

¿Evaluación o Inculpación?

F. Javier Merchán Iglesias¹

Resumen

La mejora de la educación a gran escala parece una meta inalcanzable y una promesa difícil de cumplir. Supuestamente con ese objetivo, las políticas educativas hoy dominantes en el panorama mundial, promueven un dispositivo en el que la evaluación y autoevaluación constituyen la pieza clave. Inspirándose en la perspectiva de las escuelas eficaces y en los modelos de gestión de calidad, su virtualidad no consiste tanto en que realmente mejoren la educación de forma significativa, sino en que exime al estado de esa responsabilidad y se la endosa a centros y profesores. Así, a la vista de los pobres resultados que el mecanismo de la evaluación ha conseguido en años, más que de evaluación cabría hablar de inculpación.

Palabras clave

Evaluación, autoevaluación, mejora de la educación, política educativa, profesores, eficacia escolar, gestión de calidad

Abstract

The improvement of education on a large scale seems an unattainable goal and a difficult promise to fulfill. With this objective, education policies today dominant on the world scene, promote a device in which the evaluation and self-evaluation are the key piece. Inspired by the prospect of effective schools and quality management models, its virtuality does not consist of both actually improve education significantly, but it exempts the State from this responsibility and endorses it centers and teachers. Thus, in view of the poor results the evaluation mechanism has achieved in years, instead of evaluation could be talking of indictment.

Key words

Evaluation, self-evaluation, improve of education, education policies, teachers, school effectiveness, quality management.

Recibido: 11-02-2015
Aceptado: 17-03-2015

¹ Universidad de Sevilla, jmerchan@us.es.

La configuración de la evaluación como parte del dispositivo para la mejora de la educación

Cual bálsamo de Fierabrás, una forma de evaluación se ha naturalizado como mecanismo de mejora de la educación a escala mundial. Aplicándose a centros y profesores, gobiernos de uno y otro signo, delegan en este instrumento estrella de las políticas educativas neoliberales la mejora a gran escala de los aprendizajes y de los resultados académicos de los alumnos, o, como gusta decir a sus defensores, la mejora de la calidad de la educación. Transcurridos ya varios años desde su puesta en marcha, a finales de la década de los ochenta del pasado siglo, en este artículo se trata de hacer una valoración crítica de sus fundamentos, así como de sus resultados en el ámbito de la enseñanza primaria y secundaria.

El mecanismo en cuestión forma parte de un dispositivo más amplio que se ha ido configurando mediante la confluencia de varias líneas de pensamiento y acción, no siempre provenientes del campo de la educación. Sus orígenes pueden rastrearse en el conocido como *movimiento de escuelas eficaces*, movimiento que de alguna forma tuvo su punto de partida en el archiconocido *Informe Coleman*. Como se sabe, este informe, encargado por el Congreso norteamericano, llegó a la conclusión –resumida en la frase *la escuela no importa*– de que el factor determinante en el logro académico de los estudiantes era el contexto sociocultural. Frente a esto, se advirtió que, no obstante, ese factor no operaba en todos los casos, acuñándose el término *escuelas eficaces* para referirse a aquellos centros escolares que obtenían resultados mucho mejores de los que cabía esperar de su contexto sociocultural. El *movimiento de escuelas eficaces*, partiendo de este dato, se ocupó entonces en indagar cuáles eran los factores que explicaban la eficacia de esas escuelas. A renglón seguido, se desarrolló otra línea complementaria que se conoció como *mejora de la eficacia escolar*. Ahora, se trataba de trabajar, no ya en la indagación sobre los factores de éxito, sino en la implementación de fórmulas que permitieran conseguir que las escuelas ineficaces consiguieran buenos resultados. En el camino parecía haberse olvidado la relevancia del contexto sociocultural, no porque se ignorara su incidencia, sino porque se consideraba que esa variable podía controlarse.

A decir verdad, la búsqueda de factores de éxito y de estrategias de mejora basadas en la extensión de lo que hoy denominaríamos *buenas prácticas*, tuvo poco recorrido, pues, no se alcanzaba un consenso significativo acerca de cuáles eran esos factores, ni funcionaba la idea de aplicar el modelo de los centros que alcanzaban buenos resultados a los que no los obtenían. No obstante, permanecía viva la convicción de que era posible obtener buenos resultados al margen del condicionante sociocultural y de que tendría que existir una fórmula que lo hiciera posible.

Por su parte, en el campo de la industria, se habían venido desarrollado iniciativas de mejora de la productividad que acabarían denominándose Sistemas de Gestión de Calidad. Se trata de métodos que, resumidamente, se basan en procesos de evaluación y, sobre todo, de autoevaluación. El más conocido y duradero es el llamado ciclo de Deming, formulado en la década de los cincuenta. El sistema, conocido también por las siglas PDCA (Plan-Do-Check-Act), consistía en un círculo continuo en el que se parte de un diagnóstico, se fijan objetivos y planes, se ponen en marcha y se evalúan para fijar nuevos objetivos y nuevas estrategias. Siguiendo estos pasos, la European Foundation for Quality Management (EFQM), corporación fundada en 1988 y formada por grandes empresas multinacionales, editó en 1992 un modelo de referencia para la evaluación de los procesos, modelo que pronto sería trasladado al campo de la educación.

La adopción de este modelo de gestión para abordar el objetivo de la mejora de la educación a gran escala se apoyó también en algunas tesis de las Teorías de las Organizaciones aplicadas al funcionamiento de los centros escolares. Más concretamente, las teorías estructurales destacaron la importancia de la eva-

luación, y de los resultados, proponiendo el cambio mediante la autoevaluación (Perrow, 1991). Se parte del supuesto de que los centros escolares operan racionalmente, se proponen (o se les proponen) metas para maximizar el rendimiento y de que algunas formas de gestión por objetivos pueden ayudar a su consecución “Este es, por ejemplo el marco teórico subyacente al modelo de mejora TQM [*Total Quality Management*], basado en el ciclo de Deming” (Muñoz-Repiso et. al, 2001: 77). Así, se considera a la escuela como una organización inteligente que, mediante la autoevaluación, es capaz de identificar sus errores, aprender de la experiencia, planificar y desarrollar acciones hacia el logro de los fines que les *son propuestos*.

Finalmente, un nuevo ingrediente vino a sumarse al dispositivo de mejora de la educación basada en la evaluación: La formulación en los años 70 en USA de lo que se conoce como *New public management*. Efectivamente, a partir de la crisis de los años 70, se empieza a cuestionar el funcionamiento de la administración del estado, propugnando la eficiencia y rentabilidad de los servicios que ofrece mediante la puesta en marcha de nuevos modelos de gestión. Los principios de estos nuevos modelos serían: a) Los productos y servicios públicos son mensurables, lo que requiere el uso de técnicas y métodos científicos; b) los actores deben rendir cuenta ante los gestores y los ciudadanos; y c) la administración pública debe regularse en función de los resultados y no exclusivamente en función de los recursos (Moons, 2009).

Con estos materiales, en la década de los noventa se fue perfilando un dispositivo sistemático aplicable al campo de la educación. Frente a las políticas de mejora basadas en el cambio curricular que se desarrollaron en torno a las reformas de la escuela comprensiva, se apostaba ahora por la tesis de que la mejora de la educación es básicamente un problema de gestión. Ello requería disponer de un artefacto que, inspirándose en los procedimientos antes mencionado, fuera viable en el ámbito específico de la educación. No se empezaba de cero: en 1968, en el Yale Child Study Center, el psiquiatra James P. Comer, en colaboración con las New Haven Public Schools puso en marcha programas de intervención escolar mediante el *School Development Plan*, subrayando con su experiencia la pertinencia y buenos resultados de la planificación como estrategia de mejora de la educación. Sin embargo, la evaluación y autoevaluación no constituían todavía el elemento central de la planificación escolar. Será en 1988, con la *Education Reform Act* inglesa, cuando empieza a desplegarse el dispositivo que convierte a la evaluación y autoevaluación en la pieza clave de la política educativa. El proceso se completaría mediante la adaptación del ciclo de Deming y el modelo EFQM al campo de la educación.

Para el caso de España la fórmula se empieza a adoptar en 1996 –siendo ministra de Educación Esperanza Aguirre–, con la puesta en marcha del Plan General para la Gestión de Calidad en Educación y los Planes de Autoevaluación y Mejora (López, 1997)². En USA, aunque ya existía cierta tradición sobre la evaluación de centros y profesores, es en el año 2001 con la ley conocida por sus siglas NCLB (*No Children Left Behind*) cuando el dispositivo entra plenamente en funcionamiento al ensamblar sus tres principales piezas: evaluación basada en resultados, rendición de cuentas e incentivos.

Efectivamente, la evaluación, como estrategia central de las políticas neoliberales de mejora de la educación, es una pieza de un dispositivo más amplio en el que, junto a los elementos anteriores, hay que incluir también el gerencialismo. Es lo que en su conjunto se ha venido a llamar la gestión empresarial de la escuela, pues, al fin y al cabo, se trata de una traslación de la cultura de la empresa al campo de la educación (Merchán, 2010 y 2011). Inspirándose en el citado modelo EFQM y concretándose en el *School Development Plan* (En el caso de España, Plan de Autoevaluación y Mejora), el dispositivo parte del principio de que la mejora de la educación es, en primer lugar, responsabilidad de la escuela. La escuela,

2 Un estudio sobre la introducción en España de este dispositivo puede verse en Merchán, 2012.

se dice, es la que mejor conoce las posibilidades de mejora y las estrategias pertinentes para introducir cambios que logren avances en los resultados de los alumnos, para ello el procedimiento apropiado es la autoevaluación, complementada con la evaluación externa (Department for Education, 2004); es decir, el motor de la mejora de la educación (o, más bien, habría que decir, mejora de los resultados) es una reflexión eficaz y sistemática de los actores del campo que permitirá diseñar las actuaciones pertinentes en el ámbito de los centros escolares. Se le confiere a la escuela la responsabilidad –y, supuestamente, la capacidad– del cambio, lo que en los discursos dominantes se refiere como autonomía.

La fórmula en cuestión se desarrolla a partir de la cuantificación del éxito mediante indicadores que son variables según las características del ámbito en el que se aplican, pero que, generalmente se centran en los resultados que obtienen los alumnos en las evaluaciones curriculares y en evaluaciones externas. Los indicadores se establecen como objetivos fijados desde fuera, de manera que la autoevaluación se refiere fundamentalmente al grado de consecución. En el proceso se planifican y ponen en marcha estrategias, se evalúan los logros alcanzados, considerando las razones de éxito y fracaso, volviendo a fijarse nuevos objetivos y nuevas estrategias en un ciclo infinito de *mejora*. El mecanismo se complementa con políticas de incentivos que premian el logro y penalizan el fracaso, pues, como se ha dicho, se atribuye a los actores del campo la responsabilidad de los avances o retrocesos³. Bolívar ha descrito acertadamente el meollo de la nueva estrategia de cambio:

En el inveterado problema de cómo hacer gobernable la enseñanza, es decir cómo desde la política educativa se puede influir positivamente en lo que se hace en el aula superando la débil conexión, el movimiento de los estándares ha encontrado una teoría de la acción: definir los resultados esperados, medirlos y usar los datos resultantes para influir en la enseñanza, ya sea directamente mediante la intervención de la autoridad o, como es más común, indirectamente por los incentivos o sanciones derivadas de los resultados alcanzados, por la propia administración educativa o cediéndola a los clientes” (Bolívar, 2006: 38).

Por su parte Lingard, Ladwing y Luke (2001) consideran que estamos en realidad ante un artefacto del taylorismo aplicado a las escuelas, con la novedad quizás de la “autonomía” y la “autovigilancia”. Un artefacto que funciona como un mando a distancia, pues regula desde fuera lo que se pretende que ocurra dentro.

La evaluación como análisis de resultados

La promoción de la “cultura de la evaluación” y de la práctica de la autoevaluación como estrategias de mejora, se apoya en sobretendidos interesadamente publicitados que adolecen de notable debilidad teórica y carecen de apoyos empíricos suficientemente consistentes. Desde luego, la estrategia en cuestión reformula profundamente el concepto de mejora de la educación y el sentido mismo de la formación de los jóvenes escolarizados. La centralidad y relevancia que el sistema concede a los indicadores de éxito o fracaso, hace que, en la práctica, la mejora de la educación se asimile a mejora de resultados; y, puesto que, supuestamente, ese es el propósito de la evaluación –mejorar los resultados–, ocurre que los indicadores se convierten en objetivos. Pero, cómo acertadamente señala Stobart, refiriéndose a la conocida como Ley de Goodhart, sucede también que cuando una medida se convierte en objetivo, deja de ser una buena medida, pues, añade, citando a McIntre: “Por regla general, la medida de un sistema lo perturba. Cuanto

³ En algunos países la política de incentivos afecta al salario del profesorado, a su situación laboral y a la gestión del centro escolar. En España generalmente se ha limitado a algunas subvenciones de los centros. Sólo en el llamado Plan de Calidad, desarrollado en Andalucía durante algunos años, los resultados en el rendimiento de los alumnos han supuesto incentivos económicos al profesorado. En la LOMCE también se establece que los centros que no alcancen un nivel mínimo de resultados, se someterán a planes de actuación diseñados por la administración educativa.

más precisa es la medida y más corta su escala temporal, mayor es la energía de perturbación” (Stobart, 2010: 135). Es decir, por una parte, cuando el indicador se convierte en objetivo cuyo grado de consecución es objeto de medición, deja de ser un buen indicador. Por otra, la medición del fenómeno, en nuestro caso la enseñanza y el aprendizaje, lo altera significativamente.

Efectivamente, respecto al valor de los resultados como indicadores de la calidad de la educación, Yong Zaho (2014) cuestiona, por ejemplo, que los resultados de PISA indiquen realmente el grado de formación de los alumnos. Observa el autor que en los últimos quince años suelen ser los países asiáticos los que obtienen mejores resultados en las citadas pruebas, pero lo único que con ello se demuestra es eso, que obtienen buenos resultados en los exámenes, lo que atribuye al hecho de que sus sistemas educativos se centran precisamente en preparar a los estudiantes para aprobar exámenes. Sin embargo, dice Zaho, es difícil encontrar en ellos una producción científica relevante a nivel mundial. El crecimiento económico de estos países no se debe a mejoras en el nivel de formación de los jóvenes, sino a las inversiones de capital extranjero.

Por otra parte, al asimilarse el objetivo de la mejora de la educación con la mejora de los resultados y puesto que, mediante la política de incentivos, el grado de consecución tiene consecuencias profesionales y económicas sobre los docentes y sobre los centros escolares, sucede que, efectivamente, la práctica docente tiende a centrarse en la preparación de los alumnos para las pruebas. Así, los centros escolares se convierten en autoescuelas, la enseñanza en adiestramiento y el aprendizaje en una suerte de entrenamiento para el examen. Por lo demás, el conocimiento que de esta forma se pone en juego, ha de ser igualmente reconfigurado con el fin de adaptarse a los requerimientos del hecho examinatorio, descartando aquellos cuya posesión no es susceptible de medición y procurándoles un formato comprimido que haga viable su asimilación. Ciertamente estas consecuencias no siempre se producen de manera inexorable. De hecho las pruebas PISA tratan de confeccionarse de manera que se conjuren esos perversos efectos del examen sobre la formación que reciben los alumnos. Sin embargo, en las condiciones actuales de la escolarización, la política de resultados extiende por el conjunto del sistema una dinámica de medición del logro que es difícilmente viable sin el recurso al examen convencional (Merchán, 2009b).

También, se ha subrayado la incidencia negativa en el comportamiento de los profesores y centros sobre los alumnos con más dificultades. Así, se tiende a prestar más atención a los que tienen posibilidades de obtener resultados satisfactorios, abandonando a su suerte a los que seguramente tendrán malos resultados. En muchos casos los directores sugieren a los “malos alumnos” que abandonen la escuela, al menos el día de la prueba, para no bajar los resultados (Wrigley, 2007; Mons, 2009). Además, cuando la remuneración y el estatus profesional están ligados a los resultados, los profesores tienden a abandonar los centros con alumnos más difíciles para trabajar en otros con más posibilidades de éxito; es por esto por lo que Ladd y Walsh (2002:16) han constatado que la política de incentivos ligados a resultados perjudica a los centros que tienen más necesidades de mejora, de manera que, como ya señalaba Wrigley, uno de sus efectos perversos es precisamente que aumenta las diferencias entre alumnos de grupos sociales diferentes.

Barrenechea (2010), en fin, advierte de la generalización de prácticas ilícitas con el fin de mejorar los resultados. El hecho de que tengan efectos en el prestigio y financiación de los centros, en los incentivos que perciben los profesores y la influencia de su repercusión mediática en la imagen de la administración educativa hace que, en algunos casos, se “ayude a los alumnos” a responder correctamente, se baje el nivel de las pruebas o se maquillen estadísticamente los datos, lo que, desde luego, repercute en el papel formativo de la escuela y en realidad no mejoran el nivel educativo. Ciertamente se tiende a corregir estas

“desviaciones”, pero afirma Barrenechea, mientras más importancia se le da a los resultados, más se agudiza el ingenio para buscar fórmulas que reduzcan sus efectos.

Evaluar para conocer y mejorar

En el corazón de las políticas de mejora de la educación basadas en la evaluación, se encuentra la idea de que el mecanismo produce conocimiento acerca de las dificultades para alcanzar los resultados propuestos y esperados y, además, conocimiento para diseñar estrategias de mejora. Es decir, se parte del supuesto de que, analizando los resultados, el sistema emite señales significativas y relevantes sobre los déficits y potencialidades en el logro académico y de que los profesores y gobernantes son capaces de captarlas e interpretarlas. De hecho, PISA –paradigma del procedimiento evaluador– sugiere el camino a seguir, pues no consiste meramente en realizar pruebas a los alumnos y medir los resultados, sino que, además, recaba información sobre una serie de variables que hipotéticamente influyen en ellos, estableciendo correlaciones y relaciones de causalidad potencialmente capaces de explicar (y de sugerir líneas de actuación para la mejora). A este respecto, en el prólogo del informe español de PISA 2006, se expone el principal objetivo que anima a estos estudios impulsados por la OCDE: “Contribuir al mejor conocimiento de los aspectos fundamentales del funcionamiento del sistema educativo, analizar qué explican los resultados obtenidos y, sobre todo, facilitar la adopción de las políticas y las acciones educativas que permitan mejorar el sistema educativo español” (MEC, 2007: 12), es decir, se trata de establecer una relación productiva entre el conocimiento y la actuación para la mejora de la educación.

En el mismo sentido, imitando la fórmula PISA, algunas administraciones educativas establecieron sus propias pruebas con semejante intención, es decir, la de generar conocimientos que expliquen los resultados y permitan diseñar estrategias de mejora. En España es el caso, por ejemplo de la Comunidad Autónoma andaluza. Aplicando el modelo PISA, la Consejería de Educación de la Junta de Andalucía regulaba la realización de la denominada evaluación de diagnóstico en el ámbito territorial de su competencia:

[A partir de los resultados] *Los centros, a la vista de sus circunstancias y características, tras un proceso de análisis y reflexión riguroso establecerán las propuestas de mejora del rendimiento del alumnado que consideren necesarias. Se trata de una información que, utilizada en su justa medida, pueda ser útil al centro, al profesorado, al alumnado y a las familias para coordinar esfuerzos en la mejora del rendimiento escolar.* (Orden de 28 de Junio 2006).

De manera que, siguiendo las pautas de la evaluación, el análisis de los resultados proporciona información valiosa para que –en este caso los profesores y los centros– gestionen el cambio, la mejora de los rendimientos y... de la práctica educativa.

Ahora bien, ¿qué información proporciona realmente la evaluación basada en el análisis de los resultados que obtienen los alumnos en pruebas y evaluaciones curriculares? Como se ha dicho anteriormente, la indagación sobre los factores del éxito escolar desarrollada por el movimiento de las *escuelas eficaces*, no alcanzó a producir conocimientos decisivos, ni ha conseguido resolver los interrogantes sobre las escuelas eficaces al margen del factor sociocultural. De hecho hoy nos encontramos con enormes listas que han ido aumentando unos estudios tras otros, llegando hasta la cifra de 719 que encuentran Scheerens y Bosker en una revisión sobre el tema (Muñoz-Repiso y Murillo, 2001: 7). Además, los estudios sobre factores de éxito en distintos países llegan a conclusiones contradictorias (Creemers, 2001: 61), confunden correlación con causalidad (Wrigley, 2007: 28), sobreestiman las posibilidades de la estadística para hacerse cargo de los procesos de la vida social y operan con dudoso rigor al definir *ex-ante* las variables de la efectividad que serán observadas (Carrasco, 2007: 293). También se ha destacado que estos estudios se

centran en factores sobre los que se puede actuar (Scheerens, 1998), mientras que relegan aquellos sobre los que no se puede intervenir desde los centros escolares, operando como si realmente no intervinieran; sin embargo, son numerosos los trabajos de investigación que ponen de manifiesto el hecho de que entre el 85 % y el 90 % de la varianza entre las escuelas se explica por factores sobre los que no tienen control (Lauder, Jamieson, y Wikely, 2001: 75; Carabaña, 2009).

El caso es que la dinámica de la evaluación basada en el análisis de resultados no parece que haya hecho aportaciones relevantes sobre el tema. Son muchos los estudios, incluso los elaborados por las instituciones y administraciones patrocinadoras de los programas de evaluación, que llegan a conclusiones bastante similares acerca del éxito en el rendimiento académico de los alumnos: el factor socioeconómico y cultural es el más decisivo⁴. Concretamente, en un trabajo sobre los datos de PISA 2006, Carabaña (2009) afirma que ni el sistema educativo, ni la gestión pública o privada de los centros escolares, ni otras de las variables explicativas consideradas inciden de manera significativa en los resultados:

Los estudios PISA permiten afirmar que las diferencias entre los países no se deben a las características de sus escuelas, ni a nivel de sistema ni a nivel de centro. (...) El único factor con poder explicativo importante es la composición sociocultural de las poblaciones, un rasgo de los individuos que le viene de fuera a la escuela. (Carabaña, 2009: 92-93).

En consecuencia, puede decirse que si el objetivo de las evaluaciones y pruebas de diagnóstico, o de otro tipo, es el de facilitar información y producir conocimiento sobre la mejora de la educación, hay que constatar que realmente nos dicen muy poco (poco más de lo que ya sabemos) sobre las diferencias de rendimiento y nada sobre lo que habría que hacer para mejorar los resultados (Carabaña, 2009). Y si esto puede afirmarse respecto del mecanismo a escala de los países, con más razón puede hacerse cuando se aplica a la escala de los centros, pues si los análisis de los resultados de los alumnos utilizando gran aparato estadístico revelan lo que ya sabemos, difícilmente pueden los centros y profesores –generalmente con menos recursos analíticos– producir los *nuevos* conocimientos que se les demanda para aplicarlos a la mejora de los resultados y de la práctica docente.

El mecanismo de la evaluación no sólo propone explicar los resultados, sino producir también estrategias que conduzcan a la mejora del rendimiento de los alumnos. Esta idea también se ha planteado anteriormente al referirme al movimiento de *mejora de la eficacia escolar* como uno de los ingredientes que han intervenido en la configuración del conjunto del dispositivo. Se ha considerado que, conociendo los factores de éxito, es posible establecer pautas de intervención que permitan a las escuelas *ineficaces* conseguir mejoras significativas. Es decir, si conocemos las dificultades, conocemos las soluciones.

Amparándose en el supuesto de que existe una racionalidad científico-técnica capaz de conseguir escuelas eficaces mediante la aplicación de un instrumental apropiado, la perspectiva de la eficiencia ha pretendido emular lo que ocurre en el campo de las ciencias positivas y de la técnica, sin caer en la cuenta de que las variables que allí intervienen son de muy distinta naturaleza a las que actúan en los procesos de la educación y en las dinámicas que gobiernan el sistema escolar.

4 Son muchos los estudios que abundan en esta idea. Con respecto, por ejemplo, a los datos de PISA 2006, así se afirma, entre otros, en el Informe español PISA 2006 (MEC, 2007), en el estudio de López Martín y otros (2009), y en el de Julio Carabaña (2009). En el mismo sentido se pronuncia el estudio de García, Pinto y Robles (2009). Por lo demás, los informes de la administración andaluza sobre las pruebas de diagnóstico (Evaluaciones de diagnóstico, 2006-2007; 2007-2008 (avance) y 2008-2009), llegan también a la misma conclusión.

La idea de que, una vez que conocemos los factores de éxito, la mejora consiste en reproducirlos en las escuelas ineficaces, adolece de una grave debilidad conceptual. David Reynolds y Louise Stoll (2001) afirman que los estudios sobre eficacia son muy deficientes en lo que respecta al análisis de procesos, de manera que, en todo caso, muestran una instantánea de una escuela en un momento determinado, más que una “película”; pero “la mejora escolar necesita saber de qué forma las escuelas se han convertido en eficaces (o ineficaces) para poder copiar (o erradicar) los procesos” (Ibídem: 112). Por otra parte, al descuidar la importancia de factores externos a las escuelas y los docentes, no puede establecerse “de manera concluyente qué variables son causas de la eficiencia escolar y cuáles son efecto” (Ibídem: 113), así que “la determinación de las variables exactas en las que la mejora de la escuela necesita dirigirse para influir en los resultados es, hoy día, claramente imposible” (Ibídem: 114).

En fin, Lauker, Jamieson y Wikely (2001) sostienen que, aun admitiendo la discutida posibilidad de conocer las características de las escuelas de éxito y considerarlas factores de eficacia, de ahí no se desprende una estrategia de cambio. En todo caso, esas evidencias pueden servir para escuelas que *ya son de éxito*, pero carece de sentido pensar que las escuelas ineficaces se caracterizan precisamente por lo contrario, de manera que, sabiendo cómo son las escuelas eficaces, no podemos saber cómo actuar con las que son ineficaces (Carrasco, 2008).

Si nos referimos a PISA, o a las pruebas similares que a nivel local realizan algunas administraciones citadas anteriormente, resulta evidente que de las correlaciones estadísticas entre los resultados de los alumnos y un grupo de variables seleccionadas a priori no se deriva ningún tipo de actuación. Como se ha dicho anteriormente, aún en el supuesto ya discutido de que se expliquen los resultados, los gobernantes en su escala y los centros y profesores en la suya, no pueden inferir de ellos actuaciones que mejoren decisivamente la educación (Carabaña, 2009), de manera que no es arriesgado afirmar que la evaluación como análisis de resultados difícilmente puede ser un recurso⁵ potente para la mejora de la educación⁵.

Christian Maroy y Catherine Mangez (2008) consideran que la relación entre el conocimiento educativo –a partir de la información que proporciona PISA y otras pruebas similares– sobre los problemas de la educación y el desarrollo de políticas educativas forma parte de un movimiento más amplio que pretende hacer ver las decisiones políticas como decisiones técnicas. A este respecto, cuestionan la afirmación, tantas veces reiterada en las declaraciones oficiales, de que las evaluaciones contribuyan a la mejora de la educación y de la política educativa, preguntándose si los datos que proporcionan las pruebas sirven verdaderamente para racionalizar la acción pública o lo que realmente se produce es una politización del conocimiento, es decir, un uso político de esos datos. En un estudio sobre las reacciones de distintas instituciones vertidas en la prensa a propósito de las pruebas, los autores constatan cómo, efectivamente, en el caso de Bélgica, los mismos resultados PISA se utilizan para justificar, argumentar y apoyar tres discursos distintos: el de la ministra de educación, el del movimiento socialista APED y el de la Red de Escuelas católicas.

En el nivel de los centros escolares, las administraciones educativas se sirven del mecanismo de la evaluación y del uso de incentivos o penalizaciones para impeler a los profesores a que formulen y pongan en marcha propuestas de mejora, atribuyéndoles, como se decía anteriormente, la responsabilidad del cambio. Ya se ha dicho que la mera reflexión a partir de los resultados es difícil que pueda producir conocimientos potentes y nuevos para su explicación y que, difícilmente también, puede generar estrategias de mejoras significativas. Se apela entonces al conocimiento profesional y a una disposición positiva de los docentes. A este respecto, se da por supuesto que existe un conocimiento técnico capaz de diseñar y po-

⁵ Por nuestra parte, no cuestionamos la utilidad de la evaluación, sino el papel central que se le atribuye en la mejora de la educación.

ner en marcha estrategias que permitan avanzar en el logro del rendimiento de los alumnos. Un supuesto que nos devuelve al viejo paradigma tecnocrático sobre la educación, que ignora la complejidad de la vida social y sus cualidades específicas y distintas a las del mundo técnico. O a una suerte de pedagogismo, o de idealismo pedagógico, que reduce los problemas de aprendizaje a problemas relativos a los métodos de enseñanza y, por tanto, a la competencia profesional. Sin embargo, como afirma Perrenoud (2000:3), los problemas no son meramente técnicos, ni el estado del conocimiento sobre la enseñanza nos permite una acción eficaz y clara capaz de responder a tamaño requerimiento

Por otra parte, el mecanismo opera como si se actuara en un sistema cerrado, pues, aunque formalmente se asume la incidencia de factores extraescolares, en la práctica de la autoevaluación se actúa como si los centros escolares fueran islas, es decir, entidades descontextualizadas con capacidad de afrontar y resolver todas sus dificultades (Wrigley, 2007). Ciertamente, se argumenta que se trata de que las escuelas y los profesores hagan todo lo que puedan en el campo de sus posibilidades, pero, deliberadamente, se ignora el carácter sistémico de los problemas educativos y la incidencia de factores sobre los que la escuela no tiene capacidad de intervención. Es decir, perfectamente puede ocurrir que los docentes formulen propuestas de mejora cuya aplicación resulta sencillamente inviable. Limitar el campo de las propuestas a aquellas que los docentes pueden poner en marcha –como la administración sugiere–, de ninguna manera garantiza la mejora de los resultados, sino que, más bien, preconiza su fracaso. De esta forma, la política educativa, mediante el mecanismo de la autoevaluación como instrumento de mejora, no sólo se exime de cualquier responsabilidad ante el previsible fracaso, sino que se la endosa a los profesores.

Así mismo, dada las peculiares características de la práctica de la enseñanza y del trabajo de los docentes, es difícil que a las propuestas de mejora de los profesores se les pueda atribuir la exclusiva responsabilidad del cambio. Efectivamente, aunque la práctica de la enseñanza está sujeta (y constreñida) a una serie de reglas y obligaciones (de contenido, espacio, tiempo, número de alumnos, etc.), existe un amplio campo de incertidumbre en la relación cara a cara con los alumnos. Por otra parte, la enseñanza es una relación entre dos actores sociales, insertos a su vez en un contexto más complejo de relaciones; esta relación se mueve entre la cooperación y el conflicto, entre intereses comunes y divergentes. Así, el trabajo docente, aunque sea excelente, no es garante del éxito de los alumnos, pues, entre otras razones, requiere su cooperación. En todo caso, puesto que ni todo el éxito ni todo el fracaso puede atribuirse al trabajo docente, es prácticamente imposible determinar su porcentaje de responsabilidad. En fin, la peculiaridad de la profesión docente, y del marco organizativo en el que se realiza, la hace escasamente autónoma y, por lo tanto con escasas posibilidades para responder a fuertes requerimientos externos (Durán, 2008)⁶. Diríamos que no puede hablarse de una relación causal única entre enseñanza y aprendizaje, sino que este estaría determinado por un conjunto de factores además de la enseñanza (Bogotch, Miron y Biesta, 2007) y, en consecuencia, no cabe esperar cambios significativos sólo de las propuestas de mejora que formulan los docentes.

Los resultados de la política de mejora de la educación basada en la evaluación

Abstrayéndonos ahora del debate acerca de si los resultados indican realmente una mejora de la educación, parece pertinente indagar sobre los logros que el dispositivo ha conseguido en su propio campo, es decir en el rendimiento de los alumnos. A este respecto, Hopkins y Lagerweij (2001:89) se preguntaban, refiriéndose al caso de Inglaterra y Gales, sobre la relación entre la dinámica de la autoevaluación y los resultados de los alumnos, afirmando que “La respuesta más sencilla es que estas estrategias pueden crear

⁶ Sobre la singularidad de la práctica de la enseñanza en el aula puede verse Merchán, 2009a.

y, de hecho, crean las condiciones para mejorar los resultados de los alumnos, pero, por sí solas, tienen poco impacto directo en el progreso de los alumnos”.

Paradójicamente no son muchos los estudios e investigaciones sobre los logros de la evaluación como instrumento de mejora. Mons advierte que algunos trabajos sobre las políticas de rendición de cuentas y de evaluaciones externas, son poco fiables, ya que adolecen de tales fallos metodológicos –y triquiñuelas–, que en buena parte los invalidan; lo que sí queda claro es que el mecanismo de la evaluación y el uso de incentivos a profesores y centros escolares, no ofrece unos resultados suficientemente consistentes (Mons, 2009: 112-117). Más concretamente, la citada autora se refiere a uno de los programas pioneros del mecanismo de la evaluación, la aplicación de la *Texas Assessment of Academic Skills*, programa al que se le llegó a bautizar como “milagro de Texas”. Los primeros datos revelaban una mejora espectacular de los resultados. Sin embargo, este primer entusiasmo fue enfriado por investigaciones posteriores de más amplio recorrido en las que se utilizaron también los resultados de las pruebas federales. Estos ponían de manifiesto que las progresiones disminuían fuertemente y no eran significativas según las disciplinas. La experiencia tejana, afirma Mons, interroga sobre la consistencia de los efectos positivos de la evaluación estandarizada y la rendición de cuentas, su pervivencia en el tiempo, los efectos perversos que puede generar, así como sobre los escollos metodológicos de algunas investigaciones que han analizado sus consecuencias. Refiriéndose al mismo caso, y abundando en la tesis de Mons, Stobart (2010: 140), afirma irónicamente que “la cuestión es si esto [el milagro Texas] fue un milagro o un espejismo”⁷.

Por su parte, una investigación desarrollada en el País Vasco sobre la percepción de los profesores de la incidencia de la implantación de sistemas de acreditación de calidad basados en la evaluación (normas ISO)⁸ en los centros docentes (Etxague y otros, 2009), concluye que entre las variables en las que los docentes no han percibido mejora, se encuentra, además de otras, los resultados académico-educativos. En el mismo sentido concluye un estudio de Cantón (2004) sobre la aplicación de los Planes de Autoevaluación y Mejora desarrollados en la Comunidad de Castilla-León. La autora afirma que los profesores consideraban que la dinámica de la autoevaluación les exigía un trabajo excesivo y, sobre todo, que generaba una gran actividad burocrática, y que, sin embargo, las mejoras didácticas eran escasas, es decir, apenas incidían en la práctica de la enseñanza. Igualmente, Bolívar (1994) sostiene que hay evidencias como para concluir que el control por el centro del proceso de cambio mediante el recurso de la autonomía, autoevaluación y evaluación, no necesariamente afecta ni mejora la educación; en realidad genera cambios formales y, en algunos casos, cambios en las estructuras de decisión que no producen una mejora escolar significativa.

Recordando ahora el vaticinio de la administración educativa andaluza en 2006 –cuando afirmaba que la realización de las Pruebas de Diagnóstico y la introducción de la dinámica de la autoevaluación produciría mejoras en el rendimiento de los alumnos y de la práctica educativa–, nueve años después, puede afirmarse que no existe ninguna evidencia, antes al contrario, de que se hayan producido mejoras significativas en ninguno de esos objetivos. Incluso el magro informe de la propia administración sobre los resultados de la aplicación del Plan de Calidad (paradigma en España del dispositivo de mejora basado en la evaluación, autoevaluación e incentivos), pone de manifiesto que los resultados son bastante pobres⁹.

7 Este programa de evaluación de los estudiantes del estado de Texas se puso en marcha a principios de los años 90, siendo gobernador George W. Bush. Su principal impulsor fue Rodney Paige, superintendente de las escuelas de Houston y posteriormente Secretario de Educación cuando aquel accedió a la presidencia de los Estados Unidos de América (Stobart, 2010: 140).

8 La acreditación de las normas ISO es una versión *fuerte* de la aplicación del Modelo de Gestión de Calidad en los centros escolares.

9 Así, según refiere el propio informe, en la prueba de 2006 la puntuación media de todos los centros fue de 488,60, mientras que en la de 2010 fue de 491,47, apreciándose, por tanto, una diferencia positiva de 2,87. En los centros acogidos al Programa esa diferencia fue

Sin olvidar las estrategias de manipular las estadísticas para que los números justifiquen lo que previamente se había decidido justificar, es evidente que el mecanismo en cuestión no ha conseguido mejorar significativamente los resultados ni las prácticas educativas. El propio informe PISA 2006 revela que no existe una relación estadísticamente significativa entre el rendimiento de los alumnos y la existencia de mecanismos de evaluación (OCDE, 2008: 253).

A modo de conclusión

Diane Ravitch (2014) constata que, a pesar de que las políticas de evaluación se vienen aplicando en USA desde hace más de quince años, los resultados que obtienen los alumnos norteamericanos en las prestigiosas pruebas PISA apenas han cambiado. Sin embargo, afirma, lejos de sentirse aludidos, los gobernantes las siguen impulsando, sin asumir la evidencia de que no producen mejores resultados ni mejor educación. En su lugar, dice, se culpa del fracaso a los profesores y a las familias.

Si nos situáramos en una perspectiva exclusivamente técnica y nos moviéramos en un campo distinto al de la educación, hace tiempo que se habría abandonado el mecanismo de la evaluación tal y como lo aplican las políticas neoliberales. Los resultados no justifican su vigencia. Resumidamente puede decirse que su incapacidad proviene de que se trata de un mecanismo simple que trata de actuar en un sistema complejo.

Entonces ¿cómo se explica la insistencia en centrar en la evaluación la mejora de la educación? Como demuestran Malin y Lubiensnsky (2015), en el campo de la educación existe una llamativa desconexión entre los resultados de las investigaciones y la formulación de políticas. Su hipótesis es que la política educativa se rige más bien por factores mediáticos que por conocimientos expertos y contrastados. De manera que, según esto, la omnipresencia del dispositivo basado en la evaluación no obedece tanto a su potencialidad como instrumento de mejora de la educación, sino a razones de otro tipo.

A este respecto, Bolívar (1994) sugiere que la evaluación confiere a determinadas políticas un plus de legitimidad, pues las arroja en una supuesta racionalidad científica y técnica. La evaluación es un proceso de claros componentes políticos que trata del control de un campo, el de la educación, que es difícil de controlar. Parte de ese proceso consiste precisamente en legitimar la autoridad de quien lo implementa, mediante la apariencia de una actividad técnica. Bajo la cobertura científica, añade el catedrático de la Universidad de Granada, se justifican las prescripciones de una determinada política educativa, cerrándose, de paso la posibilidad de la discusión pública que, en todo caso, se remite al campo de los problemas técnicos de la estadística.

Desde una perspectiva complementaria, la dinámica de la evaluación y autoevaluación puede interpretarse como un dispositivo de control e interiorización de la culpabilidad ante el contrastado fracaso de la política educativa en la mejora significativa de la calidad de la educación¹⁰. Se trataría de una tecnología del poder que permite ocultar su incapacidad y, a la vez, gobernar a los centros escolares y a los profesores, haciéndolos responsables de los resultados de los alumnos. En este sentido, Malen (1994: 251) afirma que, a la vista de la débil relación que existe entre las iniciativas gubernamentales y la mejora de la calidad educativa y del aprendizaje de los alumnos, su persistencia se explica por la utilidad política que aportan. El énfasis en la evaluación y autoevaluación es reflejo de la nueva relación entre política y educación, en la que se ha pasado de un *Estado reformador* a un *Estado evaluador*:

de 9,86. Si tenemos en cuenta que en la escala utilizada la media era 500 y la desviación típica 100, es evidente que una diferencia de 7 puntos es claramente irrelevante.

10 Evidentemente no puede decirse lo mismo en lo que respecta al objetivo de la escolarización, pero no hay que dar por supuesto que más tiempo de escolarización produce más y mejores aprendizajes.

Podemos considerar la delegación de poderes como una completa abdicación de Estado respecto a sus responsabilidades (...) o como un abandono selectivo de áreas en las que es difícil cosechar éxitos, como la igualdad de oportunidades (...); dejar la responsabilidad de las decisiones educativas en manos de las instituciones pedagógicas y de las familias constituye una estrategia eficaz para "culpabilizar" a otros (Whitty, Power y Halpin, 1999:65).

Referencias bibliográficas

Barrenechea, I. (2010): Evaluaciones estandarizadas: seis reflexiones críticas, *Archivos Analíticos de Políticas Educativas*, 18 (8). Disponible en: <http://epaa.asu.edu/ojs/article/751>

Bogotch, Ira, Miron, Luis y Biesta, Gert (2007): Effective for what? Effective for whom? En Townsend, T. *International Handbook of School Effectiveness and Improvement* (Nueva York, Springer).

Orden de 28 de Junio 2006. BOJA de 4 de agosto de 2006.

Bolívar, A. (1994): Autoevaluación institucional para la mejora interna. En M.A. Zabalza, *Reforma educativa y organización escolar* 915-944 (Santiago de Compostela, Tórculo).

Bolívar, A. (2006): Evaluación institucional: entre el rendimiento de cuentas y la mejora interna, *Gest. Ação*, v.9, n.1, 37-60, jan./abr.

Cantón, Isabel (2004): *Planes de mejora en los centros educativos* (Archidona, Aljibe).

Carabaña, J. (2009): *Las diferencias entre países y regiones en las pruebas PISA*. Documento 2 del Colegio Libre de Eméritos. Disponible en http://www.colegiodeemeritos.es/docs/repositorio//es_ES//documentos/pisa_carabana_%28vf%29.pdf

Carrasco Rozas, Alejandro (2007): Mejoramiento escolar en contextos de pobreza. Problematizando supuestos. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*. Vol 5, N.º 5, 291-299.

Carrasco Rozas, Alejandro (2008): Investigación en Efectividad y Mejora Escolar: ¿Nueva Agenda?. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, Vol 6, N.º 4, 5-23.

Creemers, B. (2001): La base de conocimientos de eficacia escolar. En D. Reynolds y otros, *Las escuelas eficaces. Claves para mejorar la enseñanza* (51-70) (Madrid, Santillana).

Department For Education and Skills (2004): *A New Relationship with Schools: Improving Performance through School Self-Evaluation* (Nottingham, DfES Publications) Disponible en: <http://webarchive.nationalarchives.gov.uk/20130401151715/http://www.education.gov.uk/publications/eOrderingDownload/1290-2005PDF-EN-01.pdf>

Durán Jara, Diego (2008): El rol del estado en la eficacia y mejora de la educación ¿Hacia una imposición ideológica? *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*. Vol. 6, N.º 4, 70-79.

Etxague, Xavier y otros (2009): Diferencias en la percepción de variables educativas antes y después de la implantación de los sistemas de acreditación de calidad. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*. Vol. 7, N.º 3. Disponible en <http://www.rinace.net/reice/numeros/arts/vol7num3/art1.pdf>

Evaluación de diagnóstico. Informe 2006-2007. Consejería de Educación. Junta de Andalucía.

Disponible: http://www.juntadeandalucia.es/educacion/portal/com/bin/Contenidos/OEE/evaluacion/informe_evaluacion_diagnostico_2006_2007/1181722468254_informe_finalx_12_de_junio.pdf

Evaluación de diagnóstico. Informe 2007-2008. Avance Consejería de Educación. Junta de Andalucía.

Disponible en: http://www.juntadeandalucia.es/educacion/portal/com/bin/Contenidos/OEE/AvanceDiagnostico0708/1217447709315_avanceevaluaciondediagnostico0708.pdf

Evaluación de diagnóstico. Curso 2008-2009. Consejería de Educación. Junta de Andalucía. Disponible en: http://www.juntadeandalucia.es/educacion/portal/com/bin/Contenidos/OEE/evaluacion/Evaluaciondiagnostico0809/1246617499991_evaluaciondiagnostico2008_09.pdf

Evaluación del Plan de calidad (2011). Consejería de educación. Junta de Andalucía. Disponible en: (http://www.juntadeandalucia.es/educacion/portal/com/bin/Contenidos/OEE/evaluacion/evaluacionprocalymej/1317638665039_eval0pcmr.pdf)

García Pérez, J. I., Pinto Pradas, J. L. y Robles Zurita, J. A. (2009): *Diferencias regionales y rendimiento educativo. Análisis de resultados del estudio PISA-2006* (Sevilla, Centro de Estudios Andaluces).

Hopkins, D. y Lagerweij, N. (2001): La base de conocimientos de mejora de la escuela. En D. Reynolds, R. Bollen, B. Creemers, D. Hopkins, L. Stoll y N. Lagerweij (Eds.), *Las escuelas eficaces. Claves para mejorar la enseñanza* (71-101) (Madrid, Santillana-Aula XXI).

Ladd, H. F. y Walsh, R. P. (2002): Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review*, 21, 1-17.

Lauder, H., Jamieson, I. y Wikely, F. (2001): Los modelos de la escuela eficaz: limitaciones y capacidades. En Slee, R., G. Weiner, et al. (2001): *¿Eficacia para quién? crítica de los movimientos de las escuelas eficaces y de la mejora escolar* (69-95) (Madrid, Akal).

Lingard, B.; Ladwing, J. y Luke, A. (2001): Los efectos de la escuela en las condiciones posmodernas. En Slee, R.; Weiner, G. et al. (2001): *¿Eficacia para quién? crítica de los movimientos de las escuelas eficaces y de la mejora escolar* (115-135) (Madrid, Akal).

López García, L. (coord.) (1997): *Modelo Europeo de Gestión de Calidad. Guía para la Autoevaluación* (Madrid, Ministerio de Educación y Cultura).

López Martín, E., Navarro Asencio, E., Ordóñez Camacho, X. G., Romero Martínez, S. J. (2009): Estudio de variables determinantes de eficiencia a través de los modelos jerárquicos lineales en la evaluación PISA 2006: el caso de España. *Archivos Analíticos de Políticas Educativas*, Vol. 17, N.º 16. Disponible en <http://epaa.asu.edu/ojs/article/viewFile/18/18>

Malen, B. (1994): Enacting Site-Based Management: A Political Utilities Analysis, *Educational Evaluation and Policy Analysis*, 16 (3), 249-267.

Malin, J. R., & Lubienski, C. (2015): Educational expertise, advocacy, and media influence. *Education Policy Analysis Archives*, 23 (6). Disponible en: <http://dx.doi.org/10.14507/epaa.v23.1706>

Maroy C. y Mangez C. (2008): Rationalisation de l'action publique ou politisation de la connaissance? *Revue française de pédagogie*, 164, juillet-août-septembre, 87-90.

Merchán, F. Javier (2009a): La cuestión del cambio de la práctica de la enseñanza y la necesidad de una teoría de la acción en el aula. *Revista Iberoamericana de Educación*, 8/6, 10-03-08. Disponible en: <http://www.rieoei.org/2679.htm>

Merchán, F. Javier (2009b): Hacer extraño lo habitual. Microsociología del examen en la clase de Historia. *Íber. Didáctica de las Ciencias Sociales*, N.º 60 Abril-Junio. 21-34.

Merchán, F. Javier (2010): Las evaluaciones y pruebas de diagnóstico como instrumentos de la política educativa: significado y balance de su aplicación. *Con-Ciencia Social*, 14, 63-80.

Merchán, F. Javier (2011): La evaluación de centros y profesores como instrumento de la política educativa. *Organización y Gestión Educativa*, 1, Enero-Febrero, 11-15.

Merchán, F. Javier (2012): La introducción en España de la política educativa basada en la gestión empresarial de la escuela: el caso de Andalucía. *Archivos Analíticos de Políticas Educativas*, 20 (32). Disponible en: <http://epaa.asu.edu/ojs/article/view/1059>

Ministerio de Educación y Ciencia. Secretaria General de Educación. Instituto de Evaluación (2007): *PISA 2006. Programa para la Evaluación Internacional de Alumnos de la OCDE. Informe español* (Madrid, MEC).

Mons, Nathalie (2009): Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie*, 169 99-140. Disponible en: http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/111FR.pdf

Muñoz-Repiso, M., et. al (2001): Aportaciones de las Teorías de la Organización al nuevo movimiento teórico-práctico de Mejora de la Eficacia Escolar. *Revista española de Pedagogía*, 218, 68-84.

Muñoz-Repiso, M. y Murillo, F. J (2001): Un balance provisional sobre la calidad en educación. Eficacia escolar y Mejora de la Escuela. *Organización y gestión educativa*, 4 3-9.

OCDE (2008): *PISA 2006: Informe PISA 2006. Competencias científicas para el mundo de mañana* (Madrid, Santillana).

Perrenoud, P. H. (2000): *Obligation de compétence et analyse du travail: rendre compte dans le métier d'enseignant*. Disponible en: https://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_2000/2000_29.html

Perrow, C. (1991): *Sociología de las organizaciones* (Madrid, McGraw-Hill).

Ravitch, Diane (2014): The Mythe of Chinese Super Schools. Nueva York. *The New York Review of Books*. Disponible en: <http://www.nybooks.com/articles/archives/2014/nov/20/myth-chinese-super-schools/>

Reynolds, David y Stoll, Louise (2001): La fusión de eficacia escolar y mejora de la escuela. La base de conocimientos. En Reynolds, David et al. *Las escuelas eficaces. Claves para mejorar la enseñanza*, 103-118 (Madrid, Santillana).

Scheerens, J. (1998): The school effectiveness knowledge base as a guide for school improvement. En Hargreaves, A. y cols (eds): *International handbook of educational change* (1096-1115) (Dordrecht, Kluwer).

Stobart, G. (2010): *Tiempos de pruebas: Los usos y abusos de la evaluación* (Madrid, Morata).

Whitty, G.; Power, S., y Halpin, D. (1999): *La escuela, el estado y el mercado. Delegación de poderes y elección en educación* (Madrid, Fundación Paideia/Morata).

Wrigley, T. (2007): *Escuelas para la esperanza. Una nueva agenda hacia la renovación* (Madrid, Morata).

Zaho, Yong (2014): *Who's Afraid of the Big Bad Dragon? Why China Has the Best (and Worst) Education System in the World* (San Francisco, Jossey Bass).