

MÉXICO EN PISA 2015

Desaciertos de su aplicación por medios computarizados

JOSÉ ALFONSO JIMÉNEZ MORENO

Resumen:

Este artículo analiza algunas de las implicaciones que tiene para los estudiantes mexicanos la aplicación del Programa para la Evaluación Internacional de Alumnos (PISA) a través de computadoras, que se lleva a cabo desde el ciclo 2015. El texto presenta un análisis de evidencias sobre la información que esta forma de aplicar la prueba genera para el país, su viabilidad en términos de infraestructura, la inversión destinada a ello y el riesgo de obtener resultados influidos por un bajo nivel de conocimiento informático por parte de los estudiantes. Se concluye la falta de viabilidad para el uso de esta modalidad de aplicación frente a los pocos beneficios que trae para el sistema educativo mexicano.

Abstract:

This article analyzes some of the implications for Mexican students when the Programme for the International Student Assessment (PISA) uses computerized testing, first implemented in 2015. The study presents an analysis of evidence regarding the information that the computerized test generates for Mexico, its viability in terms of infrastructure, investment in the test, and the risk of obtaining results influenced by students' limited knowledge of computers. The conclusion points to the lack of viability for the computerized test and its limited benefits for Mexico's educational system.

Palabras clave: evaluación de estudiantes; evaluación de pruebas; educación y computación; infraestructura educativa; desempeño del estudiante.

Keywords: student evaluation; evaluation of testing; education and computers; educational infrastructure; student achievement.

José Alfonso Jiménez Moreno: profesor-investigador de la Universidad Autónoma de Baja California, Instituto de Investigación y Desarrollo Educativo. Carretera Tijuana-Ensenada Km. 103, 22830, Ensenada, Baja California, México. CE: jose.alfonso.jimenez.moreno@uabc.edu.mx

Introducción

La evaluación de la educación es una acción ampliamente utilizada en la actualidad por los sistemas educativos, instituciones, escuelas y los docentes mismos. Su uso se fundamenta en la necesidad de retroalimentación de los avances educativos, así como con fines de rendición de cuentas y como apoyo a las políticas educativas (Jiménez, 2016). Su objeto de análisis puede orientarse a programas, instituciones, resultados de aprendizaje, desempeño docente e incluso a los sistemas educativos en su conjunto. Dentro del ámbito de la evaluación, el desarrollo y uso de los instrumentos de medición han sido tema de diversas investigaciones a nivel mundial, debido a la importancia que tiene la medición de los constructos educativos como fundamento en la emisión de resultados del proceso evaluativo.

La medición de constructos educativos a través de una modalidad por computadora (en inglés se suele utilizar el término *Computer Based Assessment* o sus siglas: CBA) se ha desarrollado fuertemente en las últimas décadas en diversos países (Rodríguez, 2005; Ramírez, 2006; Csapó, Ainley, Bennet, Latour *et al.*, 2012; Piaw, 2012; López, Sanmartín y Méndez, 2013). Particularmente, se le ha dado un papel primordial en procesos de evaluación a gran escala, al punto que se ha considerado que esta modalidad sustituirá a la *evaluación tradicional*, en lápiz-papel, la cual se da en la mayoría de los escenarios de evaluación educativa (Csapó *et al.*, 2012).

El movimiento que sustenta el uso de medios electrónicos en evaluación se fundamenta en la idea de que la tecnología, en particular el uso de computadoras, es un catalizador del cambio en la educación (Piaw, 2012) pero, además, por las ventajas que ofrece para medir habilidades más complejas e, incluso, por las demandas sociales asociadas con la tecnología (Olea, Abad y Barrada, 2010). Por supuesto, en sí misma la tecnología no genera un cambio positivo de la educación y el aprendizaje, pero su inclusión en este terreno es, sin duda, una realidad que debe analizarse en aras que se espera facilite cambios en la naturaleza del aprendizaje (Panel on Educational Technology of the President's Committee of Advisors on Science and Technology, 1997, en Csapó, *et al.* 2012).

El uso de las computadoras con fines de evaluación de procesos educativos, particularmente el aprendizaje, ha estado relacionado con el aprovechamiento de los recursos técnicos que facilitan procesos operativos y

de calificación (Jornet, González-Such y García-Bellido, 2012), así como la reducción de costos de desarrollo (Rojas, 2001). Desde hace más de dos décadas, este modo de evaluar el aprendizaje ha virado de una forma experimental a ser uno de los medios más populares. Algunos de los elementos que mayormente han potencializado el uso de la CBA en educación han sido, entre otros, la facilidad que ofrece en la estandarización de procesos de aplicación, favorece la emisión de calificaciones automatizadas y, además, permite ofrecer formatos de reactivos o ítems que favorecen la medición de habilidades que no necesariamente son de fácil observación por medios de lápiz-papel (Barberá, 2016; Olea, Abad y Barrada, 2010; Csapó *et al.* 2012).

El uso de computadoras en evaluaciones a gran escala se ha dado en todo el mundo bajo diversas condiciones. Considerando la información de la investigación de Csapó *et al.* (2012), en Australia gran parte de la población de estudiantes tiene acceso a tecnologías de la información desde casa, lo cual ha facilitado la implementación de este tipo de evaluaciones. En Asia, las condiciones de infraestructura tecnológica entre escuelas tienen altas variaciones, por lo que la implementación de una plataforma única de evaluación en una región es prácticamente imposible. Los países europeos han sido más consistentes en el uso de la CBA, dado que la búsqueda de evaluación a través de medios tecnológicos es cada vez más demandada en diversos países de la región; incluso se han desarrollado distintas herramientas, como el caso de la *Open Source Assessment Platform* (TAO),¹ plataforma electrónica altamente utilizada tanto en Europa, como en países de otras regiones. En el caso de América, específicamente en Estado Unidos, hay una gran diversidad de instancias que desarrollan evaluaciones mediante el uso de herramientas tecnológicas. Por supuesto, los proyectos internacionales de evaluación a gran escala que utilizan las tecnologías como parte de su estrategia de aplicación no son la excepción, tal es el caso de *Trends in International Mathematics and Science Study* (TIMSS), *Progress in International Reading Literacy Study* (PIRLS), *Computer and Information Literacy* (ICILS) y el *Programme for International Student Assessment* (PISA).

En el caso de México, existen algunos antecedentes de evaluaciones informatizadas en el ámbito educativo. Uno de los más reconocidos es el Examen de Habilidades y Conocimientos Básicos (EXHCOBA), que, desde los años noventa, ha sido un instrumento altamente utilizado por escuelas y

universidades del norte del país con fines de evaluar el ingreso a los estudios profesionales. Este proyecto sigue vigente y, a la fecha, utiliza una plataforma electrónica que permite incluso construir patrones desarrolladores de reactivos (Tirado, Backhoff y Larrazolo, 2011). Por su parte, también el Centro Nacional de Evaluación para la Educación Superior (Ceneval) aplica algunos de sus exámenes de ingreso y egreso a nivel profesional por medios computarizados.²

Dentro de los ejercicios de CBA más relevantes resalta el caso de PISA. Este proyecto, coordinado por la Organización para la Cooperación y Desarrollo Económicos (OCDE), es una de las evaluaciones a gran escala con mayor impacto político y mediático desde su primera aplicación en el año 2000. Su influencia en la política educativa es tan notoria que la mayoría de los países participantes se han fundamentado de alguna forma en esta prueba para sus procesos de reformulación curricular (Taut y Palacios, 2016). PISA se enfoca en la medición de *literacy* de Matemáticas, Ciencias y Lectura en estudiantes de 15 años; se aplica cada tres años y en cada uno de ellos enfatiza la medición de un área en particular. En el caso del ciclo 2015 la evaluación se enfocó en Ciencias; pero, además, incluyó un área nueva: Solución de problemas en colaboración; lo que fue posible gracias al uso de la CBA (INEE, 2016b), y es que, como se apuntaba previamente, esta forma de presentación de las evaluaciones permite la medición de elementos que no pueden considerarse en el formato lápiz-papel (Olea, Abad y Barrada, 2010; Csapó *et al.*, 2012), que es lo que justamente sucedió con PISA.

El uso de la CBA en PISA comenzó desde 2006 con el área de Ciencias; la versión informatizada de la prueba se aplicó a países que decidían participar bajo esta modalidad (Björnsson, 2008; Csapó *et al.*, 2012; Jornet, González-Such y García-Bellido, 2012; INEE, 2016b). Desde aquella aplicación se han realizado algunas críticas respecto de las posibles diferencias existentes entre quienes presentan la prueba mediante el uso de computadoras y de aquellos que la realizan en lápiz-papel; incluso, los resultados de PISA en algunos países (particularmente Dinamarca, Islandia y Corea) han presentado diferencias de resultados en función de si los estudiantes realizaron la prueba en una modalidad u otra (Björnsson, 2008).

A partir de la aplicación de PISA 2015, México decidió sumarse a la propuesta de aplicar la prueba en su versión CBA. El Instituto Nacional

para la Evaluación de la Educación (INEE) es la instancia encargada de coordinar la adaptación, aplicación y emisión de resultados de este proyecto en México y, a partir de ese año, asumió la responsabilidad de dejar atrás la prueba en lápiz-papel. La modificación de la forma de aplicación de este instrumento de carácter internacional puede tener diversas implicaciones a diferentes niveles. Frente a esta situación, el presente artículo tiene por objetivo identificar algunas de las implicaciones que tiene para México la aplicación por computadora de PISA. La intención es aportar a la reflexión de los resultados de esta prueba y sobre las implicaciones que puede tener el uso de la tecnología en este tipo de proyectos.

Para cumplir con este objetivo se presenta el análisis de diversas evidencias obtenidas a través de fuentes oficiales y de recursos bibliográficos que orientan las implicaciones que son materia de este trabajo, mostrando a partir de ello algunos problemas sobre la decisión de aplicar PISA mediante medios informáticos en México. Se abordan temáticas circundantes al proyecto en el sentido de: *a)* la información que genera para el país; *b)* su viabilidad, dadas las características de infraestructura de las escuelas mexicanas; *c)* la inversión destinada a esta modalidad de aplicación; y *d)* aspectos técnicos respecto de las consecuencias de medición bajo la forma de CBA.

Para la presentación de estos cuatro problemas que rodean la aplicación por computadora de PISA 2015 en México, la discusión se orienta en tres apartados: *a)* la descripción de la aplicación en México, que incluye el análisis de la decisión de comenzar a usar computadoras y las implicaciones que ello tiene en términos de información disponible para el país, *b)* el análisis sobre viabilidad del proyecto en términos de infraestructura e implicaciones de gasto, y *c)* el posible impacto del uso de computadoras en los resultados del desempeño de los estudiantes participantes.

La participación de México en PISA y la reducción de información en el ciclo 2015

México ha participado en PISA desde su primer ciclo, en 2000. Esta prueba comenzó con el fin de ser un insumo de política pública y es un proyecto resultante de la intención de contar con información sobre desempeño de estudiantes, ello como forma complementaria al *Education at a Glance* (Baird, Isaacs, Johnson, Stobart *et al.*, 2011), que la OCDE publica con fines de retroalimentar los sistemas educativos de sus países miembros.

La idea central de proveer información sobre avances de los estudiantes es dar a los Estados herramientas frente los retos de la economía global (Fernández-González, 2015).

PISA evalúa *literacy*, es decir, la manera en que los estudiantes hacen uso de “los conocimientos y habilidades en temas clave, y analizar, razonar y comunicar con eficacia a medida que se identifican, interpretan y resuelven problemas en una amplia variedad de situaciones” (INEE, 2016b:12). Como es bien sabido, la evaluación se enfoca fundamentalmente en tres áreas que se consideran clave para el desarrollo de los estudiantes en la sociedad actual: Ciencias, Matemáticas y Lectura (junto con Solución de problemas en colaboración, para el caso del ciclo 2015); pero, además, se incluyen algunos cuestionarios de contexto que permiten obtener información sobre factores y variables relacionadas con los resultados. En el caso de México, en el ciclo 2015 se aplicaron los siguientes cuestionarios de contexto:

- *Cuestionario del estudiante*: enfocado en identificar elementos relevantes sobre su vida académica, así como sobre sus expectativas académicas y sus antecedentes socioeconómicos. Además, se preguntaron temas relacionados con la enseñanza y el aprendizaje de las Ciencias en su escuela.
- *Cuestionario escolar*: instrumento dirigido a los directores de las escuelas participantes. Permitió obtener información sobre la estructura, la organización y el clima escolares, así como elementos generales de la administración y del personal docente.
- *Cuestionario de familia*: centrado en recolectar información sobre la percepción que los padres tienen respecto de la escuela, el apoyo que se da en casa para el aprendizaje, las expectativas académicas de los hijos y sus antecedentes sociales.
- *Cuestionario de familiaridad con las tecnologías de la información y la comunicación (TIC)*: dirigido a estudiantes, con el fin de indagar sobre la disponibilidad y el uso que dan a las TIC, así como las habilidades para realizar actividades en computadora.

Para hacer inferencias del desempeño en *literacy* de los estudiantes de 15 años que asisten a la escuela de cada uno de los países participantes, PISA usa muestras que oscilan entre 4,500 y 10,000 personas (INEE, 2016b),

aunque este número aumenta si el país llega a utilizar sobremuestras con fines de obtener información a mayor detalle de alguna región en particular. México lo usó entre los ciclos 2003 y 2012, con el fin de dar resultados para cada una de las 32 entidades federativas que conforman el país. Justamente para el ciclo 2015 no se hizo uso de sobremuestra, por lo que solo se pudieron dar resultados a nivel nacional. Entre los ciclos 2012 y 2015 el número de estudiantes participantes bajó de 33,806 a 7,568 (INEE, 2016b).

La baja de participación de estudiantes no es en sí misma un problema, lo es la falta de información disponible para el país, ya que durante 12 años seguidos se contó con los datos de desempeño de los jóvenes de 15 años para cada entidad, permitiendo comparaciones detalladas entre cada una de ellas, así como el análisis de avances entre ciclos. La falta de consideración de la sobremuestra en el ciclo 2015 se debió a que el costo financiero de realizar esta aplicación por computadora a nivel nacional era muy alto, dado que se tenía que asegurar que cada estudiante participante contara con una computadora para poder presentar la prueba; por lo tanto, se optó por aplicar únicamente el número de escuelas necesario para dar representatividad nacional, como en el caso de PISA 2000 (INEE, 2016b).

En ese sentido, un primer problema relacionado con la aplicación de PISA por computadora en México refiere a la valoración primordial del uso de la tecnología como medio de administración de pruebas sobre la posibilidad de información a detalle de todas las entidades que conforman el país. Frente a ese problema, es menester considerar que, en 2015, de los 72 países participantes en PISA, 57 decidieron aplicar la prueba por computadora, el resto lo hizo en la modalidad lápiz-papel (Albania, Argelia, Argentina, Georgia, Indonesia, Jordania, Kazajistán, Kosovo, Líbano, Malta, Moldavia, República de Macedonia, Rumania, Trinidad y Tobago y Vietnam). Incluso, si bien la aplicación por computadora fue opcional, se consideró “implícitamente como obligatoria, dada la limitación de la modalidad impresa, que únicamente incluye los reactivos ancla de las tres áreas, y con tendencia a desaparecer a mediano plazo” (INEE, 2016b:28).

Por supuesto, el discurso del INEE es preocupante, en el sentido de que parece que la decisión de participar en la versión CBA de PISA responde a una obligación implícita, aun a costa de perder la información detallada de las 32 entidades. Si bien su informe sobre PISA 2015 (INEE, 2016b)

permite todavía valorar el desempeño de los estudiantes mexicanos de 15 años de edad en *literacy* a nivel nacional, cabe recordar que cada una de las 32 entidades federativas cuenta con secretarías de Educación locales que pueden utilizar como insumo de toma de decisiones políticas los resultados de una prueba como PISA. Afortunadamente, aún se mantienen ejercicios de evaluación nacional que permiten tener indicadores de desempeño de estudiantes en cada entidad, como el Plan Nacional para la Evaluación de los Aprendizajes (Planea), desarrollado por el propio INEE.

El uso de sobremuestras en México en el proyecto PISA reflejaba, sin duda, el interés del Estado por contar con información del resultado de los estudiantes mexicanos en todo el país, considerándola valiosa aun a pesar de las diferencias culturales por región, dada la diversidad existente en el país. Por supuesto, resulta preocupante reconocer la preferencia del uso de herramientas electrónicas como medio de aplicación sobre la posibilidad de tener información de todas las regiones de un país. Además, de esta pérdida, existen otros factores que ponen en duda la viabilidad de esta decisión, mismos que se exponen a continuación.

Aplicación de PISA por computadora: infraestructura e inversión

En lo que respecta a la operatividad de PISA 2015, sin duda el uso de computadoras para cada estudiante representaba un reto relevante en términos logísticos y de infraestructura. Para asegurar la aplicación adecuada de la prueba era necesario instalar 21 equipos de cómputo externos a los centros educativos, dado que no era posible asegurar que hubiera computadoras funcionales dentro de las aulas o que el posible equipo con el que contarán las escuelas tuviera las características necesarias de *software* para ejecutar la aplicación de la prueba.

Al respecto, y dado que PISA se enfoca en estudiantes de 15 años, en México dicha población suele distribuirse entre educación secundaria y media superior; en 2015, la participación de ambos niveles fue de 30.7% y de 69.3%, respectivamente (INEE, 2016b). Con el fin de asegurar la estandarización de las condiciones de aplicación, independientemente del nivel de estudios de la escuela seleccionada, el INEE decidió contratar a una empresa de tecnología que llevara *laptops* a cada una de las 275 escuelas participantes (INEE, 2016b).

Respecto de la temática de infraestructura, el propio INEE en su estudio *Evaluación de las Condiciones Básicas para la Enseñanza y Aprendizaje* (ECEA) (INEE, 2016a) valora el estado de infraestructura básica de las escuelas mexicanas. En su estudio de 2014 (publicado en 2016) encontró que 56.9% de las primarias del país no tienen acceso a, al menos, una computadora por estudiante. La situación se agrava al referir únicamente a escuelas de carácter público, que pueden alcanzar 93% sin acceso a este equipo en aquellas clasificadas como comunitarias. Por su parte, 89.3% de los establecimientos privados sí cuentan con al menos una computadora por escuela. Adicionalmente, los datos del *Panorama Educativo de México* (INEE, 2016c) muestran que solo 69.1% de las secundarias en el país tienen al menos una computadora. Si bien el estudio ECEA se enfocó en escuelas primarias,³ la información que arroja, junto con la correspondiente al *Panorama Educativo de México*, orienta la falta de infraestructura informática en las escuelas del país.

Si bien habrá que esperar los resultados detallados del estudio ECEA del nivel medio superior y secundaria para hacer afirmaciones más precisas, los datos disponibles sobre infraestructura muestran un segundo problema respecto de la aplicación de PISA por computadora en México. Este no es menor, ya que la falta de estructura informática en las escuelas es evidente, particularmente en las públicas, cuestión que trasciende la viabilidad de la aplicación de una prueba internacional y evidencia las brechas de inequidad aún existente en el país, lo cual pone en riesgo el derecho de los infantes y jóvenes mexicanos de recibir educación. Pero, en lo particular en el tema de PISA, sin duda la consideración de aplicar la prueba por computadora en México no tomó en cuenta la capacidad de infraestructura del país de hacer frente a tal desafío.

Por otra parte, la contratación de una empresa de tecnología, que asegurara la instalación de *laptops* con determinadas características de *software* en las 275 escuelas participantes en todo el país, implicó una inversión importante, en virtud de la premisa de la falta de infraestructura necesaria para la aplicación de la prueba en las escuelas seleccionadas. Al respecto, de acuerdo con los datos de transparencia en términos de contratos y convenios realizados por el INEE en el primer trimestre del año 2015, la aplicación definitiva de PISA de ese año demandó el gasto que se puede observar en la tabla 1.

TABLA 1
Gasto destinado para la implementación logística de PISA 2015 en México

Concepto	Monto en pesos mexicanos
Memorias USB	630,663.00
Apoyo logístico	31,608,151.23
Levantamiento de datos	6,118,041.84
Total	38,356,856.07

Fuente: INEE (2015).

La tabla 1 no refleja gastos de sueldos, impresión de cuadernillos y cuestionarios, codificación de respuestas a reactivos abiertos ni otro tipo de elementos necesarios para asegurar la adecuada aplicación de la prueba PISA. Considerando el total de \$38,356,856.07,⁴ es una inversión alta para 275 escuelas. Como referente de la dimensión del gasto, si se considera la información del Cuestionario Sobre Financiamiento Educativo Estatal (CFEE) de la Secretaría de Educación Pública⁵ (SEP, 2017), el monto destinado a la aplicación de PISA detallado en la tabla 1 es mayor que el ejercido en educación por 96.8% de los 32 estados del país en 2015 (tabla 2).

Cabe aclarar que el CFEE no considera los montos que la SEP destina a la educación, que son aquellos que mayoritariamente se utilizan para operar los gastos relacionados con la educación en el país, sino que únicamente toma en cuenta los que cada estado ejerce a partir del presupuesto federal, estatal y municipal. A pesar de esta importante delimitación, un tercer problema refiere al gasto que el INEE destinó a la renta y movilización de equipo de cómputo para la aplicación de una prueba, lo cual no solo no representó una inversión que fortaleciera la infraestructura educativa de las escuelas participantes, sino que el gasto realizado no permitió la obtención de datos a profundidad; al contrario, presentó información reducida respecto de otros ciclos. Además, el hecho de que el gasto que el país erogó para aplicar una prueba sea mayor que el destinado por algún estado a la educación no es un elemento menor.

TABLA 2

Gasto educativo realizado por cada una de las 32 entidades de México en 2015

Entidad	Federal	Estatad	Municipal	Total
Aguascalientes	5,609,188.70	1,781,654.50	0.00	7,390,843.20
Baja California Sur	4,355,074.90	925,211.10	0.00	5,280,286.00
Baja California	10,700,968.20	10,819,916.60	0.00	21,520,884.80
Campeche	6,373,119.00	1,108,714.50	7,588.40	7,489,421.90
Coahuila	10,023,931.50	5,443,323.60	0.00	15,467,255.10
Colima	3,111,554.80	762,336.80	6,527.80	3,880,419.40
Chiapas	17,403,117.00	8,358,206.50	0.00	25,761,323.50
Chihuahua	11,441,983.10	12,459,024.00	0.00	23,901,007.10
Ciudad de México	0.00	0.00	0.00	0.00
Durango	7,418,865.00	2,951,623.70	0.00	10,370,488.70
Guanajuato	17,853,560.00	9,732,517.70	0.00	27,586,077.70
Guerrero	16,266,023.80	3,991,116.90	0.00	20,257,140.70
Hidalgo	12,922,997.10	210,000.00	0.00	13,132,997.10
Jalisco	18,802,786.10	15,159,803.00	0.00	33,962,589.10
Estado de México	35,805,537.60	35,268,094.50	0.00	71,073,632.10
Michoacán	17,326,880.40	7,942,536.80	0.00	25,269,417.20
Morelos	6,692,153.60	911,076.60	0.00	7,603,230.20
Nayarit	4,951,043.50	1,051,525.20	0.00	6,002,568.70
Nuevo León	13,528,336.00	10,989,005.70	0.00	24,517,341.70
Oaxaca	21,403,301.97	0.00	0.00	21,403,301.97
Puebla	16,632,002.14	12,160,790.87	0.00	28,792,793.01
Querétaro	6,428,847.60	2,451,108.80	138,766.60	9,018,723.00
Quintana Roo	5,509,678.60	1,207,823.90	0.00	6,717,502.50
San Luis Potosí	11,510,554.50	3,860,365.80	0.00	15,370,920.30
Sinaloa	10,141,252.50	6,957,199.00	0.00	17,098,451.50
Sonora	8,865,548.20	8,133,445.00	0.00	16,998,993.20
Tabasco	7,436,010.00	5,257,192.00	167,672.40	12,860,874.40
Tamaulipas	12,613,777.30	3,406,139.30	165,738.00	16,185,654.60
Tlaxcala	5,301,528.90	2,032,596.20	0.00	7,334,125.10
Veracruz	29,309,921.50	12,691,593.70	0.00	42,001,515.20
Yucatán	5,670,981.30	5,974,941.60	0.00	11,645,922.90
Zacatecas	6,364,068.70	1,112,412.10	0.00	7,476,480.80

Fuente: Cuestionario Sobre Financiamiento Educativo Estatal (CFEE) de la Secretaría de Educación Pública⁶ (SEP, 2017).

De esta manera, en la aplicación de PISA 2015 por computadora en México, al no tener la certeza de contar con infraestructura necesaria en las escuelas participantes, se optó por rentar equipo de cómputo y transportarlo entre los 275 planteles participantes en todo el país. Esta decisión permitió asegurar las condiciones de sistematicidad que requiere este tipo de evaluaciones, sin embargo, al finalizar la aplicación, el monto utilizado con fines de la aplicación no representó una inversión para el país, en el sentido de que no se aprovechó el equipo para mejorar las condiciones de infraestructura de las escuelas, sino que solo sirvió como medio para recolectar información.

Considerando todo lo anterior, la aplicación por computadora de PISA 2015 en México resultó en una decisión con poca posibilidad de implementación en términos de infraestructura educativa, así como en un gasto importante que no contribuyó a la mejora de la infraestructura o de condiciones escolares, sino que, como se comentó previamente, limitó el acceso a la información que se ofrecía en ciclos previos.

El impacto del uso de la computadora en el desempeño de los estudiantes

Adicionalmente a las implicaciones en información, infraestructura y en temas económicos, vale considerar un aspecto relacionado con la metodología del desarrollo de este tipo de pruebas. Uno de los temas que más interesa a los psicómetras respecto de la CBA es, además de la posibilidad de desarrollar evaluaciones adaptativas,⁷ la manera en que las versiones computarizadas de las pruebas educativas pueden tener cierta influencia en el desempeño que muestran quienes la sustentan. Al estudio de este interés se le suele conocer como *mode-effects study*. Este tipo de trabajos se enfocan en el análisis de la manera en que la administración o aplicación de las pruebas influye en los resultados; es decir, se pretende indagar si la aplicación por computadora o en lápiz-papel tiene impacto en los resultados de los sustentantes.

Hasta el momento, no hay una línea concluyente respecto de si el uso de una computadora tiene efectos o no sobre el desempeño de los estudiantes (Marcenaro-Gutiérrez y López-Agudo, 2016). Sin embargo, existen investigaciones que indican que no se obtendrán los mismos resultados si se evalúa por medio de computadora o en lápiz-papel (Clariana y

Wallace, 2002). La pregunta sobre la posible diferencia de resultados entre los sustentantes de acuerdo con la forma de realizar la prueba no se debe por la computadora en sí misma, sino a la posibilidad de cuestionar si el cambio de modalidad asegura la medición de las mismas competencias, o bien, si se pone en desventaja a cierto grupo de estudiantes (Björnsson, 2008).

Existen diversas posturas al respecto. Por una parte, hay quienes apelan al uso de computadoras para el desarrollo de evaluaciones a gran escala, en aras de aprovechar las virtudes que ofrece esta modalidad. Entre las ventajas se encuentran la posibilidad de asegurar la estandarización de procesos de aplicación, la eficiencia en el procesamiento de la información, así como la viabilidad de codificar respuestas de forma automática y de generar pruebas adaptativas (Csapó *et al.*, 2012; Greiff, Kretzschmar, Müller, Spinath *et al.*, 2014). Además, esta forma de aplicación facilita la codificación de respuestas de los sustentantes y la reducción de error en dicho proceso, sin olvidar que promueve la interactividad, posibilitando la inclusión de ítems o reactivos dinámicos (Kikis-Papadikis y Kollias, 2009). Desde la perspectiva de los estudiantes permite la obtención de resultados de forma más rápida, contar con una evaluación más disfrutable y disminuir el factor de estrés (Björnsson, 2008). Incluso, en términos de precisión de medida, hay estudios que indican que la aplicación por computadora es más estable y consistente en términos de validez interna y externa (Piaw, 2012).

Frente a las posibilidades que permite la CBA, la postura del INEE (2016b) se enfoca en tres aspectos:

- La evaluación de PISA a través de computadora favorece la evaluación de competencias más amplias que *aprender a aprender*, como la colaboración y resolución de problemas. Esta postura es coincidente con los estudiosos del fenómeno de la resolución de problemas en colaboración, quienes afirman que su evaluación requiere del uso de computadoras (Greiff *et al.*, 2014).
- La aplicación lápiz-papel se ve rebasada porque no permite la medición de “tareas auténticas” (Greiff *et al.*, 2014:15). Desde luego, esta aseveración resulta riesgosa, ya que pone en duda la certeza de los resultados de ciclos previos y asume que solo la CBA permite una

evaluación auténtica, lo cual sería una contradicción para los informes que el propio INEE ha publicado.

- La evaluación por computadora permite adaptarse de manera interactiva con los jóvenes. Lo cual también es coincidente con las posturas mencionadas previamente.

Bajo esas tres premisas, la OCDE desarrolló un *mode-effects study* en el piloto de PISA 2015 (OCDE, 2016). Para identificar si la presentación de los ítems podía tener efectos en los resultados, este organismo aprovechó la aplicación piloto. En esa prueba los grupos de alumnos de las escuelas seleccionadas de todos los países participantes se segmentaron en tres subgrupos (OCDE, 2016): 1) los estudiantes que resolvieron la prueba en lápiz-papel, 2) los que la presentaron por computadora, y 3) los que la hicieron por computadora con ítems nuevos de Ciencias y de Solución de problemas en colaboración.⁸

El *mode-effects study* de la OCDE se enfocó en los estudiantes de los dos primeros grupos, con la finalidad de valorar la posibilidad de seguir usando los ítems utilizados en ciclos previos bajo la modalidad CBA. Los resultados del informe técnico no son tan concluyentes ni optimistas, sino que sugieren tener una interpretación equilibrada de los datos arrojados. El análisis se hizo a través de estimaciones basadas en la teoría de respuesta al ítem, con la finalidad de valorar si ambas versiones de la prueba se ajustan al mismo modelo de medida. Los resultados muestran que hay variaciones positivas y negativas; es decir, hay ítems o reactivos que favorecen a la aplicación en lápiz-papel y viceversa, específicamente en términos de dificultad. Por tanto, para asegurar la equivalencia entre versiones, PISA asume que el modelo de medida de ambos formatos de la prueba es el mismo, considerando únicamente el parámetro Funcionamiento Diferencial del Ítem (DIF, por sus siglas en inglés),⁹ sin la inclusión de ninguna variable relacionada con habilidad (OCDE, 2016).

Adicional a esta inferencia –que es bastante discutible debido a que el elemento definitorio para asumir equivalencia entre pruebas son los grupos de estudiantes y no sus habilidades en sí–, de acuerdo con el informe de la OCDE (2016), el tamaño de las muestras de la aplicación piloto en cada país es demasiado bajo para poder verificar si el promedio de desempeño en cada área se ve mermado por el modo de aplicación de la prueba, por

lo que, como recomendación final, se solicita al lector la interpretación ecuaníme de los datos.

En el informe del INEE (2016b) no se reporta nada en relación con los resultados del *mode-effects study*, salvo el hecho de que se llevó a cabo. Sin embargo, es necesario cuestionar si el resultado observado realmente se debe al nivel de *literacy* de los estudiantes. En términos de equidad, este es un asunto importante, ya que los estudiantes con un menor acceso a tecnologías pueden presentar problemas para contestar la prueba por computadora (Marcenaro-Gutiérrez y López-Agudo, 2016). Considerando únicamente los resultados de ciclos previos, PISA ha recibido críticas respecto de que los datos no reflejan del todo los *aprendizajes para la vida*, en función que pueden depender del factor cultural y geográfico, o que el tamaño de las muestras utilizadas entre países es cuestionable en cuanto a las inferencias que se realizan con ellos (Fernández-Cano, 2016); o bien, que PISA no propone orientaciones didácticas que guíen los procesos de aprendizaje evaluados (Carabaña, 2015). Ahora, con el uso de computadoras para desarrollar la evaluación, los resultados pueden ser más cuestionables.

De acuerdo con Clariana y Wallace (2002), la probabilidad de que se obtengan los mismos resultados por medio de las dos formas de administración es de .50. Este valor no es sutil, dado el impacto político que PISA tiene para los diversos países que participan en él y los resultados que el INEE comunica. En particular, al aseverar que México tiene poca mejoría entre ciclos en Ciencias, y que, al parecer, el aumento de 1.7 cada tres años en la media nacional se debe prácticamente a cambios en la metodología de PISA; además, no se descarta que puedan influir también las tasas de matriculación a los 15 años, o bien, la composición demográfica de la población.

Frente a ello cabe la pregunta, ¿los resultados que obtienen los estudiantes mexicanos de 15 años de edad en PISA 2015 son realmente válidos? En los llamados *factores asociados* a los resultados no se habla del uso de computadoras por parte de los evaluados, a pesar de que a todos los estudiantes se les aplicó el cuestionario sobre el uso de las TIC. La información resultante de este instrumento pudo haber ayudado a saber las condiciones de uso de las computadoras y situar así a los resultados observados. Sin embargo, la información disponible no permite hacer aseveraciones al respecto.

El problema de asegurar que los resultados de los estudiantes mexicanos no se vieron influidos por el uso de las computadoras no se resuelve dada la evidencia ofrecida por la OCDE (2016) o por el INEE (2016b), pero deja abierta la posibilidad de discusión respecto de si las habilidades de los estudiantes mexicanos reportadas no se ven influidas por la falta de infraestructura en su escuela, que se traduce en un posible desconocimiento del uso académico de las computadoras. No debe perderse de vista que una discrepancia entre la enseñanza y los métodos de evaluación compromete el significado de los resultados de una evaluación (Csapó *et al.* 2012); sobre todo debido a que las posibles brechas de los desempeños observados entre quienes presentan pruebas en papel respecto de los de computadora pueden deberse al acceso que se tiene para el conocimiento de las tecnologías, independientemente de la evaluación en sí misma (Marcenaro-Gutiérrez y López-Agudo, 2016).

Conclusiones

Hasta ahora hemos enfrentado cuatro problemas importantes en lo que respecta a la evaluación de PISA por computadora en el ciclo 2015 en México. Todos ellos abonan al cuestionamiento de la pertinencia de esta decisión dada la falta de información que implicó, la carencia de infraestructura, el valor económico que representó este esquema de aplicación, así como la posible validez de los resultados de este ciclo frente a falta de evidencias sobre la pertinencia de esta modalidad en la población evaluada. No se juzga el valor del proyecto como insumo político para el país dada la pertenencia de México en la OCDE así como su papel en rendición de cuentas (Jiménez, 2016), que son motivo de otro tipo de estudios; pero la decisión de aplicar la prueba por computadora, dados los argumentos expuestos, se torna cuestionable.

La plataforma tecnológica que permite una evaluación a gran escala con estas características es difícil de alcanzar, aun para otros continentes, como el asiático. Como lo afirman Papanastasiou, Michalinos y Charalambos (2003), incluso en un contexto con mayor acceso a la tecnología, como el caso de Estados Unidos, este por sí mismo no es suficiente; por supuesto, esto se complica cuando el país no cuenta con la infraestructura necesaria no solo para evaluar, sino para incorporar la tecnología en las aulas (Ramírez, 2006). No se puede dejar de lado la posibilidad de que, en

la CBA, el desempeño observado puede estar influido por las habilidades de manipulación de las tecnologías de informática (Marcenaro-Gutiérrez y López-Agudo, 2016).

Asimismo, la literatura presenta posibles problemas o desventajas del uso del CBA que guardan cierta relación con los problemas expuestos en este artículo, como la inversión económica que requiere su implementación, así como los posibles bajos desempeños en los puntajes obtenidos y que puedan ser resultantes del modo de administración (Kikis-Papadikis y Kollias, 2009). Respecto de este último punto, la familiaridad con el uso de las computadoras es el factor fundamental en el análisis de los factores que influyen en el efecto del modo de administración (Clariana y Wallace, 2002). En ese sentido, el uso de la tecnología en el proceso de enseñanza no puede considerarse solo como una variable más (Papanastasiou, Michalinos y Charalambos, 2003).

Desde el punto de vista político, es bien sabido que PISA se ha utilizado en diversas partes del mundo no como un medio de mejora del aprendizaje de los estudiantes, sino como una forma de rendir cuentas, justificar el gasto económico en educación, o bien, de administrar el gasto futuro en este rubro (Fernández-Cano, 2016). Sin embargo, la decisión de modificación de la forma de aplicación del instrumento no puede estar supeditado únicamente a presiones implícitas por parte de la coordinación internacional del proyecto. Sobre todo cuando la literatura parece orientar que la familiaridad con el uso de la computadora se relaciona con el efecto del modo de administración (Clariana y Wallace, 2002); o bien, cuando los datos del mismo INEE obtenidos a través del estudio ECEA (INEE, 2016a) o del *Panorama educativo de México* (INEE, 2016c) muestran la imposibilidad operativa de su desarrollo en términos de computadoras disponibles.

La aplicación de PISA 2015 por computadora conllevó a un alto gasto para el Estado mexicano que solo permitió ofrecer información con menor amplitud de la presentada en ciclos previos, lo cual no puede ser valorado como un acierto. Además, representó ser un ejercicio que no se enmarca en la realidad mexicana, dada la falta de infraestructura que permea en las escuelas del país, particularmente en lo que se refiere a las computadoras disponibles para los estudiantes. En ese sentido, si el Estado no ha logrado dotar de infraestructura informática a todas sus escuelas, es válido dudar

de que todos los estudiantes evaluados hayan tenido experiencia previa en el uso de estas herramientas con fines pedagógicos, lo cual es posible que haya impactado en sus resultados. Aunado a ello, las evidencias del *mode-effects study* mostradas en OCDE (2016) no permiten asegurar de forma concluyente el alcance de los resultados de esta aplicación.

En medio de estas problemáticas cabe la pregunta, ¿cuál fue el beneficio para el sistema educativo mexicano al haber aplicado PISA bajo esta modalidad? Más allá de las aparentes ventajas de medida establecidas por la literatura y por el mismo INEE en el contexto mexicano, ¿la evaluación basada en computadora representa una necesidad?, ¿es viable? Curiosamente, a pesar de que pareciera haber poca evidencia que sustente la viabilidad y posibles consecuencias positivas de esta decisión, para el INEE, la evaluación por computadora parece ser más relevante que la información que puede brindar, así como sobre las condiciones propias del país y sus estudiantes. Al respecto, retomando el caso asiático (Csapó, *et al*, 2012), parece prudente no usar este tipo de evaluación en función de la poca posibilidad del país de asegurar condiciones sistematizadas de aplicación. Paradójicamente, el propio INEE (2016b), en el informe de PISA 2015, recomienda que el sistema educativo mexicano debe mejorar la infraestructura de los centros escolares, así como incrementar el equipamiento informático y de conectividad a internet.

Notas

¹ Puede consultarse la página: <https://www.taotesting.com/>

² Si bien, el Ceneval no tiene información técnica publicada sobre las características de sus evaluaciones informatizadas, en sus diferentes guías de estudio se incluye una descripción general de las mismas. Esta información puede ser consultada a detalle en: <http://www.ceneval.edu.mx/>

³ El estudio ECEA sobre educación media superior y el correspondiente a secundaria no habían sido publicados al momento del desarrollo de este trabajo.

⁴ Equivalente a 2,158,517.505 dólares con un tipo de cambio de \$17.77 pesos, de acuerdo con información del Banco de México al de 30 de agosto de 2017 (<http://www.banxico.org.mx/portal-mercado-cambiario/>).

⁵ Información pública disponible en: <http://www.planeacion.sep.gob.mx/recuesfinaeduct.aspx>

⁶ La Ciudad de México ejerce su gasto educativo a partir del presupuesto de la SEP Federal, por ello no contempla gastos de las categorías analizadas por el CFEE.

⁷ Las llamadas *pruebas adaptativas* son instrumentos que, mediante el uso de la Teoría de Respuesta al Ítem (TRI), estiman la habilidad del sustentante en función de las respuestas que la persona da a un reactivo computarizado. La idea es que la computadora estime la habilidad del estudiante para presentarle reactivos cuyo nivel de dificultad sea acorde con su habilidad (Sierra-Matamoros, Valdelamar-Jiménez, Hernández-Tamayo y Sarmiento-García, 2007).

⁸ En caso de que los países no aplicaran la prueba por computadora, solo tuvieron el grupo de estudiantes que presentó la prueba en papel.

⁹ El funcionamiento diferencial del ítem refiere al análisis de las posibles diferencias en las funciones de los reactivos entre grupos sociales (Osterlind y Everson, 2009).

Referencias

- Baird, Jo-Anne; Isaacs, Talia; Johnson, Sandra; Stobart, Gordon; Yu, Guoxing; Sprague, Terra y Daugherty, Richard (2011). *Policy effects of PISA*, Londres: Oxford University Center for Educational Assessment.
- Barberá, Elena (2016). “Aportaciones de la tecnología a la e-Evaluación”, *Revista de Educación a Distancia*, núm. 50, pp. 1-13. Disponible en: <http://www.um.es/ead/red/50/barbera.pdf>
- Björnsson, Juluis (2008). “Changing Iceland national testing from traditional paper and pencil based tests to computer based assessment: some background, challenges and problems to overcome”, en Friedrich Scheuermann y Angela Guimaraes (eds.), *Towards a research agenda on Computer-Based Assessment*, Ispra: European Commission-Joint Research Centre-Institute for the Protection and Security of the Citizen, pp. 10-14.
- Carabaña, Julio (2015). *La inutilidad de PISA para las escuelas*, Madrid: Catarata.
- Clariana, Roy y Wallace, Patricia (2002). “Paper-based versus computer-based assessment: key factors associated with the test mode effect”, *British Journal of Educational Technology*, vol. 33, núm. 55, pp. 593-602. Disponible en <https://pdfs.semanticscholar.org/3e8a/ff00cb390786995d0abb473d8233866b3f9e.pdf>
- Csapó, Benó; Ainley, John; Bennet, Randy; Latour, Thibaud y Law, Nancy (2012). “Technological issues of Computer-Based Assessment”, en Patrick Griffi y Esther Care (ed.), *Assessment and teaching of 21st century skills*, Nueva York: Springer, pp. 143-230.
- Fernández-Cano, Antonio (2016). “Una crítica metodológica a las evaluaciones PISA”, *Revista Electrónica de Investigación Educativa*, vol. 22, núm. 1, pp. 1-17.
- Fernández-González, Noelia (2015). “PISA como instrumento de legitimación de la Reforma del LOMCE”, *Bordón*, vol. 67, núm. 1, pp. 165-178.
- Greiff, Samuel; Kretzschmar, André; Müller, Jonas; Spinath, Birgit y Martin, Romain (2014). “The Computer-Based Assessment of complex problem solving and how it is influenced by students’ Information and Communication Technology literacy”, *Journal of Educational Psychology*, vol. 106, núm. 3, pp. 1-15.
- INEE (2015). *Relación de pedidos, contratos y convenios del primer semestre de 2015*, Ciudad de México: Instituto Nacional para la Evaluación de la Educación.
- INEE (2016a). *Infraestructura, mobiliario y materiales de apoyo educativo en las escuelas primarias. ECEA 2014*, Ciudad de México: Instituto Nacional para la Evaluación de la Educación.
- INEE (2016b). *México en PISA 2015*, Ciudad de México: Instituto Nacional para la Evaluación de la Educación.
- INEE (2016c). *Panorama educativo de México 2015*, Ciudad de México: Instituto Nacional para la Evaluación de la Educación.

- Jiménez, Alfonso (2016). "El papel de la evaluación a gran escala como política de rendición de cuentas en el sistema educativo mexicano", *Revista Iberoamericana de Evaluación Educativa*, vol. 9, núm. 1, pp. 109-126.
- Jornet, Jesús; González-Such, José y García-Bellido, María Rosario (2012). "La investigación evaluativa y las Tecnologías de la Información y Comunicación (TIC)", *Revista Española de Pedagogía*, vol. 70, núm. 251, pp. 93-110.
- Kikis-Papadikis, Katerina y Kollias, Andreas (2009). "Reflections on paper-and-pencil test to eAssessments: narrow and broadband paths to 21st century challenges", en Friedrich Scheuermann y Julius Björnsson (eds.), *The transition to Computer-Based Assessment. New approaches to skills assessment and implications for large-scale testing*, Ispra: European Commission-Joint Research Centre-Institute for the Protection and Security of the Citizen, pp. 92-96.
- López, Ruber; Sanmartín, Paul y Méndez, Fernando (2013). "Revisión de las evaluaciones adaptativas computarizadas (CAT)", *Educación y Humanismo*, vol. 16, núm. 26, pp. 27-40.
- Marcenaro-Gutiérrez, Óscar y López-Agudo, Luis Alejandro (2016). "Mind the gap: analysing the factors behind the gap in student's performance between pencil and computer based assessment methods", *Revista de Economía Aplicada*, vol. 71, núm. 24, pp. 93-120.
- OCDE (2016). *PISA 2015. Results (volume I)*, París: Organización para la Cooperación y el Desarrollo Económicos.
- Olea, Julio; Abad, Francisco y Barrada, Juan (2010). "Tests informatizados y otros nuevos tipos de tests", *Papeles del Psicólogo*, vol. 31, núm. 1, pp. 97-107.
- Osterlind, Steven y Everson, Howard (2009). *Differential item functioning*, Thousand Oaks: Sage Publications.
- Papanastasiou, Elena; Michalinos, Zembylas y Charalambos, Vrasidas (2003). "Can computer use hurt science achievement? The USA results from PISA", *Journal of Science Education and Technology*, vol. 12, núm. 3, pp. 325-332.
- Piaw, Chua Yan (2012). "Replacing paper-based testing with computer-based testing in assessment: Are we doing wrong?", *Procedia. Social and Behavioral Sciences*, núm. 64, pp. 655-664.
- Ramírez, José Luis (2006). "Las tecnologías de la información y comunicación en la educación en cuatro países latinoamericanos", *Revista Mexicana de Investigación Educativa*, vol. 11, núm. 28, pp. 61-90.
- Rodríguez, María José (2005). "Aplicación de las TIC a la evaluación de alumnos universitarios", *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, vol. 2, núm. 6.
- Rojas, Antonio (2001). "Pasado, presente y futuro de los Tests Adaptativos Informatizados: entrevista con Isaac I. Bejar", *Psicothema*, vol. 14, núm. 4, pp. 685-690.
- SEP (2017). *Resultado del Cuestionario de Financiamiento Educativo Estatal*, Ciudad de México: Secretaría de Educación Pública. Disponible en: <http://www.planeacion.sep.gob.mx/recuesfinaeduest.aspx> (consultado el 30 de agosto de 2017).

- Sierra-Matamoros, Fabio Alexandre; Valdelamar-Jiménez, Juliet Rocío; Hernández-Tamayo, Félix Adrián y Sarmiento-García, Luz Miriam (2007). “Test adaptativos informatizados”, *Avances en Medición*, vol. 5, pp. 157-162.
- Taut, Sandy y Palacios, Diego (2016). “Interpretaciones no intencionadas e intencionadas y usos de los resultados de PISA: una perspectiva de validez consecencial”, *Revista Electrónica de Investigación y Evaluación Educativa*, vol. 22, núm. 1.
- Tirado, Felipe; Backhoff, Eduardo y Larrazolo, Norma (2011). “Las complejas relaciones en los procesos de evaluación: un análisis basado en el Examen de Habilidades y Conocimientos Básicos (EXHCOBA)”, *XI Congreso Nacional de Investigación Educativa*, Ciudad de México: Consejo Mexicano de Investigación Educativa.

Artículo recibido: 5 de septiembre de 2017

Dictaminado: 6 de abril de 2018

Segunda versión: 9 de abril de 2018

Aceptado: 11 de abril de 2018