

Una Propuesta de Ejes para el Análisis del Uso de Resultados de Evaluaciones a Gran Escala del Aprendizaje por Docentes y Directores

A Proposal of Axes for the Analysis of the Use of Results of Large-Scale Evaluations of Learning by Teachers and Principals

Guadalupe Ruiz Cuéllar *
Paloma Prado Robledo

Universidad Autónoma de Aguascalientes - México

Las evaluaciones a gran escala forman parte de los esfuerzos de los sistemas educativos por identificar su situación actual, y, según el caso, aprovechar sus resultados para fines de rendición de cuentas o de mejora pedagógica. Para generar en efecto algún impacto tendiente a la mejora, se requieren procesos de difusión y uso de resultados que aseguren el aprovechamiento de esta información. Este artículo presenta una síntesis de la literatura revisada en el marco de un estudio realizado en Aguascalientes, México, durante 2013, que tuvo como objetivo general, describir los usos que supervisores, directores y docentes hacían de la información derivada de la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE), vigente en el sistema educativo mexicano precisamente hasta ese año. A partir de esa revisión y habiéndolos utilizado en ese estudio, se proponen cinco ejes que pueden orientar la obtención y análisis de información en investigaciones futuras sobre el tema: difusión de la información producida por las evaluaciones a gran escala; análisis e interpretación de la información; uso de los resultados en tres niveles: prácticas docentes, gestión escolar y supervisión; efectos del uso de resultados y factores asociados al aprovechamiento de la información.

Palabras clave: Evaluación educativa; Evaluación del aprendizaje; Evaluaciones estandarizadas; Programa de evaluación; Educación básica; Calidad de la educación.

Large-scale assessments are one of the efforts to identify the current situation of the educational systems, and, depending on the case, build on their results for accountability purposes or pedagogical improvement. In order to promote improvement, there is a need to develop and implement specific processes for communicating and using large-scale assessments results. This paper is aimed at presenting a synthesis of the literature reviewed for a study undertaken in Aguascalientes, Mexico, during 2013. The main objective of the study was to describe how the information derived from the Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE) was used by supervisors, headteachers and teachers. ENLACE was a census large-scale in force in the Mexican educational system until 2013. As a result of this literature review, five axes are proposed for guiding the collection and analysis of information in future research: dissemination of information produced by large-scale evaluations; analysis and interpretation of information; use of results in teaching practices, school management and supervision; effects of the use of large-scale assessments results; and, factors associated with the use of information.

Keywords: Educational assessment; Learning assessment; Standardized assessment; Program evaluation; Basic education; Quality education.

*Contacto: gruiz@correo.uaa.mx

1. Introducción

En la actualidad, las evaluaciones a gran escala forman parte de las políticas educativas de un gran número de países, ya sea a través de operativos nacionales o de su participación en proyectos internacionales. Su importancia, entre otros aspectos, se cifra en su capacidad para hacer visibles aspectos fundamentales de los sistemas educativos, tales como el grado en que los alumnos aprenden lo que se espera que adquieran en cada grado o nivel; la medida en que el logro de los aprendizajes se produce de forma equitativa o inequitativa; la diversidad de prácticas educativas entre escuelas y profesores; o el acceso de los estudiantes al conocimiento y los aprendizajes, independientemente de los años de estudio que tengan. Asimismo, aportan información a los diversos actores educativos, misma que puede contribuir a la toma de decisiones y a definir estándares, niveles de desempeño o metas de aprendizaje (Ravela et al., 2007).

Si bien es cierto que estas evaluaciones son relevantes, debido a su capacidad para dar cuenta del estado de los sistemas educativos, también lo es que su aplicación no es suficiente para propiciar una mejora de la realidad que reflejan.

La evaluación no es, en sí misma, un fin. Emplear instrumentos objetivos de medición del aprendizaje que alcanzan los alumnos de una institución o sistema educativo tendrá valor en la medida en que los resultados se aprovechen para el mejoramiento de tales instituciones y sistemas (Martínez Rizo, 2001, p. 83).

Así pues, es necesario que la información derivada de las evaluaciones a gran escala sea difundida de tal manera que llegue a todos los actores educativos para los que está destinada o para quienes sería relevante disponer de la información, y que éstos participen en su análisis e interpretación y hagan un uso adecuado de ella, en aras de propiciar una mejora en las distintas esferas de la educación.

Hace casi dos décadas, Ravela, Wolfe, Valverde y Esquivel (2000) señalaban que una vez instalados formalmente los sistemas nacionales de evaluación en la región de América Latina, aún quedaba mucho por hacer en este ámbito. De manera particular, estos autores indicaban que, en materia de uso y difusión de los resultados de las evaluaciones, se requería “propiciar el desarrollo de trabajos de investigación que permitan recoger evidencia empírica acerca de los efectos que los distintos tipos de reportes de resultados de las evaluaciones nacionales tienen en diversos públicos” (p. 70). Asimismo, enfatizaban la importancia de recabar información sobre lo que los diversos actores educativos esperan de estos sistemas de evaluación, sus percepciones y comprensiones de la información que deriva de éstos, el grado en que la usan, así como sus expectativas y prejuicios acerca de los mismos.

En México, la evaluación externa de resultados de aprendizaje tuvo un gran auge en el periodo 2006-2013, a través de dos programas: la Evaluación Nacional del Logro Académico en Centros Escolares (ENLACE) y los Exámenes de la Calidad y el Logro Educativo (Excale), aplicados por la Secretaría de Educación Pública (SEP) y el Instituto Nacional para la Evaluación de la Educación (INEE), respectivamente (Backoff y Contreras, 2014; Martínez Rizo, 2015; Ruiz y Prado, 2017). Durante la presente administración federal (2012-2018), el énfasis de la política de evaluación se movió claramente de la evaluación de alumnos a la de docentes en el marco de la implementación en el país del Servicio Profesional Docente, si bien, la primera no dejó de realizarse, pero

a través de un nuevo instrumento, el Plan Nacional para la Evaluación de Aprendizajes (PLANEA), que busca superar las limitaciones de los programas de evaluación que le precedieron, sobre todo ENLACE.

Este artículo presenta una síntesis de la literatura revisada en el marco de un estudio realizado en Aguascalientes, México, durante 2013 -último año en que se aplicó ENLACE-, que tuvo como objetivo general, describir los usos que supervisores, directores y docentes hacían de la información derivada de esta evaluación, única que la producía de manera censal y que, por ende, llegaba a todas las escuelas participantes cada ciclo escolar.

A fin de mostrar de manera sistemática las contribuciones de la literatura científica que guiaron el análisis de la información obtenida en ese estudio, el artículo se divide en dos grandes apartados. En el primero se presentan algunas investigaciones acerca de la difusión y uso de resultados de evaluaciones a gran escala, desarrolladas tanto en América Latina como en México. En el segundo se proponen, incorporando investigaciones adicionales, cinco ejes para el estudio de la difusión y uso de los resultados de las evaluaciones a gran escala: 1) difusión de los resultados; 2) análisis e interpretación de los resultados; 3) uso de los resultados; 4) impacto del uso de los resultados y 5) factores asociados al aprovechamiento de la información de evaluaciones a gran escala. Este último eje es relevante, debido a que su indagación permitiría trascender la mera descripción de los usos y avanzar hacia la comprensión de las razones o factores que condicionan su aprovechamiento en la práctica.

A riesgo de reiterar, pero con el propósito de delimitar los alcances de esta contribución, es importante señalar que los ejes utilizados en el estudio sobre ENLACE, están sustentados en la revisión de los referentes teóricos y conceptuales, así como de los hallazgos de varios estudios previos, realizados tanto en México como en otros países, sobre la difusión y uso de resultados de evaluaciones externas. La pretensión del artículo no es, en sentido estricto, proponer un marco analítico para nuevas investigaciones sobre este objeto de estudio, pero sí aportar algunos elementos que parece relevante considerar por su capacidad informativa del tema, particularmente si el interés está puesto en estudiar los usos que pueden hacer de esos resultados profesores y directores, y si las pruebas a que corresponden dan prioridad al propósito de “brindar información oportuna y relevante a los docentes que les ayude a mejorar tanto su enseñanza como el aprendizaje de sus estudiantes (función pedagógica)” (Backhoff y Contreras, 2014, p. 1269).

Finalmente, cabe añadir que en lo que resta del texto se hace alusión de manera indistinta a los términos “evaluación” y “prueba”, pero se reconoce que no son equivalentes. Se coincide en este sentido con el argumento de Moreno (2011), quien precisa a propósito de un texto suyo:

Este trabajo se alude a las pruebas estandarizadas como evaluación porque así es como habitualmente se les denomina, pero es preciso advertir que no son sinónimos; las pruebas o exámenes son uno más de los diversos instrumentos que la evaluación puede emplear para la valoración de un determinado objeto. (p. 117)

2. Estudios previos

En América Latina se han realizado diversas investigaciones orientadas a dar cuenta de algunos aspectos relativos al uso de resultados de las evaluaciones a gran escala. Una de estas iniciativas fue llevada a cabo en 1999 por el Grupo de Trabajo sobre Estándares y

Evaluación (GTEE) del PREAL, cuyo objetivo fue analizar la situación de los sistemas nacionales de evaluación de diez países de la región: Argentina, Brasil, Bolivia, Costa Rica, Chile, Ecuador, México, Paraguay, Uruguay y Perú.

Fruto de este proyecto es el documento *¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?* (Ravela, 2001). Uno de los temas abordados, referido a la información producida y difundida por los sistemas nacionales de evaluación, es el de los usos que se pretende se hagan de este tipo de datos y si efectivamente se llevan a la práctica.

Los hallazgos de esta investigación muestran que la mayor parte de los sistemas nacionales de evaluación de los aprendizajes en América Latina son de carácter formativo y su uso es de bajo riesgo,¹ salvo algunas excepciones como Chile y México,² que implementaron sistemas cuyos usos eran de alto riesgo.³ Ambos tipos tienen desventajas, dado que los usos de consecuencias fuertes o alto riesgo pueden generar que los actores educativos dediquen una cantidad excesiva de tiempo a la preparación para las pruebas, reduciendo el currículo (Backhoff y Contreras, 2014, Ravela, 2001); mientras que los usos de bajo riesgo o la ausencia de consecuencias a partir de los resultados de las evaluaciones pueden propiciar poco interés y responsabilidad por los resultados de este tipo de pruebas. En las conclusiones de este trabajo el autor menciona, entre otros aspectos, que es importante realizar investigaciones acerca de los efectos que tiene el uso de las evaluaciones a gran escala para establecer incentivos.

El Programa de Promoción de la Reforma Educativa en América Latina y el Caribe (PREAL), con el apoyo del Banco Mundial abrió, en 2002, la primera convocatoria del Fondo de Investigaciones Educativas (FIE). El propósito fue financiar proyectos en torno a la siguiente pregunta: “¿Cómo se usa y qué impacto tiene la información empírica en el mejoramiento de los sistemas educativos en América Latina?” (Cueto, 2005, pp. 11-12). Los resultados de algunos de los estudios realizados se sintetizan enseguida.

Fernández y Midaglia (2005) analizaron el uso de informes de evaluación en dos países cuyos sistemas de evaluación tiene características diferentes: México y Uruguay. Para ello, se tomaron en cuenta los aspectos organizacionales de cada centro educativo y se generó una tipología de la gestión de la información. Fernández y Midaglia (2005) concluyeron

¹ Este tipo de evaluaciones tiene un carácter formativo y un uso de bajo impacto, esto se refiere a que se emplean principalmente para dar información acerca de los sistemas educativos, ayudar a la elaboración de políticas públicas y propiciar la mejora de las instituciones escolares, así como el desarrollo profesional de los docentes (Ravela, 2001).

² La investigación considera el caso del Programa “Carrera Magisterial” y la asignación que conllevaba de incentivos económicos para docentes, considerando, entre otros factores, los resultados de sus alumnos en la evaluación ENLACE. El sistema de evaluación que actualmente se aplica en México (PLANEA) no tiene asociado este uso.

³ Ravela (2001) señala que a diferencia de las evaluaciones de bajo riesgo, las de alto impacto tienen un carácter sumativo y se emplean principalmente para la acreditación y selección de alumnos, el establecimiento de incentivos para docentes e instituciones escolares y la creación de *rankings* de escuelas.

que en ambos países la lectura de informes de evaluación es escasa y, de manera general, los docentes que leen los informes son los que exhiben un mayor nivel de profesionalismo. La principal dificultad reportada por los docentes para leer los informes es la falta de orientaciones claras del director del centro o del supervisor.

Aristimuño y Kaztman (2005), a través de entrevistas en profundidad y grupos focales, indagaron en torno al uso y comprensión de los informes de evaluación del rendimiento de los estudiantes en Uruguay. Los autores describen las reformas educativas emprendidas en el país, especialmente a partir de 1995. Dos características importantes del sistema uruguayo consisten en que los informes de resultados son utilizados para impactar directamente en el trabajo de los maestros y directores, pues son retomados en cursos de formación en servicio y talleres, y en reuniones que tienen lugar en las mismas escuelas. Por otra parte, las escuelas que no han participado en alguna evaluación nacional reciben un informe de los resultados, así como la prueba para que ésta sea autoadministrada en las instituciones y comparada con varios referentes (por ejemplo, el promedio nacional o el promedio de las escuelas del mismo contexto sociocultural). Estas prácticas, según los autores, brindan legitimidad al operativo de evaluación entre docentes y directivos, sin embargo, encuentran también un uso insuficiente de los informes de evaluación. A partir de sus hallazgos, Aristimuño y Kaztman (2005) formulan un modelo conceptual que tiene como variable a explicar la apertura al cambio del maestro (la disposición para usar los informes de evaluación), y como variables explicativas (relacionadas entre sí) la antigüedad del docente, el tipo de escuela en que trabaja y la densidad de las redes sociales de la escuela.

Montoya, Perusia y Vera (2005) analizaron el uso de la información sobre las evaluaciones del rendimiento escolar realizadas por el Ministerio de Educación Nacional en Argentina. Los autores encontraron un conocimiento diferencial de los informes según la asignatura a la que se refieren (mayor en el caso de español y matemáticas y menor en el de otras asignaturas evaluadas). La oportunidad con que llegó la información a los establecimientos también fue diferente según las pautas de entrega seguidas en tres provincias distintas. Se advierte que no existe un plan que guíe a las instituciones respecto al uso de la información proporcionada. Los autores concluyen que para que los informes de evaluación sean insumos relevantes para el cambio debe trabajarse con los docentes en la interpretación de la información reportada y en el análisis de alternativas pertinentes a los resultados.

Aunque estos trabajos abordaron desde diferentes ángulos un mismo objeto de estudio y emplearon para ellos diversas metodologías, Cueto (2005) señala que, en conjunto, estas investigaciones dejan ver que en América Latina prevalece una preocupación por la información que se debe recabar y la manera de hacerlo, mientras que hay poco desarrollo a propósito de cómo emplearla para contribuir a la mejora de la educación. De acuerdo con el autor, pese a que estos sistemas fungen como insumos de información que es vital para los sistemas educativos, hay poco uso de estos datos por parte de las audiencias a las que se dirigen. Para Martinic (2010) esto deriva, en parte, de “las insuficiencias en la comunicación hacia los actores y [la falta de] interpretaciones que les ayuden a comprender los problemas que verifican los datos reunidos” (p. 38).

Finalmente, cabe hacer alusión al informe “Las evaluaciones educativas que América Latina necesita” (Ravela, 2010), documento que derivó del proyecto “La evaluación de aprendizajes en las aulas de primaria en América Latina. Enfoques y prácticas”, llevado a

cabo por el Instituto de Evaluación Educativa de la Universidad Católica del Uruguay, con apoyo del Grupo de Trabajo sobre Estándares y Evaluación (GTEE) de PREAL. Este trabajo fue un estudio exploratorio y descriptivo elaborado en ocho países de América Latina (Argentina, Colombia, Costa Rica, El Salvador, México, Perú y Uruguay) con maestros de sexto grado de primaria, cuyo énfasis estuvo en los enfoques de la evaluación en las áreas de matemáticas, lectura y escritura.

El trabajo tuvo entre sus objetivos estudiar cómo las evaluaciones a gran escala ayudan a los profesores a enriquecer sus propias prácticas de evaluación. Gran parte de los docentes indicaron que los resultados de este tipo de evaluaciones efectivamente contribuyen a la mejora de esas prácticas. A partir de la información reunida, se generó una tipología de usos, misma que incluye la toma de conciencia acerca de los contenidos a abordar en el aula; la apertura a nuevos enfoques conceptuales y prácticos; el uso de instrumentos para implementarlos en el aula; la creación de nuevos instrumentos de evaluación; la preparación de los alumnos para las mismas pruebas en aplicaciones posteriores; y la creación de procesos de seguimiento para los alumnos, de manera general o individual (Ravela, 2010).

De igual manera, el estudio identificó factores que pueden afectar el aprovechamiento que los profesores hacen de la información derivada de evaluaciones a gran escala, tales como su comprensión de los datos, la formación de los propios docentes, el exceso de carga de trabajo, las fallas técnicas de las pruebas, las opiniones de los profesores acerca de las evaluaciones, entre otros (Ravela, 2010).

En el contexto mexicano son pocas las investigaciones que se han dado a la tarea de indagar el uso de resultados de las evaluaciones a gran escala, si bien, este campo de estudio ha cobrado mayor importancia en las últimas décadas. Uno de los trabajos que abordó este tema es previo a la implementación de ENLACE y su objetivo fue “[...] captar, registrar y sistematizar significados que profesores, directores, supervisores y funcionarios otorgan a la evaluación externa a partir de su experiencia” (Figueroa, 2003, p. 354). Los resultados de esta investigación apuntan a que existían “pocas y aisladas prácticas de difusión y uso de los resultados de los programas” (p. 372).

Al cabo de algunos años del surgimiento de ENLACE, empezaron a realizarse estudios sobre diferentes aspectos y dimensiones de la evaluación y de la prueba en particular (Martínez Rizo y otros, 2015; Santiago et al., 2014;). A propósito del uso de resultados, Santiago y otros (2014) señalan. en los hallazgos de su investigación, que en sus orígenes ENLACE fue planteada como una evaluación de carácter formativo, cuyos usos eran de bajo impacto; sin embargo, debido a los usos reales que se hicieron de esta información, entre ellos la asignación de incentivos para docentes y la difusión pública de los resultados de las escuelas, se convirtió en una evaluación de alto impacto. La preparación para la prueba, o la reducción del currículo, prácticas antes mencionadas, se hicieron patentes en el caso de ENLACE. Por otra parte, esta prueba también propició una toma de conciencia por parte de los profesores acerca de la necesidad de fortalecer la preparación de los alumnos en determinadas áreas y sirvió como motivación para los estudiantes y los docentes.

Martínez Rizo (2015), con la colaboración de un grupo internacional de expertos coordinó un estudio de validación de las pruebas ENLACE para educación básica (ENLACE-B) y media superior (ENLACE-MS), así como de los Exámenes para la Calidad y el Logro

Educativo (Excale). Una de las dimensiones incluidas en el estudio es la del uso de los resultados; entre los hallazgos principales cabe mencionar los siguientes: la carencia de un modelo que pueda alinear los aspectos técnicos de las pruebas a sus usos previstos, así como la diversidad de usos y su nivel de especificidad; la falta de acceso a la información de manera oportuna y efectiva; la carencia de organización y coordinación para fomentar la capacidad de los diversos actores educativos en la interpretación de los datos; el poco desarrollo de mecanismos eficientes y robustos que permitan a los investigadores acceder y hacer uso de estas bases de datos; y la falta de seguimiento a los usos y consecuencias derivadas de éstos.

3. Ejes para el estudio de la difusión y uso de resultados de evaluaciones a gran escala

Los ejes que se presentan en esta parte del artículo fueron identificados en la revisión de literatura realizada en el marco del proyecto de investigación al que se hace referencia en la Introducción. Estos ejes fueron de utilidad para sistematizar la información reunida y se considera que resultan sugerentes para indagaciones posteriores sobre el tema, advirtiendo que quizá sería necesario revisar si los ejes propuestos son suficientes para abarcar de forma más exhaustiva el tema y si son pertinentes dada la evaluación que se esté considerando en particular. Aunque el proyecto en que fueron utilizados como referentes teóricos centró su atención en los usos de los resultados por profesores y directores de escuelas primarias, se incluyen dos ejes previos a esta acción, la difusión y el análisis e interpretación de los datos, debido a que, en la literatura revisada, aparecen como procesos de claro impacto sobre el uso que efectivamente se puede hacer de los resultados; igualmente se plantea un eje relacionado con los factores que influyen en su uso por parte de los actores mencionados.

3.1. Difusión de la información de evaluaciones a gran escala

La evaluación cobra sentido en la medida en que contribuye a que la educación mejore (INEE, 2006b) y para lograr esto, entre otras acciones, es necesario que los resultados se difundan de manera oportuna y amplia a todos los actores educativos, quienes, a su vez, deben poseer una adecuada cultura de la evaluación, entendiendo este término como:

[...] Un conocimiento de las características, alcances y limitaciones de los distintos tipos de evaluación, con base en lo cual tanto las autoridades como los maestros, los padres de familia y la sociedad en general puedan interpretar los resultados de las evaluaciones correctamente, y los puedan utilizar para sustentar decisiones adecuadas, en los respectivos ámbitos de cada uno. (Martínez Rizo y Santos, 2009, p. 10)

La difusión de resultados depende del enfoque y las concepciones asumidas por los sistemas de evaluación nacional y los fines para los que fueron creados. Ravela (2012) indica que hay dos enfoques para llevar a cabo esta tarea: data-céntrico y agere-céntrico. En el primero de éstos el punto central son los datos de la evaluación, vistos desde la lógica de la estadística y la psicometría; se procesan los datos obtenidos y se generan informes. Este enfoque enfatiza la difusión de los resultados, pero no su uso, dado que éste se encuentra en manos de otros actores. Desde esta postura, se asume que la información estadística compleja podrá ser comprendida por todos los actores educativos y que sabrán como emplearla, además de presuponer que la información será de utilidad para la toma

de decisiones y que los usuarios tienen tanto los incentivos como la motivación para usar la información.

A diferencia del anterior, el segundo enfoque se orienta a los actores, de tal manera que se busca, en primera instancia, familiarizarse con sus modos de actuar respecto a la educación, sus percepciones e intereses, así como con el tipo de decisiones que pueden tomar. Con base en esto, se decide cómo se hará el tratamiento de los datos, buscando difundir la información que les pueda ser útil a los actores para tomar decisiones. El enfoque agerecéntrico es menos usual en América Latina, dado que los sistemas de evaluación nacionales han puesto mayor énfasis en la difusión de los resultados que en su uso (Ravela, 2012).

A partir de un estudio acerca del uso de la información de los sistemas de evaluación de educación primaria en México y Uruguay, Fernández y Midaglia (2005) también señalan que hay diversas estrategias para comunicar los resultados. La más frecuente se denomina *trickle down* o de chorreo y consiste fundamentalmente en la distribución de los reportes de resultados de uno a otro actor, comenzando por las autoridades educativas o instancias encargadas de iniciar esta tarea, hasta llegar a las escuelas⁴.

La segunda estrategia es la *resignificación*, que parte del supuesto “[...] de la comunicación transparente y someter en distintas instancias a reflexión el sentido de lo comunicado” (Fernández y Midaglia, 2005, p. 129). Dicha estrategia tiene dos subtipos, en el primero, el informe de los resultados se acompaña por un documento de trabajo dirigido a profesores y directores, el cual puede referirse a los objetivos de la evaluación o a la interpretación de los datos. El segundo subtipo tiene como punto de partida un cuestionamiento, ya sea implícito o explícito, y en este sentido, “[...] las organizaciones intermedias simplemente adaptarían y aprehenderían los informes bajo las premisas estructuralmente disponibles, que son típicamente burocráticas” (p. 130).

El último tipo de difusión de resultados es llamado estrategia de *reestructura*, la cual se define de la siguiente manera:

[...] una política de reforma estructural de la supervisión escolar que pone en cuestión las relaciones interorganizacionales entre gerencia, dirección de supervisión y escuelas. Entre otros impactos deseados, contribuiría a jerarquizar el uso de las evaluaciones como un insumo de las prácticas de asesoramiento académico, de diagnóstico y monitoreo de las escuelas y sus maestros. (Fernández y Midaglia, 2005, p. 130)

3.2. Análisis e interpretación de la información

Uno de los aspectos a abordar cuando se indagan los usos que diversos actores hacen de los resultados de pruebas aplicadas en evaluaciones a gran escala es el análisis que realizan a partir de la información que reciben, y lo que este proceso les permite ver, es decir, las interpretaciones de los resultados, mismas que influirán previsiblemente, en la toma de decisiones.

Cawelti y Protheroe (2000, en Protheroe, 2009), Fernández y Midaglia (2005), Means y otros (2009), Ravela (2010) y Ravela y otros (2007) señalan, como parte de los usos de los

⁴ La investigación realizada por Fernández y Midaglia tomó como referente en el caso de México, la evaluación Aprovechamiento Escolar, que formaba parte del Programa Carrera Magisterial.

resultados de las pruebas a gran escala, la generación de diagnósticos del aprendizaje de los estudiantes y/o de los grupos, así como del funcionamiento de las escuelas o regiones.

En el estudio llevado a cabo por el Council of the Great City Schools (2012), se señala que dar sentido a los datos provenientes de las evaluaciones intermedias y finales aplicadas a los alumnos desemboca idealmente, en una mejora del conocimiento y comprensión de las necesidades y retos de los alumnos, de manera individual y como grupo; de las fortalezas y áreas de oportunidad de los profesores y las escuelas, así como de las estrategias y recursos para dar atención a las dificultades de los alumnos, los profesores y las escuelas (Council of the Great City Schools, 2012).

A propósito de esta misma cuestión, Griffin (2009) indica que las evaluaciones cuyos fines son formativos pueden contribuir a identificar el nivel de desarrollo de los estudiantes y, con base en esta información, es posible implementar intervenciones que contribuyan al progreso de los alumnos. Para llevar a cabo esto se requiere, como condiciones indispensables, que el propio instrumento de evaluación posea suficiente capacidad para generar un diagnóstico del aprendizaje de los alumnos y que los profesores y, de manera general, los actores educativos que reciben, analizan y usan los resultados, tengan habilidades para la interpretación y el uso de este tipo de datos (Griffin, 2009).

Matters (2009) indica que las pruebas estandarizadas pueden ofrecer información a propósito no sólo de lo que los alumnos aprendieron, sino también de aquello que no aprendieron. Ante esto, es necesario que los docentes puedan identificar las condiciones en las cuales los alumnos llegaron o no a determinado nivel de logro. Love (2004) también indica que en el proceso de uso de datos se deben identificar los problemas de los estudiantes y verificar sus causas.

Al igual que Love (2004), el Institute of Education Sciences (2009) señala que la interpretación de la información derivada de este tipo de evaluaciones consiste en identificar las fortalezas y debilidades de los alumnos, ya sea de forma individual o del grupo, y, subsecuentemente, desarrollar hipótesis acerca de cómo mejorar el aprendizaje de los alumnos.

Tanto el Institute of Education Sciences (2009) como Love (2004) y Ravela (2001) señalan la importancia de triangular información acerca del aprendizaje de los alumnos, con el fin de comprobar los datos que ofrecen distintas fuentes y así llegar a conclusiones fundamentadas acerca de las necesidades educativas de los alumnos.

3.3. Uso de los resultados

Si bien es cierto que la difusión de los resultados de las evaluaciones a gran escala o estandarizadas y su análisis e interpretación son condiciones indispensables para asegurar su adecuado aprovechamiento, el uso que se haga de esta información es lo que da sentido a la propia tarea de evaluar.

Además de la distinción ya mencionada entre usos de alto y bajo riesgo o impacto, a partir de la literatura consultada también se pueden distinguir tres niveles en los que se usan los resultados de las evaluaciones a gran escala y diversos matices u orientaciones en cada caso.

3.3.1. Nivel de las prácticas docentes

En el primer nivel se engloban todos aquellos usos y acciones que se espera que los profesores lleven a cabo desde sus prácticas docentes, ya sean de enseñanza o evaluación. Dentro de este nivel, se encuentran diversos planteamientos que se pueden agrupar en cinco categorías:

1. Planeación de clase
2. Modificación y mejora e intervenciones
3. Prácticas de evaluación
4. Apoyo y seguimiento de estudiantes
5. Pruebas estandarizadas

El diseño y/o ajuste de la planeación de clase incluye diversos aspectos que van desde la modificación de los planes de clase (Council of The Great City Schools, 2012), reordenar los planes de estudios, su secuencia y el tiempo dedicado a cada tema (Institute of Education Sciences, 2009, Hoover, 2009, en Council of The Great City Schools, 2012), hasta cuidar la alineación del currículo o de las expectativas de aprendizaje (Institute of Education Sciences, 2009, Protheroe, 2009).

La segunda categoría está sustentada en las contribuciones de autores como Coburn y Talbert (2006), Hoover (2009, en Council of the Great City Schools, 2012) y Protheroe (2009), que se refieren a la modificación, mejora o diferenciación de las prácticas y enfoques de enseñanza de los profesores. Las intervenciones consisten en el diseño e implementación de diversas estrategias y actividades. En algunos casos se habla de manera general de la realización de “intervenciones” (Griffin, 2009; Institute of Education Sciences, 2009), mientras que, en otros, se señala de manera específica, que un uso que se da a la información de este tipo de evaluaciones consiste en volver a enseñar determinados temas o conceptos (Hoover, 2009).

La tercera categoría alude a la modificación de las prácticas evaluativas de los docentes, ya se trate de retomar los instrumentos de las evaluaciones a gran escala para aplicarlas en el aula o que éstas sirvan como modelo para generar nuevas evaluaciones (Ravela, 2010). Autores como Cawelti y Protheroe (2000, en Protheroe, 2009) señalan que, a partir de los resultados de este tipo de pruebas, se pueden aplicar evaluaciones intermedias; mientras que Hoover (2009, en Council of The Great City Schools, 2012) menciona que una práctica común es aplicar nuevamente la evaluación para abordar determinados temas.

La cuarta categoría comprende el apoyo y seguimiento de los estudiantes con bajos niveles de logro. Diversos autores señalan que, a partir del diagnóstico generado con base en la información de las pruebas, se puede dar seguimiento o tutoría a los alumnos que presentan dificultades, así como ubicarlos en grupos especiales para brindarles apoyo adicional (Cawelti y Protheroe, 2000, en Protheroe, 2009; Council of The Great City Schools, 2012; Insitute of Education Sciences, 2009; Hoover, 2009, en *Council of The Great City Schools*, 2012; Ravela, 2010).

Finalmente, Ravela (2010) también incluye, como uno de los usos de los resultados de evaluaciones a gran escala, la preparación para este tipo de pruebas, aunque el mismo autor indica que:

Este uso puede ser legítimo si la finalidad es familiarizar a los alumnos con un cierto tipo de preguntas y de evaluaciones, pero tiene el problema de que muchas veces termina transformando en el objetivo principal de la enseñanza lo que pretende ser solo un indicado
(p. 14)

3.3.2. Nivel de la gestión escolar

En el segundo nivel se encuentran todos aquellos usos de los resultados de evaluaciones a gran escala que son atribuidos al colectivo escolar. Autores como Ravela (2001) señalan, de manera general, que los resultados de evaluaciones a gran escala pueden contribuir al mejoramiento de las escuelas, aunque no se especifica a través de qué medios se lleva a cabo esto. Algunas otras contribuciones llevan a identificar las siguientes categorías en este nivel:

1. Mejora escolar, planes escolares e intervenciones a nivel escuela
2. Revisión del currículo
3. Comunicación con los padres

En la primera categoría se encuentran las contribuciones de Ravela (2001) y de Santiago y otros (2014), autores que señalan que los resultados deben usarse para propiciar el fortalecimiento o mejora de las escuelas. También comprende el uso de resultados para elaborar o ajustar planes escolares o de mejora, así como para plantear nuevas metas (Institute of Education Sciences, 2009; Department of Education, 2010, en Council of The Great City Schools, 2012 y Council of The Great City Schools, 2012). Finalmente, en esta categoría se pueden ubicar usos relacionados con la implementación de estrategias y actividades a nivel escolar (Henke, 2005, en Council of The Great City Schools, 2012 y Council of The Great City Schools, 2012).

La segunda categoría alude a la revisión y ajuste del currículo a partir de la información de pruebas a gran escala (Department of Education, 2010, en Council of The Great City Schools, 2012 y Council of The Great City Schools, 2012), lo cual tiene sentido en sistemas educativos como el norteamericano donde el currículum se encuentra descentralizado.

Finalmente, la tercera categoría consiste en informar a los padres de los alumnos acerca del progreso académico de sus hijos (Means et al., 2009). Este punto forma parte de los propósitos y usos sugeridos de ENLACE, que de manera particular proponía la realización de actividades, tales como talleres para propiciar la participación de los padres en el proceso de aprendizaje de los alumnos.

3.3.3. Nivel de la supervisión

En el nivel de la supervisión se incluyen aquellos usos de la información que hacen las instancias externas a las escuelas; en este caso también se advierten diferentes posibilidades:

1. Desarrollo profesional docente
2. Apoyo a las instituciones escolares
3. Diseño de políticas públicas

La primera categoría engloba, de manera general, los usos encaminados a fortalecer el desarrollo de los docentes a través de diversas estrategias. En la investigación de Council of The Great City Schools (2012), se habla de las acciones que emprenden los distritos con

base en los resultados de evaluaciones intermedias y finales del aprendizaje de los alumnos y entre ellas se encuentran las estrategias encaminadas al desarrollo profesional de los docentes. En este punto coinciden Ravela (2001), SEP (2013) y Santiago y otros (2014), al igual que Kerr, Marsh, Darilek y Barney (2006, en Young y Kim, 2010).

La segunda categoría se refiere al apoyo a las instituciones escolares, que es uno de los usos a los que hace alusión la investigación de Council of The Great City Schools (2012), particularmente el acompañamiento a las escuelas en la generación del plan de mejora escolar. Por último, la tercera categoría comprende la generación de programas, políticas públicas y educativas por parte de las autoridades escolares o del gobierno (Coburn y Talbert, 2006; Dussel et al., 2005; Ravela, 2001; Santiago et al., 2014).

3.4. Efectos del uso de los resultados

Los usos que se den a los resultados de las evaluaciones a gran escala producen un efecto o impacto en diferentes niveles de los sistemas educativos. Con respecto a este tema, los modelos de uso de resultados presentados por Love (2004) y el Institute of Education Sciences (2009) señalan que el efecto esperado del proceso de analizar, interpretar y usar los datos de este tipo de evaluaciones consiste en una mejora tanto en la enseñanza como en el aprendizaje de los alumnos.

Para Ravela (2010), el uso de los resultados de pruebas estandarizadas puede propiciar que los profesores tomen conciencia de los conocimientos y capacidades que los alumnos deben adquirir. Por otra parte, contribuye a que los docentes se abran ante otros horizontes conceptuales en el ámbito de la didáctica y la evaluación; les brinda acceso a instrumentos de evaluación que les pueden servir de apoyo en el trabajo en el aula o para crear nuevos instrumentos; y les permite reflexionar a propósito de sus prácticas de evaluación en el aula, poniendo mayor énfasis en los procesos cognitivos de sus alumnos.

Si bien lo deseable es que el uso de este tipo de evaluaciones derive en una mejora general de los sistemas educativos, hay usos que desembocan en efectos no deseados. Matters (2009) indica que una acción común es que las ciudades y países se definan en función de los resultados de evaluaciones como PISA.

Por su parte, Dussel y otros (2005) encontraron, a partir de una investigación realizada en las provincias de Córdoba, Chaco y la Ciudad Autónoma de Buenos Aires en Argentina, que los docentes consideran que el impacto de los sistemas de evaluación ha sido poco visible o incluso nulo, debido a situaciones tales como “[...] la falta de acceso, el desconocimiento, la escasa necesidad de estadísticas para la labor docente, la falta de presupuesto para modificar situaciones ya identificadas como problemáticas, o porque cada docente está en sus temas específicos” (p. 377).

Matters (2009) afirma que, en este siglo, la información derivada de las evaluaciones es extremadamente poderosa, siempre y cuando cumpla con dos condiciones. La primera de ellas consiste en que la prueba, por sí misma, sea de calidad y lo suficientemente fuerte para fundamentar los usos que se hacen a partir de la información que arroja; mientras que la segunda condición señala que los usuarios de esta información deben poseer suficiente pericia para reconstruir adecuadamente la realidad que subyace a los resultados.

3.5. Factores asociados al aprovechamiento de la información de evaluaciones a gran escala

La mayor parte de la literatura revisada en esta investigación refiere una amplia cantidad de factores que impactan en el uso de los resultados de evaluaciones a gran escala, mismos que se pueden agrupar en dos categorías los factores de orden individual, relativos a los actores educativos, y los que se ubican en el nivel de la institución escolar.

3.5.1. Factores individuales

Los factores individuales se definen como las percepciones y actitudes de los actores educativos, mismas que pueden favorecer u obstaculizar el trabajo con los datos. Cabe señalar que esto no alude ni involucra elementos relacionados con el contexto sociodemográfico de los docentes, variables que diversas investigaciones (Fernández y Midaglia, 2005) han demostrado no intervenir en aspectos como la lectura de los resultados por parte de los profesores.

Ravela (2010) afirma que es pertinente considerar las percepciones de los docentes, dado que “[...] la evaluación externa puede ser percibida de distintas maneras: como una amenaza, como una oportunidad de aprendizaje o como un requisito administrativo más [...]” (p. 18). Esto depende de la cultura de evaluación de cada institución y de la estrategia de comunicación de los resultados, pero también se atribuye a las ideas previas de los docentes. Asimismo, factores como la formación inicial de los profesores, el clima institucional y profesional de las escuelas, la complejidad de la labor docente “[...] pueden favorecer u obstaculizar la comprensión y el uso de evaluaciones [...]” deben ser tenidas en cuenta a la hora de promover el uso de los resultados de las evaluaciones” (Ravela, 2010, p. 18). Datnow, Park y Wohlstetter (2007, en Young y Kim, 2010) enfatizan que la capacidad de los docentes para usar los datos de las evaluaciones está “inextricablemente ligado a su conocimiento instruccional. El análisis de los datos puede indicar las áreas en las que los profesores deben hacer mayores esfuerzos, pero no les dicen cómo hacerlo” (p. 10).

En la investigación de Council of The Great City Schools (2012) también se indica que las percepciones negativas a propósito de la calidad de los datos de la evaluación pueden llevar a que los actores educativos se muestren renuentes ante las estrategias implementadas a partir de esta información, lo que trae consigo una disminución del aprovechamiento de los resultados de las evaluaciones. De acuerdo con Young y Kim (2010), “aunque los distritos pueden usar muchas estrategias para estimular y apoyar a las escuelas en el uso de los datos, éstos deben tener legitimidad ante los profesores” (p. 18) para que los usen.

La investigación de Council of The Great City Schools (2012) reporta que otro tipo de factores que intervienen en el uso de las evaluaciones intermedias y finales son las limitaciones de las capacidades del personal, mismas que en otros estudios también se ha encontrado que obstaculizan el uso de los resultados (Departamento de Educación de Estados Unidos, 2010a; Herencia, 2007; Herman et al., 2008; Ingram et al., 2004; Lachat y Smith, 2005; Sharkey y Murnane, 2006; Patrimonio y Bailey, 2006; Wayman et al., 2007). En el mismo sentido, esto es, las debilidades en la formación de los profesores para la evaluación, apuntan las conclusiones de Impara y otros (1993) y Stiggings (1991) (citados en Young y Kim, 2010).

Por su parte, Dussel y otros (2005) señalan que hay varios factores relevantes para poder interpretar el comportamiento de los docentes ante la información empírica, uno de éstos se relaciona con el escepticismo de los profesores ante la utilidad de la información. De igual manera, estos autores reportan diversas opiniones ante las estadísticas que reportan las evaluaciones, algunas de ellas positivas y otras que apuntan a una conceptualización de este tipo de información como datos en los que hay sesgo y que, por tanto, no reflejan la complejidad de la realidad evaluada. Finalmente, también hay opiniones que los autores denominan “ausentes” y consiste en una postura que concibe las estadísticas como una actividad carente de sentido.

3.5.2. Factores institucionales

Los factores institucionales se definen como aquellos aspectos de la dinámica escolar o de la logística de la evaluación que intervienen en el aprovechamiento de los resultados de evaluaciones a gran escala. La literatura consultada en este estudio apunta principalmente a tres factores que puede determinar la difusión, análisis, interpretación y uso de los resultados: 1) la entrega oportuna de la información, 2) el trabajo colegiado, y 3) la existencia de liderazgos.

A propósito del primer punto, la investigación de Council of the Great City Schools (2012) señala que diversos autores (Clune y White, 2008; Lachar y Smith, 2005; Wayman et al., 2007) han encontrado que algunas de las barreras para el uso de los resultados de las evaluaciones intermedias es la falta de puntualidad y accesibilidad en la entrega de los resultados, así como la carencia de precisión de éstos. Rojas y Esquivel (1988, en Cruz, Montoya y Vera, 2005) coinciden en este punto y afirman que, de no hacer una entrega oportuna de los resultados, éste se torna “[...] un ejercicio mecánico y costoso, que genera una gran cantidad de información que en últimas no es utilizada, hecho que incide en la seriedad y prestigio que se le asigne al proceso” (p. 88).

Sobre el trabajo colegiado, son varios los investigadores (Council of the Great City Schools, 2012; Griffin, 2009; Ravela, 2010) que señalan la importancia de que exista una dinámica de trabajo colectivo en torno al análisis, interpretación y uso de los resultados de evaluación. Con respecto a este tema, Griffin (2009) indica que los equipos profesionales de enseñanza permiten a los profesores exponer y compartir sus interpretaciones acerca de los datos, así como proponer estrategias e intervenciones y, sobre todo, hacer un trabajo reflexivo y crítico en torno a estos aspectos.

El modelo de uso de evaluaciones expuesto por el Institute of Education *Sciences* (2009) también enfatiza la importancia de establecer equipos de análisis de datos que dirijan su uso y definan conceptos cruciales para ello, tales como las nociones de enseñanza y aprendizaje. Por otra parte, en la investigación de Council of the Great City Schools (2012) se indica que una de las barreras para el uso de los datos, de acuerdo con el Department of Education (2010), es la falta de tiempo para participar en la examinación y reflexión de los resultados.

Finalmente, la literatura también apunta a la existencia de liderazgos dentro de las instituciones escolares como un elemento que incide en el proceso del uso de datos de las evaluaciones a gran escala. Fernández y Midaglia (2005) establecieron como una de las hipótesis de su investigación que: “los estilos de gestión del director de la escuela tienen efectos diferenciales sobre la lectura y uso de los informes por parte de los maestros” (p.

157). Los resultados de este estudio indicaron que para muchos de los actores educativos encuestados el éxito de la lectura de los reportes de resultados “[...] no dependería de características personales, sino principalmente del director y de su capacidad de motivación, fomento de la lectura, supervisión de la labor docente, su rol en la formación de equipos estables y académicamente sólidos, entre otros aspectos” (p. 156).

A propósito de este tema, el Institute of Education Sciences (2009) señala que es indispensable proporcionar liderazgo en el proceso de uso de datos. Asimismo, Christman y otros (2009) y Kerr y otros (2006), ambos citados en Council of the Great City Schools (2012) afirman que uno de los factores determinantes en el uso de resultados, en este caso de evaluaciones intermedias, es el liderazgo a nivel de enseñanza y a nivel escolar.

Con base en diversos autores, Young y Kim (2010) establecen que:

El liderazgo del director es crucial para establecer expectativas entre el colectivo escolar sobre el uso de los datos como insumos para la toma de decisiones (Wayman y Stringfield, 2006a, Young, 2006) y para crear ambientes de apoyo en los que los profesores puedan compartir los éxitos y los fracasos asociados con la evaluación. (p. 19)

Además de expectativas, un adecuado liderazgo precisa las actividades a desarrollar, modela los usos apropiados de la información, proporciona una teoría de la acción para el uso de los datos y despliega los recursos necesarios para hacer todo esto (Young y Kim, 2010).

4. Conclusiones

Durante los últimos años se han dado notables avances en materia de difusión de resultados de las evaluaciones a gran escala. En 2006, el INEE afirmaba que instancias como la Subsecretaría de Educación Básica (SEB), la Dirección General de Evaluación de Políticas (DGEP) y el propio INEE habían contribuido a ampliar el acceso a la información a través de diversos materiales (publicaciones, análisis de resultados, manuales técnicos, entre otros). A pesar de esto, el INEE (2006b) llamaba la atención sobre la inexistencia de mecanismos sistemáticos de difusión entre los actores educativos: alumnos y sus familias, maestros y directores:

Lógicamente, esto impide que los resultados sean utilizados en forma sistemática para mejorar los procesos educativos básicos, es decir las actividades de aprendizaje. A esto se agrega que, cuando los resultados son difundidos hacia la sociedad a través de los medios de comunicación, suelen generarse debates estériles derivados de interpretaciones incorrectas o sesgadas. (p. 2)

En el marco de la evaluación ENLACE, la difusión de los resultados se constituyó como elemento fundamental, lo que contribuyó a su gran notoriedad entre la sociedad en general y en el sector educativo, a pesar de que muchas de las dificultades en el aprovechamiento de esta información se originaban precisamente desde esta etapa. A propósito de los mecanismos de difusión de resultados de ENLACE y Excale, Martínez Rizo (2015) señala que “[...] se encontraron carencias técnicas e inconsistencias entre la información que se ofrece y los objetivos que se siguen, lo que explica que estos esfuerzos no estén teniendo el impacto esperado” (p. 89).

El autor indica también que la expectativa radica en que la difusión de resultados propiciará el aumento del aprendizaje de los estudiantes; en el mismo sentido apunta

Martinic (2010) cuando apunta que “[...] a puesta en marcha de los sistemas de evaluación tiene, como uno de sus supuestos, el hecho de que una vez conocidos los resultados, éstos serían ampliamente empleados por los usuarios en distintos ámbitos decisionales” (p. 37). Aunque la difusión de los resultados es una parte imprescindible de cualquier tipo de evaluación, no se reduce a la tarea de garantizar el acceso a la información, “[...] sino que deben ofrecerse las herramientas adecuadas para su interpretación, así como evitar usos inapropiados” (INEE, 2006, p. 1).

La evaluación a gran escala que actualmente opera en el sistema educativo mexicano, PLANEA, ha buscado, a través de su diseño e implementación, superar las limitaciones de las evaluaciones que la precedieron y mantener sus fortalezas (INEE, 20015). En los documentos que brindan el marco general de PLANEA, el INEE reconoce que la tarea de evaluar no se reduce a la mera producción de información por relevante que ésta sea, sino que se requiere conformar y consolidar procesos de difusión que garanticen el acceso oportuno de todos los actores educativos a la información, así como orientaciones que guíen su interpretación y análisis. De igual manera, se advierte que:

Actualmente, uno de los grandes retos de las evaluaciones consiste en diseñarlas en función de modelos lógicos donde se prevean los usos que los distintos actores pueden dar a los resultados para producir los efectos deseados en el sistema educativo, al mismo tiempo que anticipar y desalentar usos inapropiados o no pertinentes. (INEE, 2015, p. 2)

Ciertamente, los elementos involucrados en la difusión y uso de resultados de evaluaciones a gran escala, como muestra la revisión de literatura expuesta en este artículo, se constituyen como procesos cuya interacción es compleja, lo que hace indispensable la existencia de mecanismos que garanticen su congruencia, en aras de garantizar un adecuado aprovechamiento de la información que proporciona este tipo de evaluaciones.

En este sentido, también se torna imprescindible que las evaluaciones a gran escala sean objeto de estudio de investigaciones sistemáticas y rigurosas que examinen tanto su calidad técnica, implementación y procesamiento de datos, como las prácticas en torno a su análisis, uso y las consecuencias que derivan de ellos.

La propuesta que se hace en este artículo aporta los ejes que fueron de utilidad para orientar el acopio de información, así como la sistematización y análisis de ésta, en el marco de un estudio que tomó a la prueba ENLACE como su referente. Los ejes se refieren a procesos que abarcan un ciclo amplio una vez que los resultados están disponibles para hacerse públicos: difusión, análisis e interpretación, uso de los resultados propiamente, efecto del uso de éstos, y factores asociados al aprovechamiento de la información. Algunos se pudieron describir mejor que otros; pero en conjunto, sugieren vetas de indagación potencialmente relevantes para determinar si los esfuerzos y recursos que hoy en día se invierten en la evaluación de los resultados de aprendizaje están permitiendo alcanzar las expectativas que los alientan.

Financiación

Proyecto financiado por el Consejo Nacional de Ciencia y Tecnología (CONACYT) y la Secretaría de Educación Pública (SEP).

Referencias

- Aristimuño, A. y Kaztman, R. (2005). La evaluación de aprendizajes y su impacto en las políticas y las escuelas públicas de Uruguay. S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina*, (pp. 173-243). Santiago de Chile: PREAL.
- Backhoff, E. y Contreras, S. (2014). "Corrupción de la medida" e inflación de los resultados de ENLACE. *Revista Mexicana de Investigación Educativa*, 19(63), 1267-1283.
- Cervera, L., Martín, G. y Sánchez, C. (2008). Estudio georreferencial de la Evaluación Nacional de Logro Académico en Centros Escolares (ENLACE) en el municipio de Juárez, Chihuahua: Análisis espacial. *Revista Electrónica de Investigación Educativa*, 10(1), 1-23.
- Council of the Great City Schools. (2012). *Charting success: Data Use and Student Achievement in Urban Schools*. Recuperado de http://www.cgcs.org/cms/lib/DC00001581/Centricity/Domain/87/Charting_Success.pdf
- Cueto, S. (Ed.). (2005). *Uso e impacto de la información educativa en América Latina*. Santiago de Chile: PREAL.
- Dussel, I., Legarralde, M., Marcalain, G., Montes, N. y Pinkaz, D. (2005). Los usos de la información empírica en el sistema educativo. Un estudio de caso en tres jurisdicciones de la Argentina. En S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina*, (pp. 345-394). Santiago de Chile: PREAL.
- Fernández, T. y Midaglia, C. (2005). El uso de los informes generados por los sistemas de evaluación de aprendizaje en la educación primaria. Los casos de México y Uruguay. En S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina* (pp. 11-171). Santiago de Chile: PREAL.
- Figuroa, A. (2003). La perspectiva de maestros, directores y supervisores escolares sobre la evaluación externa. En M. Zorrilla (Coord.), *La evaluación de la educación básica en México 1990-2000. Una mirada a contraluz* (pp. 351-406). Ciudad de México: Universidad Autónoma de Aguascalientes.
- Griffin, P. (2009). Teacher's Use of Assessment Data. En C. Wyatt-Smith y J. J. Cumming (Eds.), *Educational Assessment in the 21st Century* (pp. 183-208). Amsterdam. Springer. https://doi.org/10.1007/978-1-4020-9964-9_10
- INEE. (2006). La difusión de los resultados de las evaluaciones. *Colección Folletos: Los temas de la evaluación*. Ciudad de México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/C/610/P1C610.pdf>
- INEE. (2015a). *Planea: una generación de pruebas. ¿Qué es PLANEA?*. Ciudad de México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P2/A/303/P2A303.pdf>
- Institute of Education Sciences. (2009). *Using Student Achievement Data to Support Instructional Decision Making*. Recuperado de http://ies.ed.gov/ncee/wwc/pdf/practice_guides/dddm_pg_092909.pdf
- Love, N. (2004). Taking Data to New Depths. *National Staff Development Council*, 25(4), 22-26.
- Martínez Rizo, F. (2001). Evaluación educativa y pruebas estandarizadas. Elementos para enriquecer el debate. *Revista de la Educación Superior*, 30(120), 71-85.
- Martínez Rizo, F. (Coord.). (2015). *Las pruebas ENLACE y EXCALE. Un estudio de validación*. Ciudad de México: INEE. Recuperado de <http://publicaciones.inee.edu.mx/buscadorPub/P1/C/148/P1C148.pdf>

- Martinic, S. (2010). La evaluación y las reformas educativas en América Latina. *RIEE. Revista Iberoamericana de Evaluación Educativa*, 3(3), 30-43.
- Martínez Rizo, F. y Santos, A. (2009). Evaluación e investigación educativa. En A. Alba y R. Glazman (Coords.), *¿Qué dice la investigación educativa?* (pp. 265-304). Ciudad de México: COMIE.
- Matters, G. (2009). A Problematic Leap in the Use of Test Data: From Performance to Inference. En C. Wyatt-Smith, C. y J. J. Cumming (Eds.), *Educational Assessment in the 21st Century* (pp. 209-225). Amsterdam: Springer.
- Means, B., Padilla, C., DeBarger, A. y Bakia, M. (2009). *Implementing Data-Informed Decision Making in Schools-Teacher Access, Supports and Use*. Recuperado de <http://files.eric.ed.gov/fulltext/ED504191.pdf>
- Montoya, S., Cruz, J. y Vera, A. (2005). Evaluación de la calidad educativa: de los sistemas centrales al aula. Estudio del impacto de las políticas de devolución de información personalizada a las escuelas en tres provincias de Argentina. En S. Cueto (Ed.), *Uso e impacto de la información educativa en América Latina* (pp. 295- 343). Santiago de Chile: PREAL.
- Moreno, T. (2011). La cultura de la evaluación y la mejora de la escuela. *Perfiles Educativos*, XXXIII(31), 116-130.
- Protheroe, N. (2009). Improving teaching and learning with data-based decisions: Asking the right questions and acting on the answers. *ERS Spectrum*, 19(3), 4-9.
- Ravela, P., Wolfe, R., Valverde, G. y Esquivel, J. (2000). *Los próximos pasos: ¿Hacia dónde y cómo avanzar en la evaluación de aprendizajes en América Latina?*. Santiago de Chile: PREAL-GRADE-GTEE.
- Ravela, P. (2001). *¿Cómo presentan sus resultados los sistemas nacionales de evaluación educativa en América Latina?* Santiago de Chile: PREAL.
- Ravela, P. (2010). ¿Qué pueden aportar las evaluaciones estandarizadas a la evaluación en el aula? *Documentos de Trabajo PREAL 47*. Santiago de Chile: PREAL.
- Ravela, P. (2012). La experiencia del INEE en la difusión de resultados para docentes. En INEE, *INEE. Una década de Evaluación 2002-2012*. Ciudad de México: INEE.
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez, F., Aylwin, M. y Wolff, L. (2007). *Las evaluaciones educativas que América Latina necesita. Documentos de trabajo PREAL 40*. Santiago de Chile: PREAL.
- Ruiz, G. y Prado, P. (2017, noviembre). El uso en el nivel de educación primaria de los resultados de la evaluación ENLACE. Un estudio intensivo en Aguascalientes. Comunicación presentada en el *XIV Congreso Nacional de Investigación Educativa*, Consejo Mexicano de Investigación Educativa, San Luis Potosí. Recuperado de <http://comie.org.mx/congreso/memoriaelectronica/v14/doc/1037.pdf>
- Santiago, P., McGregor, I., Nusche, D., Ravela, P. y Toledo, D. (2014). *Revisiones de la OCDE sobre la Evaluación en Educación. México 2012*. Ciudad de México: INEE.
- SEP. (2013). *Evaluación Nacional del Logro Académico en Centros Escolares. Educación Básica: manual técnico 2013*. Ciudad de México: SEP.
- Young, V. y Kim, D.H. (2010). Using assessments for instructional improvement: A literature review. *Education Policy Analysis Archives*, 18(19), 1-39. <https://doi.org/10.14507/epaa.v18n19.2010>

Breve CV de las autoras

Guadalupe Ruiz Cuéllar

Doctora en Educación egresada del programa interinstitucional de la Universidad Autónoma de Aguascalientes. Es Investigadora Nivel 1 del Sistema Nacional de Investigadores. Es docente de pregrado y posgrado en las áreas de Metodología de la Investigación, Gestión de la Educación y Evaluación Educativa. Su trabajo de investigación se desarrolla fundamentalmente en esta última área. Actualmente es Directora General de Investigación y Posgrado de la UAA. Email: gruiz@correo.uaa.mx

Paloma Prado Robledo

Licenciada en Letras Españolas por la Universidad de Guanajuato y Maestra en Investigación Educativa por la Universidad Autónoma de Aguascalientes. Se ha dedicado al diseño y elaboración de materiales didácticos y evaluaciones de aprendizaje en contextos de educación para adultos. Fungió como formadora de profesores en diversos temas pedagógicos, fue coordinadora académica y docente en educación media superior. Actualmente, colabora en proyectos de investigación en la línea de práctica docentes y evaluación educativa. Email: pradopaloma@hotmail.com