

El Papel de la Evaluación a Gran Escala como Política de Rendición de Cuentas en el Sistema Educativo Mexicano

The Role of the Large-Scale Assessment as Political Accountability in the Mexican Educational System

José Alfonso Jiménez Moreno*

Universidad Nacional Autónoma de México

En la actualidad, la evaluación a gran escala que se realiza mediante instrumentos estandarizados de aplicación nacional e internacional es una de las herramientas más utilizadas para el conocimiento y monitoreo del estatus de los sistemas educativos. Los resultados que este tipo de evaluación arrojan suelen ser esperados por los diversos actores educativos, dado que son un insumo importante para conocer los avances de los estudiantes. Con la finalidad de comprender de manera particular el papel de este tipo de evaluaciones dentro de las políticas educativas en México, el presente artículo hace uso de datos teóricos e históricos para analizar la relación entre las políticas económicas de las últimas décadas a nivel internacional y el uso de la evaluación a gran escala como mecanismo de rendición de cuentas dentro del sistema educativo mexicano. Derivado del análisis presentado, se concluye que el uso de este tipo de evaluación por parte del sistema educativo mexicano se da como consecuencia de una constante búsqueda de competitividad económica internacional.

Palabras clave: Evaluación basada en el rendimiento, Política educativa, Rendición de cuentas, Evaluación internacional, Pruebas nacionales de competencia.

At present, the large-scale assessment that is performed by standardized instruments of national and international administration is one of the most relevant tools for understanding and monitoring the status of the educational systems. The results that this kind of assessment release are usually expected by a diversity of stakeholders, as those results are an important input to understand the progress of students. In order to understand the role of this kind of assessment in educational policy in México, this article makes use of theoretical and historical data to analyze the relationship between the economic policies of recent decades at the international level and the use of large-scale assessment as a mechanism of accountability within the Mexican education system. Derived from the presented analysis, it is concluded that the use of this type of evaluation by the Mexican educational system is given as a result of a constant search for international economic competitiveness.

Keywords: Performance based assessment, Educational policy, Accountability, International assessment, National competency tests.

1. Introducción

La evaluación a gran escala es una de las formas de evaluación educativa de mayor uso en los sistemas educativos y, técnicamente hablando, es caracterizada particularmente por su alta afinidad con la psicometría como parte de su metodología de desarrollo. Sin embargo, no es su único elemento definitorio, ya que en la práctica cotidiana dentro de los sistemas educativos sus resultados suelen ser utilizados como herramienta de decisión en materia de política educativa y, en algunos casos, como proveedora de información para toma de decisiones que involucran el ingreso o egreso de estudiantes a un nivel educativo en particular.

Además de la relevancia académica y social que este tipo de evaluación puede tener, su importancia como herramienta definitoria de políticas educativas es innegable. Frente a ello, el presente artículo tiene por objeto mostrar el papel que la evaluación a gran escala ha tenido como herramienta de consolidación de la política educativa mexicana basada en la competitividad y rendición de cuentas, para ello se realiza un análisis del papel de este tipo de evaluación en la consolidación de la política educativa influida por los procesos de globalización económica, mediante el uso de datos teóricos e históricos relacionados con los antecedentes más representativos de este tipo de evaluación, así como un recuento y una crítica de las políticas educativas que han sido determinantes en el uso de la evaluación a gran escala como herramienta de consolidación de una política neoliberal en México.

2. Antecedentes y conceptualización

La evaluación a gran escala basa su quehacer en la teoría de la medición psicológica (Martínez, 2001). Es a partir del siglo XX que su metodología comienza a consolidarse mediante el uso de elementos estadísticos como confiabilidad y error de medida (Martínez Arias, 1995, citado en Martínez, 2001). El uso de pruebas psicométricas centradas en la educación inició bajo la intención de medir el aprendizaje como rasgo psicológico, ya que se consideraba que este era el indicador que mejor podía reflejar los logros educativos, dado que mostraba de manera directa la influencia que los programas podían tener sobre los estudiantes. La evaluación a gran escala que comenzó a desarrollarse desde el siglo XX, particularmente en Estados Unidos, tenía como interés identificar si los logros de los estudiantes eran acordes con los mínimos aceptables (Martínez y Santos, 2009). De manera concreta, la evaluación a gran escala consiste en la aplicación de pruebas a una gran cantidad de alumnos, ello con la finalidad de identificar el nivel de aprendizaje logrado por un país o región (Martínez, 2008b).

La evaluación a gran escala se fundamenta en la medición del logro del aprendizaje bajo el respaldo de técnicas estadísticas. Esto ha favorecido la obtención de evidencia empírica objetiva, aspecto de gran aprecio por parte de los actores involucrados con la toma de decisiones en educación debido a la confianza que el uso de métodos objetivos trae para la definición de políticas y acciones educativas. De esta forma, el uso de técnicas psicométricas y estadísticas ha facilitado el desarrollo de la evaluación, ya que conjuntan un proceso metodológico que privilegia la objetividad con una forma eficiente de procesamiento de datos. Actualmente la metodología de la evaluación a gran escala permite incluso estimar las habilidades de una persona en función de un manejo probabilístico de los datos (Martínez, 2001).

El objetivo de la evaluación a gran escala, de acuerdo con Ravela (2006) es que permite “conocer en qué medida están siendo logrados los conocimientos y competencias que se espera que los estudiantes hayan adquirido cuando finalizan un nivel o ciclo educativo” (p. 18). Según este autor, existen diversas finalidades de una evaluación a gran escala. La primera de ellas es la certificación, que se centra en la identificación de los conocimientos y competencias estipulados para un nivel educativo en particular. Un segundo objetivo es la selección, que se utiliza con fines de ingreso a instituciones educativas. Otra de sus metas es la toma de decisiones “blandas”, que refiere al sentido formativo de la evaluación. El autor las llama “blandas” debido a que su uso no conlleva una consecuencia formal. En contraparte se encuentra el objetivo de toma de decisiones “duras”, que se relaciona con las evaluaciones que traen como consecuencia la continuidad de un proyecto o programa. Finalmente, Ravela considera el establecimiento de incentivos y la rendición de cuentas y responsabilización, que permiten valorar el avance de las acciones educativas de mejora.

La clasificación realizada por Ravela refleja a grandes rasgos las diferentes formas de comprender los objetivos de la evaluación a gran escala. Sin embargo, no representa una clasificación rígida con categorías excluyentes entre sí, dado que –de acuerdo con esta categorización– una evaluación cuyos resultados sean condicionante para que los estudiantes puedan egresar del nivel licenciatura o tengan posibilidades de obtener un título profesional puede ser considerada como de certificación, de toma de decisiones *duras* o, incluso, como elemento de rendición de cuentas para efectos de los indicadores que las instituciones utilizan dentro de sus procesos de acreditación.

Por su parte, Martínez y Santos (2009) clasifican a las evaluaciones a gran escala en dos tipos: a) Selección: Se aplican a aspirantes de ingreso a una institución educativa y centran su atención en la validez predictiva respecto al posible desempeño de los aspirantes; y b) Aprovechamiento: Centradas en los alumnos que terminan un ciclo o grado escolar y en el aseguramiento de la validez y los contenidos con una fuerte carga curricular.

Considerando los elementos que se han descrito, los aspectos definitorios de la evaluación a gran escala están supeditados por cuestiones técnicas definidas por la psicometría y los usos que se dan a los resultados. Ahora bien, cabe aclarar una precisión terminológica: en el idioma inglés a la evaluación a gran escala se la conoce como *large scale assessment*; sin embargo, en dicho idioma existe una diferencia conceptual entre *assessment* y *evaluation*, lo cual permite diferenciar al *assessment* como forma de medición u observación constante de un fenómeno en particular de la *evaluation* como elemento de retroalimentación a partir de las mediciones observadas. En castellano, se hace uso del término *evaluación* sin diferencia respecto al *assessment*, ya que no existe una traducción fiel para el término. Evidentemente, ello también representa un problema, ya que el término *evaluación* se utiliza de manera indistinta en las prácticas de los evaluadores profesionales al referirse a procesos propios de medición. Sin embargo, más allá del término utilizado para el fenómeno a gran escala en castellano (*large scale assessment* vs *evaluación a gran escala*), la evaluación a gran escala asume su base psicométrica y su ejercicio en pro de las decisiones que se quieran tomar con los resultados debido a la objetividad que con la elaboración de sus instrumentos pretende alcanzar.

3. La evaluación a gran escala y su relación con la rendición de cuentas

El uso de la evaluación a gran escala se ha implementado y mantenido por necesidades educativas y sociales de estandarización de contenidos y su verificación por medio del aprendizaje logrado por los alumnos, alentadas básicamente por decisiones de política gubernamental centrada en la rendición de cuentas, como a continuación se describe. Adicionalmente, la masificación de la educación ha traído consigo la necesidad de contar con elementos objetivos para la regulación del ingreso y egreso a las instituciones educativas.

La rendición de cuentas implica la obligación de reportar a otros, explicar o justificar sobre cómo se usan los recursos y los resultados que se obtienen con ellos. En el caso de la educación, la rendición de cuentas se relaciona con los resultados que de ella se generan y cómo se aprovechan con fines de una posible potencialización del mercado (Trow, 1996). La rendición de cuentas permite dar legitimidad a las instituciones, dándoles fuerza respecto a la forma en que manejan sus recursos y los resultados que se obtienen, incluso existe una tendencia en Europa respecto a enfocarse en la calidad (entendida como resultados) para la designación del presupuesto educativo, aunque de acuerdo con Trow (1996), la intención de relacionar la rendición de cuentas con la masificación de la educación superior se da en todo el mundo.

De acuerdo con Popham (2001, citado en Depascale, 2003), la evaluación a gran escala tiene la misión primaria de rendir cuentas, ello a través de la observación de los logros de los estudiantes como elemento de análisis del avance educativo. Esta facultad de la evaluación a gran escala de poner especial atención en los resultados logrados por los estudiantes favorece su implementación como un elemento de gobernanza, basado en el llamado *policy by number* (Saura y Luengo, 2015), en el cual el uso de mediciones permite establecer un monitoreo sobre las escuelas y el sistema educativo en general y así establecer acciones orientadas a cumplir las políticas educativas de importancia para el Estado.

De manera particular, la gobernanza debe entenderse como cualquier forma de gobierno que involucre a agentes públicos y privados, incluyendo las formas tradicionales de gobierno, así como aquella derivada de la imposición jerárquica de las medidas de información (Héritier, 2011; Khem, 2011). La gobernanza implica entonces la consideración de la participación de diversos agentes en la toma de decisiones sobre procesos importantes que anteriormente corrían únicamente a cargo del Estado (Khem, 2011). El Estado, entendido como el conjunto de instituciones dentro de un territorio y que ejerce poder sobre su población (Mann, 2007; O'Donnell, 2008) tradicionalmente se encarga de centralizar sus decisiones respecto a la política educativa de impacto a su población. En particular en América Latina, antes de la década de los ochenta del siglo pasado, los Estados se caracterizaban por ser proteccionistas y ofrecer una educación democrática (Vargas, 2002); sin embargo, como parte de las políticas económicas de dicha década, los gobiernos de dicha región se orientaron a disminuir el aparato estatal debido a los riesgos inflacionarios que las políticas proteccionistas podrían representar (Centro Latinoamericano de Administración para el Desarrollo, 2007). Como consecuencia, en la década de los noventa se desarrolló un proceso de integración económica internacional, que afectaría al ámbito educativo mediante la búsqueda de homogeneización de los sistemas educativos. Ello conllevaría a que la educación respondiera a los intereses del mercado.

El resultado se manifestó en que algunos elementos de gobernanza influirían en las políticas educativas en la región latinoamericana, aunque ello es un fenómeno que también se vive en otras regiones del mundo, tal es el caso de Europa. En los países de esta región también existe un proceso de estandarización de sistemas educativos, no como una imposición externa, sino como una acción voluntaria (Dale, 1999, citado en Bonal y Tarabini, 2013). Por ejemplo, en el caso particular de España, debido a su condición geográfica, ser parte de la región europea implica buscar compartir objetivos educativos para su población con otros países de la región, en particular siguiendo a aquellos países que han obtenido buenos resultados en evaluaciones internacionales (Bonal y Tarabini, 2013).

Las influencias internacionales como elemento relevante en el establecimiento de las políticas educativas han provocado que a nivel mundial se vivan cambios en la forma respecto en que se conceptualiza a la escuela pública y su financiamiento, así como la privatización escolar. Eso es a lo que Sahlberg (2011, citado en Saura y Luengo, 2015) denomina *Global Education Reform Movement*. El movimiento globalizado que se vive desde las últimas décadas del siglo pasado se basa en el libre mercado, es decir, en el factor competencia como desarrollo de mejores opciones para el consumidor, aspecto que no escapa a la educación, ya que de esa manera se fundamenta el discurso de una búsqueda constante de la calidad (Viñao, 2012). Estas prácticas neoliberalistas (Viñao, 2012) se basan en la idea de que lo privado se relaciona directamente con el bienestar social, mientras que lo público es algo ineficaz; como consecuencia, buscan la liberación y la desregulación de las transacciones económicas, en particular entre fronteras (Jessop, 2002). Ello se manifiesta en la educación mediante la reducción de políticas proteccionistas con el fin de promover una gestión basada en la competitividad que favorezca que los países se mantengan dentro de la dinámica económica global. En el caso de la administración pública en América Latina, bajo este esquema se comenzó a percibir a la educación, a la ciencia y a la tecnología como motores de desarrollo (Centro Latinoamericano de Administración para el Desarrollo, 2007).

La perspectiva neoliberal en educación defiende que los países de América Latina han tenido una deficiente capacidad administrativa y que el hecho de contar con sindicatos fuertes ha limitado la capacidad de rendir cuentas. En ese sentido, Winkler (2006) defiende que los Estados abiertos a la dinámica global deberán establecer objetivos mensurables, mejorar su monitoreo, así como diseñar esquemas de incentivos que los orienten hacia la obtención de financiamiento.

Bajo esta política comenzaron a establecerse mecanismos de monitoreo para que los Estados pudieran conocer los avances hacia los objetivos económicos deseados. De manera particular en lo que al ámbito educativo refiere, los sistemas educativos basados en la rendición de cuentas buscarían así el monitoreo de los logros alcanzados por la escuela, reflejados en los resultados de los estudiantes (Fuhrman y Elmore, 2004).

En este contexto, en América Latina se empieza a cuestionar el papel de la educación pública (en particular en educación superior) y empieza a ponerse especial atención al término de calidad en educación (Vargas, 2002). La calidad de la educación comienza a ser un factor importante, debido a la relevancia de asegurar la homogeneización de la educación superior en aras de alcanzar los estándares de formación que los procesos de globalización exigen. En este proceso de homogeneización o estandarización educativa, la evaluación juega un papel primordial, y es que la evaluación, al hacer uso de

estándares como elemento de juicio sobre el objeto evaluado¹, favorece que los sistemas educativos que se encuentran bajo la dinámica de rendición de cuentas puedan definir si los resultados educativos son suficientes o no respecto a lo esperado (Fuhrman y Elmore, 2004).

Las evaluaciones a gran escala permiten garantizar que el Estado tenga control y orden sobre el currículum escolar (Saura y Luengo, 2015), lo que facilita el cumplimiento de metas basadas en una economía globalizada. De manera particular, las evaluaciones de nivel internacional juegan un papel relevante en esta dinámica. La participación de los diversos países en evaluaciones internacionales se da por razones diversas, desde la búsqueda de fortalecer las políticas educativas nacionales, así como compararse con otros sistemas educativos o tener acceso a financiamiento internacional (Bonal y Tarabini, 2013). El desempeño de los estudiantes puede conocerse de manera fiable mediante instrumentos de evaluación de este tipo, por lo que se infiere que el avance escolar puede ser medido de forma justa mediante este tipo de herramientas (Fuhrman y Elmore, 2004); por supuesto, ello no excluye a las evaluaciones que se desarrollan de forma interna en cada país, ya que la rendición de cuentas puede realizarse de forma externa o interna, por lo que ambas implican una observación constante respecto al avance en alcanzar los resultados esperados (Trow, 1996). La evaluación en todos los niveles educativos, desarrollada tanto de forma externa como interna, facilitará entonces la consolidación de los sistemas educativos basados en la rendición de cuentas.

Actualmente, dentro de la sociedad se asume la relevancia de conocer el avance de resultados educativos, prueba de ello es, en el caso de México, el otorgamiento de la autonomía constitucional al Instituto Nacional para la Evaluación de la Educación (INEE) (Secretaría de Gobernación, 2013), lo que le da facultades para la evaluación del Sistema Educativo Nacional y con ello establecer políticas educativas. De esta manera, en México el Estado ha oficializado una preocupación centrada en que los alumnos alcancen un nivel mínimo de resultados educativos, la cual, como se ha comentado, se consolida mediante pruebas a gran escala. De acuerdo con autores como Díaz-Barriga (1994) y Villaseñor (2003), esta preocupación resulta una consecuencia que deriva de las políticas económicas neoliberales que, al menos en el caso de América Latina, comenzaron en los años ochenta.

Bajo una política neoliberal, la evaluación cumple una función de legalización de la regulación de la educación; es decir, es un reconocimiento administrativo al conocimiento (Díaz-Barriga, 1994), la idea es la de obtener un conocimiento *objetivo* del saber de cada uno de los estudiantes. Según Calero y Choi (2012), la evaluación a gran escala responde fuertemente a una política en la que se espera conocer los resultados de los insumos que los gobiernos destinan a la educación, siendo así un instrumento importante de la difusión y mantenimiento de políticas gubernamentales. Por lo tanto, el conocimiento de los resultados de los estudiantes será un indicador que favorezca la toma de decisiones respecto a la inversión que los Estados realizan directamente en sus sistemas educativos.

Como se comentó, algunos de los ejercicios más representativos de este tipo de evaluación son aquellos que se dan como una iniciativa internacional. En Latinoamérica

¹ Normalmente expresados en niveles de desempeño definidos dentro de las pruebas sistematizadas, o bien en puntajes mínimos esperados.

destacan los trabajos del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de la Oficina Regional de la UNESCO para América Latina y el Caribe (OREALC), centrado en la evaluación de elementos en común de los currículos de los países latinoamericanos participantes en las áreas de Lenguaje, Matemáticas y Ciencias (Solano-Flores y Bonk, 2008). Los estudios que se han realizado en el LLECE constituyen el proyecto de evaluación internacional más importante hasta la fecha en Latinoamérica. Por su parte, la *Association for the Evaluation of Educational Achievement* (IEA) se desarrolló desde mediados de los años cincuenta en el Instituto de Educación de la UNESCO, en Hamburgo, hacia la posibilidad de un trabajo de evaluación a nivel internacional que permitiera obtener resultados estrictamente comparables, ello a través de la utilización de instrumentos equivalentes en contenido y grado de dificultad ante la insatisfacción por el uso de tasas de graduación o eficiencia terminal como único indicador de la calidad educativa. Con ese fin, en 1964 la IEA desarrolló el *First International Mathematics Study* (FIMS) (Saura y Luengo, 2015). En la actualidad, la IEA sigue desarrollando diversas evaluaciones a nivel internacional, como el *International Civic and Citizenship Education Study* (ICCS) (Schulz, Ainley, Fraillon, Kerr y Losito, 2010) o el *Trend in International Mathematics and Science Study* (TIMSS) (Ministerio de Educación de Chile, 2011). En ambos ejemplos hay participación de países pertenecientes a América Latina.

Por su parte, la Organización para la Cooperación y Desarrollo Económicos (OCDE) emprendió el proyecto denominado *Programme for the International Student Assessment* (PISA). Las pruebas se aplican cada tres años a estudiantes de 15 años desde el año 2000, centrando la atención en tres grandes áreas: Lectura, Matemáticas y Ciencias, con mayor énfasis en una de ellas en cada uno de los ciclos de aplicación² (INEE, 2013). La intención de la evaluación de PISA no se enfoca en los currículos de los países participantes, y es que es un programa centrado en conocer el avance de los jóvenes para hacer frente a las demandas de las sociedades de la información y del conocimiento. Independientemente del currículo del sistema educativo en cada país, se espera que los egresados del sistema educativo tengan herramientas que potencialicen su adaptación y éxito en un mundo globalizante (UNESCO, 2005), y PISA ha puesto énfasis en ello.

Los ejercicios internacionales son un indicador importante para la observación de los avances de los sistemas educativos, ya que con un estándar en común entre países participantes (ya sea conformado por temáticas coincidentes en los diversos currículos, como en el caso del LLECE, o bien por conceptos académicos como eje de comparación, como en el caso de PISA) se pueden realizar tendencias de los resultados de los estudiantes entre ciclos de aplicación y con ello analizar posibles tendencias de los avances que muestran los estudiantes, comparando los puntajes propios con contextos diferentes y similares al propio.

El monitoreo del avance educativo que se realiza mediante este tipo de evaluaciones es una práctica que permite conocer elementos genéricos de la educación para la toma de decisiones en materia de política educativa, dado que hace posible la identificación de avances o retrocesos en el aprendizaje de la población de un nivel educativo en

² Desde la aplicación de PISA 2012, en algunos países comenzó a evaluarse adicionalmente la Competencia financiera. En ciclos recientes se considerará también el área de Solución de problemas en colaboración como elemento de evaluación.

particular en una dimensión nacional o internacional, ya que hay ejemplos de pruebas nacionales en cada país que se usan con ese mismo fin, los cuales se abordan más adelante.

Al respecto, la reflexión que realiza Anderson (2005) permite comprender de mejor manera el interés de los ministerios de Educación en la evaluación a gran escala de corte nacional e internacional, al afirmar que en los sistemas basados en resultados reconocen públicamente el desempeño con el fin de otorgar recompensas (en algunos casos económicas) a las escuelas; de tal suerte que cuando los resultados de las evaluaciones otorgan beneficios para las escuelas o programas educativos, las evaluaciones comienzan a ser más solicitadas, incluso se pueden considerar como parte fundamental del sistema educativo (como en el caso del INEE en México), así como del proceso de toma de decisiones en materia de política educativa.

De esta manera, la evaluación (y en particular la evaluación a gran escala) representa una herramienta que ha permitido la consolidación de sistemas educativos basados en la rendición de cuentas, ya que favorecen la legitimación de una educación enfocada en la competencia económica. Por supuesto, bajo una política gubernamental globalizadora, ello responde a motivaciones de corte internacional. Como lo comenta Cepeda (2009) en su tesis doctoral, a consecuencia de la Declaración Mundial sobre Educación para Todos en el año de 1990, la UNESCO (a través de la Oficina Regional de Educación para América Latina y el Caribe, OREALC) desarrolló un discurso enfocado en evaluación, orientado particularmente en la definición de acciones de medición de resultados educativos que facilitarían la comparación entre escuelas e informar públicamente los resultados. De la misma manera, el Banco Mundial señaló la importancia de contar con mediciones para valorar la calidad del aprendizaje (Banco Mundial, 2000, citado en Cepeda, 2009). De esta forma se fomentaría que los países participaran en evaluaciones internacionales sobre desempeño educativo y se fomentara la creación de instituciones evaluadoras que ayudarían a consolidar una política de rendición de cuentas en función de los resultados obtenidos en educación. De igual modo, el Banco Interamericano de Desarrollo (2010) recomendó en 1996 la reestructuración de los sistemas educativos, argumentando la necesidad de que la formación de estudiantes se enfocara en el desarrollo de competencias entre escuelas, enfatizando así el discurso en la evaluación de la educación y sugiriendo orientar esta actividad hacia la participación de los países en proyectos internacionales con fines de obtener puntos de referencia para establecer metas nacionales de aprendizaje.

Claramente, en los casos mencionados, el fomento por la evaluación a gran escala se da por organismos que impulsan la apertura económica de los países en vías de desarrollo a los capitales extranjeros y, dado que la evaluación internacional fomenta la competencia económica, las sugerencias realizadas hacia los países en desarrollo favorecerán la apertura económica (Cepeda, 2009). Las políticas educativas no están entonces alejadas de las políticas económicas globalizadoras, por lo que en los discursos oficiales se busca que la enseñanza favorezca que los jóvenes sean competentes para hacer frente a un mercado mundial. Bajo esta mirada, en los años noventa se enfatizó el uso de pruebas a gran escala como el mejor indicador del avance de los sistemas educativos orientados a las competencias requeridas para una competitividad económica. A partir de esta visión es que esta práctica evaluativa se comenzó a considerar como una valiosa herramienta de monitoreo de las políticas educativas influenciadas por procesos de globalización económica.

En este sentido, la cultura mundializadora (Villaseñor, 2003) que tiene como eje la economía, permea a otras esferas de lo público y social, ya que es necesario que las instituciones (entre ellas las educativas) orienten sus esfuerzos a la obtención de los mayores beneficios económicos posibles. Al respecto, algunos de los proyectos de evaluación a gran escala (como los ya mencionados PISA y LLECE) se encuadran en la llamada sociedad del conocimiento, la cual, de acuerdo con Villaseñor (2003), se manifiesta también en la Economía del conocimiento. Es decir, el conocimiento que se genere se orienta a la producción de bienes y servicios, así como a su comercialización; de esta forma, el conocimiento se convierte en un elemento importante para la obtención de riqueza. Entonces, otro elemento de relevancia para la proliferación en el uso de la evaluación a gran escala también se da en este sentido, ya que es la herramienta que permite valorar si los sistemas educativos están cumpliendo con la función de preparar a los estudiantes para hacer frente a las demandas de competitividad que la economía del conocimiento implica.

La manera en que este tipo de evaluaciones influyen en acciones educativas en concreto puede verse reflejado en varios ejemplos. En España se han establecido algunos programas educativos como consecuencia de los resultados de PISA, como el Plan Nacional para Promover la Lectura (Bonafant y Tarabini, 2013). Por su parte, los análisis de Pereyra (2014), Tröhler (2009) y Meyer y Benavot (2013) dejan al descubierto la manera en que PISA ha sido un instrumento clave en la búsqueda de unificación de sistemas educativos bajo la mirada de políticas económicas neoliberales. PISA representa un medio de gobernanza educativa de los sistemas educativos, estableciendo las naciones cuyos estudiantes se asemejan al ideal establecido por las naciones que controlan el 80% del PIB mundial (Pereyra, 2014). Mediante este instrumento a gran escala la OCDE ha logrado que los países que presentan la prueba orienten sus políticas educativas hacia donde la organización indica, debido a que los niveles de desempeño se desarrollan bajo un marco de competencia europeo. Dicho estándar se convierte entonces en el ideal educativo por alcanzar por parte de los ministerios de Educación.

Bajo esta lógica, la evaluación del aprendizaje a gran escala es una herramienta importante para las decisiones políticas relacionadas con la educación, ya que permite que el Estado valore la calidad de la educación a partir de los resultados de sus estudiantes en las pruebas aplicadas. En este sentido, la calidad irá en función de una estrategia educativa basada en la competencia, influenciada por políticas internacionales. Evidentemente, es un modelo de evaluación que favorece la implementación de acciones orientadas hacia la rendición de cuentas de un Estado, la cual establece como eje principal de competencia y calidad la observación del cumplimiento de estándares que el sistema educativo determina como adecuados.

4. Evaluación a gran escala en el sistema educativo mexicano

En el caso de México, la historia de la evaluación educativa de manera formal no es tan amplia como en otros países, como en Estados Unidos, por ejemplo. Sin embargo, los aspectos políticos que dieron pauta a su auge no distan mucho de lo ocurrido en otros contextos. Si bien desde la década de los sesenta en México comenzaron a utilizarse pruebas estandarizadas para la admisión de estudiantes a instituciones como la UNAM, la UIA y el ITESM (Martínez, 2008a), a partir de los años ochenta, a pesar de los

problemas económicos por los que atravesaba el país, México se dio a la tarea de enfrentar la masificación de la educación, así como la necesidad de atender las demandas de competitividad (Cepeda, 2009; Martínez, Cepeda, Fuentes y Burgos, 1995). Para explicar con mayor detalle la manera en que en México la evaluación (y en particular la evaluación a gran escala) comenzó a cobrar relevancia, vale recordar de forma general las políticas educativas del país en las últimas décadas.

Considerando la información contenida en la investigación desarrollada por Cepeda (2009), en México las políticas públicas en relación a la educación y que impactan directamente sobre los procesos de evaluación datan de los años setenta. A inicios de los años setenta ya se destinaba una gran proporción del Producto Interno Bruto del país a la educación, incluso comienza a plantearse la posibilidad de incrementar la educación básica hasta el nivel secundaria y el tema de la preocupación por la calidad aparece dentro del discurso público. Como parte de ese interés, en 1971 se crea el Instituto de Investigación Educativa, encargado de conformar estadísticas sobre la cobertura y reprobación en la educación básica, incluyendo también estudios explicativos sobre ese tipo de problemas. También se crea la Subsecretaría de Planeación Educativa dentro de la Secretaría de Educación Pública (SEP) y en 1974 aparece la Subdirección de Evaluación y Acreditación, área que comienza a desarrollar pruebas para el ingreso de alumnos a secundaria (aunque se habían iniciado desde 1972).

Ya en los años ochenta, debido a la crisis económica de principio de década en el gobierno de Miguel de la Madrid, el país comienza una importante apertura económica hacia otras naciones. Como consecuencia, hubo un debilitamiento de la industria local, un control de la inflación y la privatización de empresas públicas, así como una dependencia a instituciones crediticias, como el Fondo Monetario Internacional. El impacto de estas decisiones en términos educativos se manifestó en enfocar un interés hacia la calidad, eficiencia terminal y cobertura que el sistema educativo podía lograr. Desde ese momento comienza a valorarse la necesidad de un Sistema Nacional de Evaluación Educativa para sustentar mediante métodos objetivos la toma de decisiones en materia de política educativa enfocada a la calidad. Como consecuencia de esta intención, comenzaron a desarrollarse pruebas con fines de medición con diversos objetivos; por ejemplo, en el año 1984 se desarrolló el Instrumento de Diagnóstico y Clasificación para el Ingreso a la Educación Normal (IDCIEN).

Ya para el inicio de los años noventa se formalizó aún más el ingreso de empresas extranjeras a México a través del Tratado de Libre Comercio y, además, el país se incorporó como nación miembro a la OCDE. Con la finalidad de que el país se mantuviera competitivo frente a otros países en el marco de esos compromisos, se inició el Programa para la Modernización Educativa en 1992, enfocado en una reestructuración del sistema educativo y, a su vez, explicitando la necesidad de identificar las causas de la deserción educativa y el rezago escolar. En este marco se desarrolla el Diagnóstico para Alumnos de Nuevo Ingreso a la Educación Secundaria (IDANIS) y el programa de Carrera Magisterial, un programa de estímulos para docentes (Horbath y Gracia, 2014). También se desarrolla el Programa para Abatir el Rezago Educativo, realizando evaluaciones de aprendizajes en el aula, así como las Olimpiadas del Conocimiento Infantil, que comenzaron en 1993 y que implicaban ya una evaluación basada en el currículum. Entre los años 1995 y 2000 se realizó el Acuerdo Nacional para la Modernización de la Educación Básica, así como el Programa

Evaluación de la Educación Primaria en 1996, con la idea de generar estadísticas de tendencias educativas de 3° y 6° de Primaria.

De gran importancia fue la creación de la Ley General de Educación de 1993, la cual, bajo este contexto de competencia económica internacional, resalta la evaluación de la educación. Incluso, en el Plan Nacional de Desarrollo de la gestión de Carlos Salinas de Gortari, se señaló que el camino por seguir para determinar el financiamiento de las instituciones de educación superior sería a través de la evaluación. Como consecuencia se instauraron diversos organismos dedicados a la evaluación: por ejemplo, la Comisión Nacional de Evaluación (CONAEVA), los Comités Interinstitucionales para la Evaluación de la Educación Superior (CIEES) y el Centro Nacional de Evaluación para la Educación Superior (Ceneval) (Villaseñor, 2003). Adicional a los mecanismos de financiamiento que la evaluación educativa permitía establecer, uno de los mayores logros de esta política pública fue la implementación de un imaginario social de relación entre evaluación y financiamiento que posteriormente permitiría la asimilación de la sociedad de la evaluación como elemento fundamental de orientación hacia la mejora educativa.

Entre los años noventa y dos mil comenzó un especial interés en el desarrollo de instancias externas de evaluación de los resultados educativos de los estudiantes mexicanos. Con la creación del Ceneval se unificaron los procesos de selección de alumnos a la educación media superior de la zona metropolitana de la Ciudad de México a través de la Comisión Metropolitana de Instituciones Públicas de Educación Media Superior (COMIPEMS). Bajo la misma lógica, en el año 2002 se crea el INEE, que comenzó a desarrollar las actividades de evaluación que la SEP había llevado a cabo desde años antes (Horbath y Gracia, 2014), la cual, como se comentó, adquiriría posteriormente autonomía constitucional.

A partir del 2004 el INEE diseñó los Exámenes de Calidad y Logro Educativo (EXCALE) con el fin de informar sobre el avance del sistema educativo básico como conjunto, no respecto a cada estudiante en lo particular, así como identificar factores asociados al aprendizaje para brindar información para la toma de decisiones políticas. Por su parte, la SEP en el sexenio 2000-2006 desarrolló la Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE). La prueba se desarrolló con el interés formativo de identificar el logro académico de alumnos de Primaria y Secundaria en algunas asignaturas clave con base en los planes y programas de la SEP, con el fin de atender las necesidades de cada alumno. Esta evaluación se consideró de forma censal con la finalidad de que los padres de familia y docentes tuvieran información directa sobre los logros alcanzados por los estudiantes, aprovechando así la experiencia que la Dirección General de Evaluación de Políticas de la SEP había alcanzado con evaluaciones desarrolladas desde los años setenta (SEP, 2007). Los objetivos primordiales de esta evaluación se centraban no solo en que los docentes y padres de familia tuvieran información sobre el avance de los logros educativos de sus hijos, sino que se desarrollaba con la intención de contribuir al desarrollo de mejores prácticas educativas desde el salón de clases, facilitando con esta información la toma de decisiones del proceso educativo y, con ello, “fortalecer la convergencia de esfuerzos nacionales y estatales a través del Sistema Nacional de Pruebas de Aprendizaje, de la Política Nacional de Evaluación y del Sistema Nacional de Evaluación” (SEP, 2007, p. 13). ENLACE se aplicó durante siete años seguidos en todo México a estudiantes tercero, cuarto, quinto y sexto de Primaria en las asignaturas: Español, Matemáticas y

Formación Cívica y Ética. También se aplicó en los tres años de Secundaria en las mismas asignaturas y cada año de aplicación se centraba en un área en particular.

La prueba ENLACE estuvo siempre sujeta a diversos comentarios por parte de la opinión pública: tanto positivos, considerándolo una medición relevante del estatus del logro de los estudiantes y necesaria para el conocimiento de la realidad educativa dentro de las escuelas del país, como negativos, debido a su uso enfocado únicamente como factor de reconocimiento social de las escuelas participantes. En enero del año 2014, el Secretario de Educación en turno, Emilio Chuayffett, anunció la cancelación de ENLACE, de manera que el INEE, ya con autonomía constitucional, sería el encargado de elaborar y aplicar las pruebas educativas en México para continuar con el monitoreo del logro de los estudiantes. Frente a ello, en el año 2015 el INEE sustituyó EXCALE y ENLACE por el Plan Nacional para la Evaluación de los Aprendizajes (Planea). Planea es una evaluación a gran escala enfocada en la educación obligatoria que pretende informar a la sociedad sobre el estado de la educación en función del logro de los estudiantes, así como brindar información a las autoridades educativas para el monitoreo y operación del sistema educativo de México y, por supuesto, contribuir al desarrollo de directrices sobre política educativa (INEE, 2016).

En México existe también el Ceneval, institución que desde 1994 comenzó, bajo la iniciativa de la Asociación de Universidades e Institutos de Educación Superior (ANUIES), a fungir como un agente externo de evaluación de la educación superior. Desde el año de su creación, el Ceneval desarrolla y califica el Examen Nacional de Ingreso a la Educación Media Superior (EXANI-I) a solicitud de la COMIPEMS (Hernández, Márquez y Palomar, 2006). Dicho instrumento permite desde hace más de 20 años tener un referente objetivo para la toma de decisiones respecto al ingreso a educación media superior en México, ya que actualmente se aplica en las instituciones públicas y privadas que así lo solicitan en todo el país. El EXANI también tiene dos versiones más, una centrada en el ingreso a educación superior (EXANI-II) y otra en estudios de posgrado (EXANI-III). Además del EXANI, el Ceneval desarrolla otro tipo de instrumentos a gran escala, incluso los elabora de acuerdo con las solicitudes propias de las instituciones educativas. Uno de los proyectos más grandes del centro es el Examen General para el Egreso de la Licenciatura (EGEL), el cual se centra en la medición de conocimiento y habilidades mínimos necesarios para el ejercicio profesional. De esta manera los EXANI favorecen la regulación del ingreso de estudiantes, mientras que los EGEL favorecen la del egreso de los programas educativos. En el proceso de consolidación de la evaluación a gran escala como herramienta de los sistemas educativos basados en competencia y en rendición de cuentas, los instrumentos desarrollados por el Ceneval favorecen la consolidación de una política de competitividad. De esta forma, la función de evaluación y acreditación de la calidad de la educación es un papel que el Estado canaliza a diversas agencias —en particular en educación superior—, fenómeno que se desarrolla en América Latina y el Caribe (en particular en Argentina, Chile, Brasil, Paraguay, Bolivia, Ecuador, Colombia, Costa Rica, El Salvador, México, Cuba, Trinidad y Tobago) (De la Garza, 2008).

Considerando las evaluaciones a gran escala internacionales, resalta la influencia del Banco Mundial y de la OCDE en programas estratégicos del sistema educativo en México. Por ejemplo, el Programa de desarrollo de la educación básica es un préstamo del Banco Mundial dirigido a apoyar la educación compensatoria en México, con el fin de buscar que la educación mexicana se acerque al ideal de los indicadores

internacionales establecidos por la OCDE. En particular, con este programa se ha apoyado al Consejo Nacional de Fomento Educativo (CONAFE) (Banco Mundial, 2005). Mediante esta acción el Banco Mundial afirma que está acercándose al cumplimiento de su objetivo de mejorar la calidad de la educación (entendida como la mejora del resultado de los estudiantes en evaluaciones estandarizadas) mediante los apoyos que brinda al Estado mexicano.

Por otra parte, mediante el financiamiento otorgado por el Banco Internacional de Reconstrucción y Fomento (BIRF, dependiente del Banco Mundial) se obtuvieron donaciones y becas y se apoyó a la evaluación de la educación que permitió una mejora en los resultados en la prueba PISA. Incluso, desde los años noventa, el Banco Mundial ha apoyado el Préstamo para Políticas de Desarrollo para la Educación Secundaria, con la finalidad de mejorar el acceso a la educación para mantener su competitividad (Banco Mundial, 2010). De esta forma, como lo comenta Jessop (2008, citado en Bonal y Tarabini, 2013), los tomadores de decisiones han hecho uso de PISA para construir nuevas formas de hegemonía. Además, este es un claro ejemplo de como la búsqueda de mejorar en los puntajes en evaluaciones se ha convertido en la principal meta de los cambios que los gobiernos realizan en materia educativa (Saura y Luengo, 2015).

De manera similar, la OCDE, como agente de gobernanza en la educación mexicana, ha emitido diversas recomendaciones que han facilitado la consolidación del uso de la evaluación a gran escala como medio de regulación educativa. De acuerdo con la OCDE (2010a), los elementos básicos que deben seguir los gobiernos son, entre otros aspectos, un sistema de rendición de cuentas basado en información objetiva de los resultados. Adicionalmente, se establece la necesidad de una reforma educativa en México basada en el hecho de que una gran proporción de estudiantes no finaliza la educación media superior, además de los bajos resultados mostrados en PISA. De la misma manera, dicho organismo sugiere continuar con el desarrollo de indicadores de resultados educativos nacionales (como ENLACE y EXCALE), ya que favorecen la atención a la mejora (OCDE, 2010b).

Este tipo de recomendaciones han facilitado la implementación de la evaluación a gran escala con fines de homogeneización de la educación. Según Anthony (2007), México debe trabajar para que sus estudiantes alcancen un nivel similar al de otros países de la OCDE, ya que la evaluación (junto con la autonomía de las escuelas y la rendición de cuentas de las mismas) ayudará a la mejora de la educación, propiciando así que la implementación de políticas sea más efectiva. Ello se ve reflejado en el desarrollo de programas tales como: Programa de Escuelas de Calidad (orientado a fortalecer la gestión escolar y el capital social de las escuelas), Programa Nacional de Lectura (que busca el fomento de la lectura y mejorar las prácticas pedagógicas relacionadas con ello) y el Programa Nacional de Formación Continua de los Maestros en Educación Básica (que pretende fortalecer la práctica docente). Estos se han desarrollado con la intención de mejorar el aprendizaje de los alumnos, aunque para ello es necesario establecer estándares de logro de los estudiantes y las escuelas como referente de su calidad (Miranda, 2007), lo que, como se comentó, son herramientas que favorecen la consolidación de un sistema educativo basado en la rendición de cuentas.

Previo a la Reforma Educativa mexicana que facilitó la autonomía del INEE en el año 2013, la OCDE (Santiago, McGregor, Nusche, Ravela y Toledo, 2012) sugirió continuar con el desarrollo de las evaluaciones nacionales como forma de comparación entre las calificaciones obtenidas por los estudiantes y los estándares nacionales que estas pruebas definen. También se sugirió a México continuar con el uso formativo de

evaluaciones a gran escala, con el fin de que los docentes reflexionen sobre los procedimientos necesarios para mejorar su enseñanza. En estas recomendaciones se da tal relevancia a la evaluación a gran escala que se sugiere evitar los enfoques normativos de evaluación (es decir, aquellos que se realizan dentro del aula por parte de los docentes) y centrarse en la relevancia de la evaluación por criterios, argumentando que algunos de los elementos clave para el éxito de una reforma educativa parten de la introducción exitosa de la alineación de los resultados de aprendizaje de los estudiantes y los estándares en el plan de estudios.

Considerando estos elementos históricos y de contexto global, la Reforma Educativa vivida en México en el año 2013 da un lugar preponderante a la evaluación de la educación. Esta tomó aún mayor fuerza que la mostrada en décadas previas, pues se considera que facilita la organización de todo el sistema educativo, así como el conocimiento del estatus *real* de la educación mexicana, para la implementación de políticas públicas basadas en evidencias o *policy by number*. De esta manera, los proyectos de evaluación a gran escala que se han generado desde la década de los ochenta y aquellos que se aplican en esta década representan una forma de control de la gestión del sistema educativo, en virtud de que es una herramienta que favorece el establecimiento de estándares unívocos, los cuales, a su vez, devienen de las políticas basadas en competencia consecuentes de la dinámica económica globalizadora. De esta forma, los resultados de las evaluaciones a gran escala fomentan esa determinación de un estándar o ideal de los resultados de aprendizaje de los jóvenes orientado hacia las dinámicas económicas globales, persiguiendo así una homogeneización de la formación de los jóvenes mexicanos, ya sea con fines de alcanzar estándares internacionales, o bien como medio de regulación interna del sistema educativo.

5. Conclusiones

Como se puede observar, el interés en evaluación como elemento al servicio del Estado mexicano se da a partir de la política educativa de dicho país a principios de los ochenta, ya que permitiría identificar problemas y jerarquizar las necesidades específicas de los ámbitos educativos con el fin de alcanzar la competitividad establecida por recomendaciones internacionales. Adicionalmente, desde estos años se vislumbraba ya la creciente demanda de educación como forma de movilidad social, la cual también podría ser regulada mediante la evaluación a gran escala, como sucede con el EXANI y el EGEL del Ceneval.

La evaluación a gran escala tiene entonces un papel crucial en el desarrollo de la política educativa mexicana, ya que asegura la verificación de que el aprendizaje que se pretendía enseñar mediante los programas educativos ha sido adquirido por los estudiantes, lo cual constituiría el elemento más palpable mediante el cual el sistema educativo podrá valorar sus propios avances hacia una sociedad globalizada. La evaluación a gran escala enfocada a los resultados de los estudiantes se convirtió en una herramienta efectiva de instrumentación de la cultura de la competencia enfocada a cumplir estándares internacionales. En México, además de los estudios como Planea del INEE, también hay participación en evaluaciones de corte internacional, tal es el caso de PISA o los estudios desarrollados por el LLECE, los cuales han tomado una importante fuerza política, dada su cualidad de analizar el avance del sistema educativo mexicano en términos de los resultados de aprendizaje de los estudiantes y su comparación con los de otras naciones.

De esta manera, la evaluación a gran escala en México ha tenido como fundamento de consolidación la intención de toma de decisiones, en particular como elemento de implementación de las políticas educativas orientadas a alcanzar estándares recomendados por instancias internacionales.

La consideración de la evaluación a gran escala como medio de implementación de políticas de competitividad y rendición de cuentas resulta inevitable, ya que las políticas educativas en torno a la evaluación permiten legitimar decisiones de gobernanza y ningún país ni su población puede mantenerse alejado de la dinámica económica globalizada. Frente a este panorama, vale repensar la utilidad de las pruebas a gran escala bajo un enfoque pedagógico centrado en lo formativo; es decir, valorar los resultados bajo una mirada crítica que tenga presente el papel de este tipo de evaluaciones. Resulta necesario considerar que la calidad de la educación no solo estará en función de factores exógenos (como los cataloga Márquez, 2004) que se encuadran únicamente en lo que socialmente se espera de la educación, en particular bajo un contexto globalizado. Como consecuencia, hay que recordar que la educación y su calidad no pueden referirse únicamente a los resultados de aprendizaje de los estudiantes bajo estándares unívocos (sobre todo en consideración de una sociedad tan desigual, como es el caso de la sociedad mexicana), sino también en función de la infraestructura escolar o el acceso y la permanencia en la escuela, entre otros elementos. Vale también considerar modelos de rendición de cuentas orientados hacia la responsabilidad que los sistemas educativos tienen sobre la población. En esa línea, Darling-Hammond y Snyder (2015) proponen un modelo orientado a facilitar el fomento a la responsabilidad que los sistemas educativos tienen sobre el aseguramiento de la educación que los niños merecen. Este tipo de propuestas pueden generar una alternativa que permita avanzar hacia la creación de sistemas alineados de evaluación y responsabilidad educativa. Resulta necesario considerar la posibilidad de reflexión sobre este tipo de alternativas, que ayudarán a valorar la evaluación a gran escala no solo como una herramienta para el control y la consolidación de políticas gubernamentales neoliberales.

Referencias

- Anderson, J. (2005). *La rendición de cuentas en la educación*. París: Instituto Internacional para la Planeación de la Educación.
- Anthony, H. (2007). Factores determinantes del aprendizaje y calidad de la educación en México. En F. Miranda, H. Anthony y A. López (Coord.), *Mejora de la calidad educativa en México. Posiciones y propuestas* (pp. 13-26). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- Banco Interamericano de Desarrollo. (2010). *Reforma a la educación primaria y secundaria en América Latina y el Caribe*. Nueva York: BID.
- Banco Mundial. (2010). *México: Fortalecimiento del aporte de la educación a la economía del conocimiento*. Recuperado de <http://web.worldbank.org/WBSITE/EXTERNAL/BANCOMUNDIAL/NEWSSPANISH/0,,contentMDK:22899181~menuPK:3327642~pagePK:34370~piPK:34424~theSitePK:1074568,00.html>
- Bonal, X. y Tarabini, A. (2013). The role of PISA on shaping hegemonic educational discourses, policies and practices: The case of Spain. *Research in Comparative and International Education*, 8(3), 335-341. doi:10.2304/rcie.2013.8.3.335.

- Calero, J. y Choi, A. (2012). La evaluación como instrumento de política educativa. *Presupuesto y Gasto Público*, 67, 29-42.
- Centro Latinoamericano de Administración para el Desarrollo. (2007). Una nueva gestión pública para América Latina. En C. H. Acuña (Comp.), *Lecturas sobre el Estado y las políticas públicas: Retomando el debate de ayer para fortalecer el actual* (pp. 403-424). Buenos Aires: Jefatura del Gabinete de Ministros.
- Cepeda, B. (2009). *La política de evaluación externa de los aprendizajes en educación básica. Una aproximación a la comprensión de su racionalidad. Periodo 2000-2006* (Tesis doctoral no publicada). Universidad Nacional Autónoma de México, Ciudad de México.
- Darling-Hammond, L. y Snyder, J. (2015). Meaningful learning in a new paradigm for educational accountability: An introduction. *Education Policy Analysis Archives*, 23(7). doi:10.14507/epaa.v23.1982
- DePascale, C. (2003). The ideal role of large-scale testing in a comprehensive assessment system. *Association of Test Publishers*, 5(1), 1-11.
- De la Garza, J. (2008). Evaluación y acreditación de la educación superior en América Latina y el Caribe. En C. Tünnermann (Ed.), *La educación superior en América Latina y el Caribe: Diez años después de la Conferencia Mundial de 1998* (pp. 175-222). Cali: Sello Editorial Javeriano.
- Díaz Barriga, A. (1994). Una polémica en relación al examen. *Revista Iberoamericana de Educación*, 5, 161-181.
- Fuhrman, S. y Elmore, R. (Eds.). (2004). *Redesigning accountability systems for education*. Nueva York: Teachers College Press.
- Héritier, A. (2011). New modes of governance in Europe: Policy-making without legislating? *Preprints aus der Max-Planck-Projektgruppe Recht der Gemeinschaftsgüter*, 14.
- Hernández, J., Márquez, A. y Palomar, J. (2006). Factores asociados con el desempeño académico en el EXANI-I. *Revista Mexicana de Investigación Educativa*, 11(29), 547-581.
- Horbath, J. y Gracia, M. (2014). La evaluación educativa en México. *Revista de Relaciones Internacionales, Estrategia y Seguridad*, 9(1), 59-85.
- Instituto Nacional para la Evaluación de la Educación. (2013). *México en PISA 2012*. Ciudad de México: Autor.
- Instituto Nacional para la Evaluación de la Educación. (2016). *Planea: Una nueva generación de pruebas. ¿Cómo y cuándo se evalúa?* Recuperado de http://www.inee.edu.mx/images/stories/2015/planea/fasciculosnov/Planea_2.pdf
- Jessop, B. (2002). Liberalism, neoliberalism, and urban governance: A state-theoretical perspective. *Antipode*, 34(3), 452-472. doi:10.1111/1467-8330.00250.
- Khem, B. M. (2011). Gobernanza, ¿qué es? ¿Es importante? En B. M. Khem (Ed.), *La gobernanza en la enseñanza superior* (pp. 21-25). Barcelona: Universitat de Barcelona.
- Mann, M. (2007). El poder autónomo del Estado: Sus orígenes, mecanismos y resultados. En M. Mann (Ed.), *Lecturas sobre el Estado y las políticas públicas* (pp. 55-78). Buenos Aires: Jefatura del Gabinete de Ministros.
- Martínez, F. (2001). Evaluación educativa y pruebas estandarizadas. Elementos para enriquecer el debate. *Revista de la Educación Superior*, 30(4), 71-85.
- Martínez, F. (2008a). *Evaluación a gran escala y evaluación en el aula. Los usos de Excale para la mejora educativa*. Recuperado de <http://www.inee.edu.mx/explorador/propuesta2008/Textos/EvaluacionGranEscala.pdf>

- Martínez, F. (2008b). *La evaluación de aprendizajes en América Latina*. Ciudad de México: Instituto Nacional para la Evaluación de la Educación.
- Martínez, F. y Santos, A. (2009). Evaluación e investigación educativa. En A. Alba y R. Glazman (Coord.), *¿Qué dice la investigación educativa?* (pp. 265-304). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- Martínez, F., Cepeda, B., Fuentes, G. y Burgos, R. (1995). Evaluación del aprendizaje. En M. Rueda (Coord.), *La investigación educativa en los ochenta, perspectivas para los noventa. Procesos de enseñanza y aprendizaje I* (pp. 11-19). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- Márquez, A. (2004). Calidad de la educación superior en México. ¿Es posible un sistema eficaz, eficiente y equitativo? Las políticas de financiamiento de la educación superior en los noventa. *Revista Mexicana de Investigación Educativa*, 9(21), 477-500.
- Meyer, H. y Benavot, A. (2013). *PISA, power and policy: The appear of global governance of education*. Oxford: Symposuim Books.
- Ministerio de Educación de Chile. (2011). *Resultados de TIMSS 2011. Chile. Estudio Internacional de Tendencias en Matemática y Ciencia*. Santiago de Chile: Autor.
- Miranda, F. (2007). Mejora del aprendizaje y la calidad de la educación: Una visión política educativa. En F. Miranda, H. Anthony y A. López (Coord.), *Mejora de la calidad educativa en México. Posiciones y propuestas* (pp. 119-138). Ciudad de México: Consejo Mexicano de Investigación Educativa.
- O'Donnell, G. (2008). Algunas reflexiones acerca de la democracia, el Estado y sus múltiples caras. *Revista del CLAD Reforma y Democracia*, 42, 5-30.
- Organización para la Cooperación y el Desarrollo Económicos. (2010a). *Acuerdo de cooperación México-OCDE para mejorar la educación de las escuelas mexicanas. Resúmenes ejecutivos*. París: Autor.
- Organización para la Cooperación y el Desarrollo Económicos. (2010b). *Mejorar las escuelas. Estrategias para la acción en México. Resumen ejecutivo*. París: Autor.
- Patrinos, H. A., Shapiro, J. y Moreno Trevino, J. (2005). *Educación compensatoria para estudiantes desfavorecidos: Pruebas de un estudio de evaluación de impacto llevado a cabo en México*. Washington D. C.: Banco Mundial.
- Pereyra, M. (2014, septiembre). PISA a examen: ¿Evento mediático, nueva forma de gobernanza, o propuesta de cambio escolar? Comunicación presentada en el *Congreso Internacional de Educación*. Evaluación. Universidad Autónoma de Tlaxcala, Tlaxcala.
- Ravela, P. (2006). *Para comprender las evaluaciones educativas. Fichas didácticas. Programa de promoción de la Reforma Educativa en América Latina y el Caribe*. Recuperado de http://www.iae.org.mx/2012/Dip1DGIFA/nov1/S102_Ravela_Pedro_Fichas_didacticas.pdf
- Santiago, P., McGregor, I., Nusche, D., Ravela, P. y Toledo, D. (2012). *Revisiones de la OCDE sobre evaluación en la educación*. Ciudad de México: INEE.
- Saura, G. y Luengo, J. (2015). Política global más allá de lo nacional. Reforma educativa (LOMCE) y el régimen de estandarización (OECD). *Bordón*, 67(1), 135-148. doi:10.13042/bordon.2015.67109.
- Schulz, W., Ainley, J., Fraillon, J., Kerr, D. y Losito, B. (2010). *ICCS 2009 International report: Civic knowledge, attitudes, and engagement among lower-secondary school students in 38 countries*. Amsterdam: IEA.

- Secretaría de Gobernación. (2013). DECRETO por el que se reforman los artículos 3o. en sus fracciones III, VII y VIII; y 73, fracción XXV, y se adiciona un párrafo tercero, un inciso d) al párrafo segundo de la fracción II y una fracción IX al artículo 3o. de la Constitución Política de los Estados Unidos Mexicanos. Ciudad de México: Diario Oficial de la Federación.
- Secretaría de Educación Pública. (2007). *Manual técnico ENLACE*. Ciudad de México: Autor.
- Solano-Flores, G. y Bonk, W. (2008). *Evaluation of the Latin American Laboratory for the Evaluation of Educational Quality (LLECE)*. Santiago de Chile: OREAC/UNESCO.
- Tröhler, D. (2009). Armonizar el globo educativo. Política mundial, rasgos culturales y los desafíos de la investigación educativa. *Profesorado. Revista de Currículum y Formación del Profesorado*, 13(2), 1-14.
- Trow, M. (1996). Trust, markets and accountability in higher education: A comparative perspective. *Higher Education Policy*, 9(4), 309-324. doi:10.1016/s0952-8733(96)00029-3.
- Vargas, A. (2002). La acreditación: Una forma de estandarizar la educación. *Revista Educación*, 26(2), 245-254. doi:10.15517/revedu.v26i2.2923.
- Villaseñor, G. (2003). La evaluación de la educación superior: Su función social. *Reencuentro*, 36, 20-29.
- Viñao, A. (2012). El desmantelamiento del derecho a la educación: Discursos y estrategias neoconservadoras. *Revista Internacional de Ciencias Sociales*, 31, 97-107.
- Winkler, D. (2006). Fortalecer accountability (rendición de cuentas) en la educación pública. En J. Corvalán y R. McMeekin (Eds.), *Accountability educacional: Posibilidades y desafíos para América Latina a partir de la experiencia internacional* (pp. 99-123). Santiago de Chile: PREAL.