. En pocas palabras, “la OCDE pudo ofrecer un margen mucho mayor de comparación, tanto para los sistemas de educación –y por tanto, las economías- más exitosos como para los menos” (Grek et al., 2009, p. 4).

Aunque los estados pueden sentirse coaccionados para participar, los resultados de PISA, cuando se combina con los indicadores económicos, no se sostienen necesariamente frente a las consecuencias que los estados-nación

pueden imaginar para su participación. Por ejemplo, para las pruebas de lectura de PISA 2006, Canadá cayó en el grupo de países que puntuaron por encima de la media de la OCDE, ningún país de las Américas cayó dentro de la media de la OCDE, mientras que Uruguay, México, Brasil, Colombia y Argentina estuvieron por debajo del promedio de la OCDE (OECD, 2007, p. 47). Para las pruebas de matemáticas y ciencias del 2006 estos patrones son casi idénticos, con la salvedad de que los Estados Unidos -que no forma parte de los datos de lectura debido a errores en el folleto- aparece al principio de la lista de países por debajo del promedio de la OCDE (OECD, 2007, pp. 22 y 53). Patrones similares han persistido a través de las distintas aplicaciones de PISA (Vegas y Petrow, 2008).

El Producto Interno Bruto y PISA

El Producto Interno Bruto (PIB) se presenta a veces como una “explicación” para los patrones de los resultados de PISA y, a su vez, puede convertirse en una fuerza motivadora para la participación en PISA con el supuesto tácito de que si los resultados PISA mejoran también lo hará PIB. Sin embargo, el hecho de que una economía grande como la de los Estados Unidos caiga por debajo del rango promedio plantea interrogantes acerca de si el PIB es de hecho una motivación para la participación en PISA. Los reacios a desistir de la asociación entre PIB y el desempeño en PISA ofrecen explicaciones para el desempeño de los Estados Unidos. Por ejemplo, Hopmann (2008) sugiere que los Estados Unidos está muy centrado en la evaluaciones trascendentales (high-stake testing) asociadas con la legislación de Que Ningún Niño se Quede Atrás (No Child Left Behind), por lo que el rendimiento en PISA no es una prioridad importante, mientras que Pelham, Crabtree y Nyiri (2008) indican que la explicación más común para el desempeño



de los Estados Unidos es “la distribución relativamente inequitativa de recursos entre los que tienen y los que no tienen en la sociedad estadounidense. Los niños estadounidenses en zonas prósperas se desempeñan igual de bien que los niños en cualquier parte del mundo, mientras que en los barrios pobres, en promedio lo hacen peor, bajando el promedio general del país” (p. 76). Por otra parte, Smithson (2009) sugiere como explicación para el rendimiento en la porción de matemáticas de PISA que hay un desencuentro entre el contenido entre PISA y el plan de estudios enseñado en los Estados Unidos.

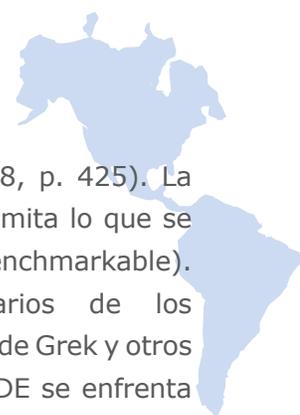
Otras explicaciones para las diferencias en el desempeño a lo largo de las Américas incluyen el hecho de que la inversión en educación de los países latinoamericanos es menor como proporción del PIB que en países con PIB similares; adicionalmente, los puntajes promedio de PISA son inferiores a los que se podrían predecir sobre la base de los niveles de gasto por alumno (Vegas y Petrow, 2008). Documentos de la OCDE, en particular, ofrecen una motivación explícita para la participación en PISA pues enmarcan los resultados de la prueba como mecanismos por medio de los cuales los estados-nación pueden mejorar “el aprendizaje” (por ejemplo, Vegas y Petrow, 2008); sin embargo, el aprendizaje, a su vez, es definido circularmente como una mejora de las puntuaciones de PISA en el futuro, y el PIB no se aborda. Pero ni el PIB ni un nivel de gasto por estudiante inferior a la media cuentan la historia completa de PISA.

Tomemos, por ejemplo, el caso de Finlandia, un país con un PIB razonablemente saludable, que se ha desempeñado rutinariamente bien a lo largo de varias aplicaciones de PISA. El éxito de Finlandia ha sido explicado por “una red de factores

interrelacionados que tienen que ver con una pedagogía abarcante, los intereses de los propios estudiantes y sus actividades de ocio, la estructura del sistema educativo, la formación del profesorado, las prácticas escolares y, en últimas, la cultura finlandesa” (Valijarvi, Linnakyla, Kupari, Reinikainen, y Arffman, 2002, p. 8). Sin embargo, Finlandia se ha visto presionada para mantener su posición de alto rango (Livingston y McCall, 2005) hasta un punto tal que nuevas políticas educativas con énfasis en el individualismo y la competencia –lo opuesto de las condiciones en las que Finlandia alcanzó su éxito (y Kivirauma Ruoho, 2007, p. 298)— son las que ahora se abogan. Al igual que algunos de sus homólogos en las Américas, a pesar de haber obtenido altas puntuaciones, Finlandia también parece haber sido empujada por PISA a tomar prestadas políticas educativas; al responder al arrastre de PISA, Finlandia está en riesgo de ceder parte de su autonomía como estado-nación frente a la educación por tomar en préstamo prácticas educativas en línea con la filosofía subyacente de PISA.

La crisis del Estado de Bienestar y PISA

Una explicación económica diferente para el arrastre de PISA puede encontrarse en las discusiones que describen a PISA como una herramienta para la “gestión de las expectativas” en una “era de rendición de cuentas” (Hopmann, 2008). En la raíz de esta caracterización hay un argumento económico basado en la crisis del Estado de Bienestar, en la que los límites para la provisión de las instituciones sociales están dados por la capacidad del país para pagar por ello. Hopmann (2008) sugiere que la crisis del Estado de Bienestar² cambió cómo los estados-nación manejan el apoyo a su ciudadanía, de la “gestión de las colocaciones” (management of placements) a la “gestión de las expectativas” (management



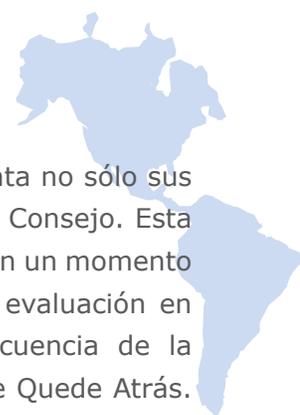
of expectations) (p. 424). En la gestión de las colocaciones, los ciudadanos podían dirigirse a instituciones sociales específicas (educación, salud, etc.) con la expectativa de que sus necesidades fueran cubiertas; los profesionales dentro de estas instituciones determinaban cómo se utilizarían los recursos y la variedad de cuestiones que podían tratar. Al verse confrontado por una creciente expansión de las necesidades identificadas, un tesoro público contraído, y una ciudadanía preocupada por la sostenibilidad de sus soportes sociales, el Estado moderno pasó a gestionar las expectativas de sus ciudadanos. Esta gestión en una variedad de sectores se traduce en parámetros y estándares de prestación de servicio y desempeño que “permiten una gestión más orientada hacia objetivos y la rendición de cuentas lo cual, sin embargo, tiene sus costos: todo lo que no se ajusta a un régimen de expectativas es marginalizado” (Hopmann de 2008, p. 424). PISA y la educación se ajustan bien dentro de un modelo adelgazado del estado de bienestar; pero se sacrifican entonces los propósitos más amplios de la educación tales como la conducta ética, la responsabilidad social, y la ciudadanía, así como el discurso sobre la educación se centra en los indicadores estrechos ejemplificados en PISA.

Según Hopmann (2008), PISA toma la educación –algo que es relativamente grande y difícil de contener– y la transforma en algo que, al menos en apariencia, parece estar bien definido y ser alcanzable. El mecanismo de rendición de cuentas proporcionado por PISA minimiza los factores independientes (por ejemplo, género, clase social, la inmigración) que operan dentro y a través de los sistemas educativos de los estados-nación. A cambio, “los economistas que usan PISA calculan los costos de transacción de la enseñanza y los modos y medios por los cuales los principales (los padres, el Estado) pueden maximizar la efectividad de los agentes elegidos (es decir, maestros, escuelas, o del

sistema escolar)” (Hopmann, 2008, p. 425). La facilidad en la medición también limita lo que se identifica como parametrizable (benchmarkable). Remitiéndonos a los comentarios de los entrevistados escoceses e ingleses de Grek y otros (2009), parece que cuando la OCDE se enfrenta a las esperanzas y los deseos de estados-nación acerca de sus sistemas educativos –esperanzas y deseos que fueron articuladas por los procesos democráticos al interior de esos estados– la OCDE es incapaz de considerar cómo se podrían evaluar estos elementos tan dispares. En cambio, en las reuniones de los coordinadores y del Consejo, la combinación de la gestión por consenso y la gran influencia de la estadística en la formulación de la evaluación, permite que los representantes de los estados-nación escuchen, pero que participen sólo de formas muy restringidas en el desarrollo de PISA.

PISA como un artefacto de la profesionalización de la evaluación educativa

Procedimientos complejos y protocolos estadísticos permean las discusiones sobre las cualidades de PISA como instrumento de evaluación, y a su vez son artefactos del discurso creciente sobre las evaluaciones de gran escala utilizado por los profesionales de la evaluación educativa. La profesionalización del acto de la evaluación en la educación es un ejemplo del “isomorfismo normativo” (DiMaggio y Powell, 1983, p. 150). Dada la profesionalización de la evaluación educativa, un discurso y unas prácticas compartidas operan atravesando fronteras nacionales y manteniendo una influencia normativa en el valor que se le asigna a estas evaluaciones; esto acarrea como consecuencia el que los tomadores de decisiones y los estados-nación vean la participación en dichas evaluaciones como algo meritorio. Para el caso de la educación basada en logros (outcomes-



based education) en el escenario internacional, Steiner-Khamsi (2004) se refiere a este discurso como un tipo de “efecto Maris O’Rourke” (p. 211), en el que las políticas promovidas por O’Rourke cuando ella lideraba el movimiento de educación basada en logros en Nueva Zelanda, la siguieron y permearon el discurso y las políticas sobre evaluación cuando ella pasó a trabajar al Banco Mundial. En países como los Estados Unidos, sin embargo, el discurso en el área de la evaluación educativa ha tenido mucho más tiempo para formarse.

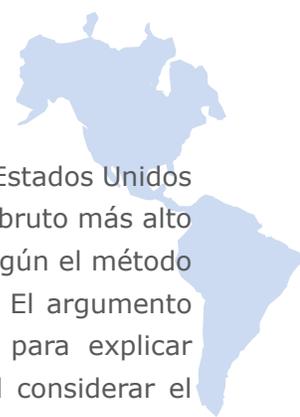
El caso de la evaluación de gran escala en los Estados Unidos

Se ha dicho que en los Estados Unidos hay una cultura de la evaluación (Hanson, 1993), y la evaluación educativa de gran escala ha sido prevalente allí por algún tiempo (Clarke, Madaus, Horn y Ramos, 2000). A pesar de que las fuerzas normalizantes de la profesionalización pueden encontrar cierta resistencia en algunos escritos académicos, es menos común que las voces en el interior de la comunidad de la evaluación educativa se preocupen públicamente o protesten por el uso de pruebas de gran escala los Estados Unidos.

Sin embargo, recientemente, el Consejo sobre Pruebas y Evaluación (Board on Testing and Assessment – BOTA) del Consejo Nacional de Investigación (National Research Council) en los Estados Unidos escribió una carta abierta al fondo Carrera hacia la Cima (Race to the Top) del Departamento de Educación de los Estados Unidos. En esta carta, BOTA (2009) le advierte a la administración del Presidente Obama sobre el uso de mediciones únicas para la toma de decisiones, y promueve un acercamiento más matizado al uso de los resultados de las pruebas. La carta viene firmada por Edward Haertel, un nombre reconocido en el campo de las

pruebas y la medición, y representa no sólo sus sentimientos sino también los del Consejo. Esta carta es importante porque llega en un momento de frenética preocupación por la evaluación en los Estados Unidos como consecuencia de la legislación de Que Ningún Niño se Quede Atrás. Esta carta también es importante porque no viene de los críticos de las pruebas estandarizadas sino de los profesionales mismos que desarrollan pruebas y mediciones.

Los esfuerzos de los profesionales de la evaluación educativa por influir en el ámbito de la política pública no son nuevos. Por ejemplo, hace casi 20 años otro experto eminente en pruebas y mediciones, Robert Linn (1991), escribió un artículo para la Oficina de Evaluación de la Tecnología (Office of Technology Assessment) del Congreso de los Estados Unidos, titulado “Mal uso de pruebas: ¿por qué es tan prevalente?” Y hace unos 40 años, uno de los principales conceptualizadores de la teoría de la validez de pruebas contemporánea, Samuel Messick, fue coautor de un artículo (Messick y Anderson, 1967) que planteó preguntas sobre las consecuencias de las prácticas de uso de pruebas. En breve, numerosos académicos provenientes del campo de las pruebas y la medición han manifestados sus propias preocupaciones dentro de sus propias comunidades de discurso sobre los usos dados a los resultados de las pruebas. Sin embargo, más que nada, la utilización de pruebas estandarizadas ha incrementado exponencialmente en los Estados Unidos en los últimos 40 años (Clarke y otros, 2000). Lo que distingue la carta de Haertel es el hecho de que la carta fue un acto muy público, reportada por *Education Week*, y que no se confinó a una revista académica o a un reporte técnico para una administración gubernamental en el poder. Dadas estas manifestaciones recientes de académicos alertando sobre la evaluación educativa, resulta



algo irónico que la profesionalización del campo de la evaluación educativa sea otra posible explicación para el arrastre de PISA.

La profesionalización de la evaluación como promulgación de la ignorancia

Una explicación adicional para el arrastre de PISA proviene de la teorización reciente adelantada por historiadores de la ciencia en el área de la "agnotología." Simplemente descrita, la agnotología es el estudio de la construcción de la ignorancia (Proctor, 2008). Smithson (2008) plantea que "la ignorancia, al igual que el conocimiento, es construida socialmente en mayor parte" (p. 212). Pensando en la ignorancia como una característica de nuestro mundo social, Proctor (2008) se enfoca en varios tipos distintos de ignorancia: a) ignorancia como "un lugar que el conocimiento no ha penetrado" (p. 4), como cuando una persona simplemente no sabe sobre algo; b) ignorancia como resultado de una "escogencia selectiva" (p. 6), como si fuera el "producto de la desatención" (p.7) que sucede cuando los individuos se enfocan en un aspecto informativo ignorando otro; c) ignorancia como "una maniobra estratégica o una construcción activa" en donde la ignorancia es "producida activamente como parte de un plan deliberado" (p. 9), como cuando se utiliza la desinformación o la duda para guiar a las personas a ver las cosas de un modo particular (como cuando la industria tabacalera norteamericana hace comentarios sugiriendo que fumar no era perjudicial para la salud a pesar de tener evidencia que indica lo contrario).

En el caso de PISA, parece que el segundo tipo de ignorancia, la ignorancia por atención selectiva, opera en distintas maneras. Consideremos por ejemplo dos factores coexistentes: a) Estados Unidos ha tendido a estar por debajo del rango medio en varias

aplicaciones sucesivas PISA, y b) Estados Unidos tenía en 2008 el ingreso nacional bruto más alto de cualquier país en el mundo (según el método Atlas del Banco Mundial [2009]). El argumento de la competitividad económica para explicar el arrastre de PISA se debilita al considerar el desempeño de los Estados Unidos. Este es un ejemplo operante de atención selectiva – los estados-nación no prestan atención a los elementos contradictorios de PISA.

No obstante, Steiner-Khamsi (2000) va aún más lejos al sugerir que el arrastre de las pruebas de gran escala se debe a la atención selectiva. Ella plantea que los estados-nación hacen un uso mucho más deliberado de los resultados de las pruebas al tomar políticas en préstamo, y utiliza el ejemplo de la respuesta alemana a una serie de resultados de pruebas internacionales para ilustrar este punto precisamente. En nombre del logro en las pruebas pueden perpetrarse crisis cuando en otras oportunidades se ignoran silenciosamente los resultados mientras los gobiernos adelantan sus propias agendas. Mientras que en el escenario internacional los gobiernos pueden parecer obligados a participar en PISA, en el nacional están utilizando los resultados de PISA para sus propios propósitos y toman ventaja del manto que les ofrece el discurso de la evaluación educativa para hacerlo. En este escenario se frustra doblemente la oportunidad para la participación democrática; primero mediante el uso de una prueba que, a pesar de su fachada de que promueve la participación de los estados-nación en su desarrollo, parece ser maniobrada y controlada por élites tras bambalinas, y es frustrada nuevamente por los representantes políticos elegidos que toman ventaja de la profesionalización de la evaluación para utilizar la información de las pruebas para satisfacer sus propias agendas políticas.



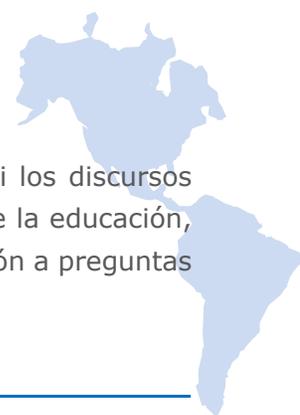
Conclusión

Las evaluaciones de gran escala como PISA son instrumentos con fallas. No son lo suficientemente robustas para sostener el tipo de comparaciones y decisiones de política que fluyen de ellas. Los estados-nación pueden verse presionados por la incertidumbre a participar en estas evaluaciones; pueden encontrar apalancamiento político o económico o influencia en la participación; pueden encontrar profesionales con conexiones burocráticas poderosas que los inducen a la participación. Como sugiere Steiner-Khamsi (2000), los mismos estados-nación pueden utilizar estratégicamente los resultados de las pruebas para avanzar en sus propias agendas. Un movimiento de este estilo parece una respuesta apropiada a los resultados de PISA dada la participación maniobrada que se le permite a los estados-nación en las fases de desarrollo de PISA.

Sin embargo, estas acciones por parte de los estados-nación son altamente similares a las acciones de los oficiales de la OCDE en el desarrollo y las reuniones directivas de PISA. Los estados-nación parecen ser abiertos pero en efecto están reduciendo la transparencia del proceso. Estas acciones permiten la manipulación del público y cierran las oportunidades para la discusión y el debate profundo, cualidades necesarias en cualquier democracia, al tomar ventaja del discurso complejo y en ocasiones confuso creado por la profesionalización de la evaluación. En el desarrollo de PISA y, para algunos estados al menos, en los usos de PISA, la participación es manipulada – para el desarrollo de las pruebas, por la OCDE y los desarrolladores, y para el uso de las pruebas, por algunos de los propios estados-nación. Ambas prácticas callan efectivamente a quienes quieren involucrarse en

una discusión sustantiva sobre los valores y los juicios subyacentes que infunden y emanan de PISA.

Mientras que más investigación que detalle los usos que los estados-nación le dan a PISA sería altamente benéfica y proveería un entendimiento más matizado de PISA y sus usos, las operaciones de la OCDE continúan. Para países grandes como los Estados Unidos, que es tan dominante que puede desplazar su atención de PISA a otros temas, las consecuencias del arrastre de PISA no son importantes, pero para democracias menos afluentes dentro y fuera de las Américas, dos temas deben discutirse. Si dirigir recursos escasos a PISA y las reformas que se originan en la participación en PISA son resultado de los efectos coercitivos de PISA, entonces, siguiendo el ejemplo de BOTA, el arrastre de PISA debería ser una fuente de determinación para los profesionales de la evaluación educativa y para los analistas de políticas públicas para extender su discurso a foros más públicos que las típicas comunidades discursivas académicas. Estos profesionales deben actuar en formas epistemológicamente responsables (Code, 1987, 2006) para educar a todos los involucrados sobre las limitaciones de los instrumentos de evaluación y el uso apropiado de dichos instrumentos para que se posibilite una participación más amplia y genuina. En otras palabras, los profesionales de la evaluación educativa no sólo deben ser sensibles frente a la creación de conocimiento sobre evaluación, sino que deben estar alerta sobre la utilización de este conocimiento. En efecto, tienen una responsabilidad ética respecto al dominio del conocimiento que están creando y dentro del cual trabajan. Si los gobiernos se están involucrando en el uso de los resultados de la evaluación cuando les son convenientes, dichas acciones deben ser discutidas por los



profesionales de la evaluación al igual que por los expertos en políticas educativas. Si los discursos tecnológicos de las evaluaciones educativas son desmantelados por los profesionales de la educación, y la evaluación educativa se torna más transparente, quizás podamos devolver la atención a preguntas más amplias e importantes que saber cómo le va a un país en PISA.

Notas

1. Un informe reciente (Lederman, 2010) indica que el Consejo Australiano para la Investigación Educativa (Australian Council for Educational Research) – el mismo grupo asociado con el desarrollo de PISA – es uno de los dos grupos financiados por la OCDE para diseñar lo que prácticamente es una variante de PISA para la educación superior llamado La Evaluación de Logros de Aprendizaje en Educación Superior (The Assessment of Higher Education Learning Outcomes – AHELO). AHELO evaluará destrezas generales y las disciplinas de ingeniería y economía. Richard Yelland, el director de la División de Administración Educativa e Infraestructura de la OCDE, presenta AHELO como “una herramienta que nos ayudará a ayudar a quienes son responsables por la educación superior en los distintos países” (Lederman, 2010, n.p.).
2. Para una discusión específica sobre América Latina, ver Shugurensky (2003).

Bibliografía

- Adams, Ray. (2002). The programme for international student assessment: An overview. En Ray Adams. y Margaret Wu (Eds.), PISA 2000 technical report (pp. 15-20). Paris, France: OECD. Board on Testing and Assessment. (2009). *Letter report to the U.S. Department of Education on the Race to the Top fund*. Washington: National Academy of Sciences. Obtenida el 28 de Octubre de 2009 en URL <http://www.nap.edu/catalog/12780.html>
- Clarke, Marguerite, Madaus, George, Horn, Catherine, y Ramos, Miguel. (2000). Retrospective on educational testing and assessment in the 20th century. *Journal of Curriculum Studies*, 32(2): 159-181.
- Code, Lorraine. (1987). *Epistemic responsibility*. London: University Press of New England.
- Code, Lorraine. (2006). *Ecological thinking: The politics of epistemic location*. New York: Oxford University Press.
- Crocker, Linda, y Algina, James. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Dale, Roger. (2000). Globalization: A new world for comparative education? En Jurgen Schriewer (Ed.), *Discourse formation in comparative education* (pp. 87-109). Frankfurt: Peter Lang.



- DiMaggio, Paul J., y Powell, Walter W. (1983). The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields. *American Sociological Review*, 48(2): 147-160.
- Dohn, Nina Bonderup. (2007). Knowledge and skills for PISA--Assessing the assessment. *Journal of Philosophy of Education*, 41(1): 1-16.
- Governing Board. (2005). *Longer term strategy for the development of PISA*. 20th Meeting of the PISA Governing Board, Reykjavik, Iceland. OECD Directorate for Education. Obtenido el 31 de Octubre de 2009 en <http://www.pisa.oecd.org/dataoecd/18/1/35519715.pdf>
- Grek, Sotria, Lawn, Martin y Ozga, Jenny (2009). Fabrication, circulation and use of a supra-national instrument of regulation based on knowledge—Education Sector: Production of OECD's 'Programme for International Student Assessment (PISA). Edinburgh: Centre for Educational Sociology, University of Edinburgh. Obtenido el 31 de Octubre de 2009 en http://www.knowandpol.eu/fileadmin/KaP/content/Scientific_reports/Orientation3/PISA.WP12.Scotland.pdf
- Grisay, Aletta. (2002). Translation and cultural appropriateness of the test and survey material. En Ray Adams y Margaret Wu. (Eds.), *PISA 2000 technical report* (pp. 57-70). Paris: OECD.
- Grisay, Aletta, y Monseur, Christian. (2007). Measuring the equivalence of item difficulty in the various versions of an international test. *Studies in Educational Evaluation*, 33(1): 69-86.
- Hanson, F. Allan. (1993). *Testing testing: Social consequences of the examined life*. Berkeley: University of California Press.
- Hopmann, Stefan Thomas. (2008). No child, no school, no state left behind: schooling in the age of accountability. *Journal of Curriculum Studies*, 40: 417-456.
- Ioannidou, Alexandra. (2007). A comparative analysis of new governance instruments in the transnational educational space: A shift to knowledge-based instruments? *European Educational Research Journal*, 6(4): 336-347.
- Kelly, Anthony. (2009). Globalisation and education: A review of conflicting perspectives and their effect on policy and professional practice in the UK. *Globalisation, Societies and Education*, 7(1): 51- 68.
- Kivirauma, Joel, y Ruoho, Kari. (2007). Excellence through special education? Lessons from the Finnish school reform. *International Review of Education*, 53(3): 283-302.
- Krawchuk, Sheila, y Rust, Keith. (2002). Sample design. En Ray Adams y Margaret Wu (Eds.), *PISA 2000 technical report* (pp. 39-56). Paris: OECD.



- Lederman, Doug. (2010, January 28). Measuring student learning, globally. Inside Higher Education. Obtenido el 11 de Febrero de 2010 en <http://www.insidehighered.com/news/2010/01/28/oecd>
- Linn, Robert L. (1991). *Test misuse: Why is it so prevalent?* Washington, DC: Office of Technology Assessment.
- Livingston, Kay, y McCall, Jim. (2005). Evaluation: Judgmental or developmental? *European Journal of Teacher Education*, 28(2): 165-178.
- McQueen, Joy, y Mendelovits, Juliette. (2003). PISA reading: Cultural equivalence in a cross-cultural study. *Language Testing*, 20(2): 208-224.
- Messick, Samuel y Anderson, Scarvia. (1967). Educational testing, individual development, and social responsibility. *The Counseling Psychologist*, 2(2): 80-87.
- Murphy, Sharon con Shannon, Patrick, Johnston, Peter, y Hansen, Jane. (1998). *Fragile evidence: A critique of reading assessment*. Mahwah, NJ: Lawrence Erlbaum.
- OECD (n.d.). *PISA: The OECD programme for international student assessment*. Paris: OECD. Obtenido el 21 de Octubre de 2009 en <http://www.pisa.oecd.org/dataoecd/51/27/37474503.pdf>
- OECD (2005). *PISA 2003 technical report*. Paris: OECD.
- OECD (2007). *Executive summary—PISA 2006: Science competencies for tomorrow's world*. Obtenido el 21 de Octubre de 2009 en <http://www.pisa.oecd.org/dataoecd/15/13/39725224.pdf>
- OECD (2009a). *About OECD*. Obtenido el 20 de Octubre de 2009 en http://www.oecd.org/pages/0,3417,en_36734052_36734103_1_1_1_1_1,00.html
- OECD (2009b). *What's new?* Obtenido el 20 de Octubre de 2009 en http://www.oecd.org/home/0,2987,en_2649_201185_1_1_1_1_1,00.html
- OECD. (2009c). *OECD: Programme for international student assessment (PISA)*. Obtenido el 22 de Octubre de 2009 en http://www.pisa.oecd.org/pages/0,2987,en_32252351_32235731_1_1_1_1_1,00.html
- Pelham, Brett, Crabtree, Steve, y Nyiri, Zsolt. (2008). Gallop presents... technology and education: The power of the personal computer. *Harvard International Review*, Summer: 74-76. Obtenido el 23 de Octubre de 2009 en www.gallup.com/poll/File/122171/TechnologyAndEducation.pdf.
- Proctor, Robert N. (2008). Agnotology: A missing term to describe the cultural production of ignorance (and its study). En Robert N. Proctor y Londa Schiebinger (Eds.), *Agnotology: The making and unmaking of ignorance* (pp. 1-33). Stanford, CA: Stanford University Press.



- Robertson, Susan L. (2009). Editorial: The shifting politics of governance, sovereignty, and citizenship. *Globalization, Societies and Education*, 7(1): 1-3.
- Rochex, Jean-Yves. (2006). Social, methodological, theoretical issues regarding assessment--lessons from a secondary analysis of PISA 2000 literacy tests. *Review of Research in Education*, 30(1): 163-212.
- Schugurensky, David. (2003). Two decades of neoliberalism in Latin America: Implications for adult education. En Stephen J. Ball, Gustavo Fischman, y Silvina Gvartz (Eds.), *Crisis and hope: The educational hopscotch of Latin America* (pp. 45-65). London: Routledge.
- Smithson, John. (2009). *Describing the academic content of PISA mathematics and science item pools*. Madison, WI: Wisconsin Center for Education Research, University of Wisconsin-Madison. Obtenido el 31 de Octubre de 2009 en https://edsurveys.rti.org/PISA/documents/SmithsonNCES_PISA_Conference_Paper_FINAL.pdf
- Smithson, Michael J. (2008). Social theories of ignorance. En Robert N. Proctor y Londa Schiebinger(Eds.), *Agnology: The making and unmaking of ignorance* (pp.209-229). Stanford, CA: Stanford University Press.
- Spring, Joel. (2008). Research on globalization and education. *Review of Educational Research*, 78 (2): 330-363.
- Steiner-Khamsi, Gita. (2000). Transferring education: Displacing reforms. En Jurgen Schriewer (Ed.), *Discourse formation in comparative education* (pp. 155-187). Frankfurt: Peter Lang.
- Steiner-Khamsi, Gita. 2004). Blazing a trail for policy theory and practice. En Gita Steiner-Khamsi (Ed.), *The global politics of educational borrowing and lending* (pp.201-220). New York: Teachers College.
- Valijarvi, Jouni, Linnakyla, Pirjo, Kupari, Pekka, Reinikainen, Pasi, y Arffman, Inga. (2002). *The Finnish success in PISA--and some reasons behind it: PISA 2000*. Jyvaskyla, Finland: Institute for Educational Research, University of Jyvaskyla, Finland.
- Vegas, Emiliana, y Petrow, Jenny. (2008). *Raising student learning in Latin America: The challenge for the 21st century*. Washington, DC: World Bank.
- World Bank. (October, 2009). *Data and statistics: Quick Reference Tables—GNI 2008 Atlas method*. Obtenido el 29 de Octubre de 2009 de <http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20399244~menuPK:1504474~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html#ranking>



Yamamoto, Kentaro (2002). *Estimating PISA students on the International Adult Literacy Survey (IALS) prose literacy scale*. Obtenido en Noviembre de 2008 en http://www.pisa.oecd.org/document/61/0,3343,en_32252351_32236159_33680573_1_1_1_1,00.html.

Wu, Margaret. (2000). Test design and test development. En Ray Adams y Margaret Wu (Eds.), *PISA 2000 technical report*. Paris: OECD.