

LOS LÍMITES Y LAS POSIBILIDADES DE LA EVALUACIÓN COMO ESTRATEGIA DE MEJORA DE LA PRÁCTICA DE LA ENSEÑANZA

F. Javier Merchán Iglesias

jmerchan@us.es

Universidad de Sevilla

¿POR QUÉ LA EVALUACIÓN SE HA CONVERTIDO EN UN DISCURSO DE MODA?

Ya en 1922 W. W. Charters había afirmado que «la historia de la educación americana es una crónica de modas pasajeras» (vid. Kliebard, 2002, p. 1), un juicio que desde entonces bien puede hacerse extensivo a otros países y épocas. Pero las modas no son meras ocurrencias, ni adquieren protagonismo de forma azarosa; en lo que respecta al menos a la dinámica de la vida social, de la que evidentemente forma parte la educación, la preponderancia de determinadas ideas, estrategias y planes de actuación debe analizarse en contextos más amplios que el mero ámbito de su aplicación o el de los argumentos que las justifican. Es el caso de la evaluación como estrategia de mejora de la educación.

Aunque la fórmula aparece en España ya en la Ley General de Educación de 1970, hace ahora cuarenta años, la evaluación está de moda o, al menos, una determinada perspectiva de lo que se entiende por tal, si bien tampoco se trata de ninguna novedad. Frente a las tesis del cambio curricular como estrategia de mejora, que predominó en los años ochenta del pasado siglo, hoy la evaluación se ha convertido en el centro de la gestión del cambio en educación, determinando, de paso, la agenda de la investigación educativa. A este respecto, puede afirmarse que, en estos tiempos, no hay política educativa que no subraye el papel de la evaluación, ni congreso o reunión científica que no dedique parte o todo su contenido al susodicho tema. Semejante consenso debe suscitar la sospecha del investigador, pues, como se ha dicho anteriormente, las modas no son productos del azar. A lo largo de mi intervención en esta mesa redonda trataré, por una parte, de indagar acerca de las circunstancias que han propiciado la centralidad de la evaluación en la educación, trataré también de establecer cuáles son los ingredientes teórico-prácticos de los que se nutren las ideas dominantes sobre la evaluación

y aportar algunas consideraciones acerca de los resultados que producen en la práctica la aplicación de las políticas educativas basadas en la evaluación como instrumento de mejora. Por otra parte, es mi intención someter a discusión las posibilidades, dificultades y límites de la evaluación de la práctica de la enseñanza, para concluir con algunas propuestas sobre el objeto de este debate.

En un interesante trabajo de revisión sobre el papel de la evaluación en el pilotaje de los sistemas educativos, Natalie Mons (2009) identifica los precedentes sobre los que se va configurando la nueva y central dimensión de la evaluación. Se refiere, por una parte a la conocida como teoría del capital humano sobre la educación. Básicamente esta teoría, sostenía la necesidad de vincular la educación con el desarrollo económico, de manera que esta se entendía como una inversión que repercute en la mejora de la capacidad productiva de un país. Desde esta perspectiva se ha considerado también que «las escuelas pueden ser vistas como las firmas que se especializan en <la producción> de la educación. El establecimiento educativo, que incluye todas las escuelas, puede ser visto como una industria» (Schultz, 1963, p. 4, citado en Tröhler, 2009, p. 7). Lejos de haber pasado a la historia, esta tesis sobre el lugar de la educación en la vida social ha cobrado notable actualidad en nuestros días. Así, en el texto del Consejo de Europa celebrado en Lisboa en el año 2000, se plantea como objetivo estratégico de la educación para el año 2010: «Llegar a ser la economía basada en el conocimiento más competitiva y dinámica del mundo, capaz de tener un crecimiento sostenible con más y mejores trabajos y con una mayor cohesión social». Al entender la educación como inversión económica y la escuela como una industria, cae por su peso que se revisen periódicamente los logros alcanzados, de manera que la evaluación toma la forma de análisis de resultados, una tendencia que dará lugar a la proliferación de pruebas, exámenes e indicadores de todo tipo que darán la medida del éxito y el fracaso. De esta forma se redefine el sentido de la educación, formulándose como la consecución de unos objetivos empíricamente demostrables.

Varios son los modelos de gestión que han circulado por el mundo de la empresa para gestionar la valoración de los logros y la mejora en la consecución de sus objetivos. El más generalizado es el conocido como modelo EFQM, siglas que responden al nombre de la institución que lo impulsó: *European Foundation for Quality Management*. Se trata de una agrupación de empresas centroeuropeas que institucionalizaron un premio a la mejor gestión, premio que obtuvieron corporaciones como Nestlé, Volkswagen o Rank Xerox. La idea es establecer un procedimiento de gestión basado en ciclos inacabables de autoevaluación, siempre referidos a indicadores precisos, medibles y reformulables. Este modelo inicialmente desarrollado en empresas industriales y después en empresas de servicio, es el que, de una u otra forma, se traslada al campo de la educación. En esa transferencia tuvo mucho que ver el desarrollo en los años 70 en Estados Unidos de la *New public management* y el movimiento de *Policy evaluation*. En un contexto de crisis económica y de cuestionamiento del papel del estado. El objetivo ahora es sanear y racionalizar la actuación de la administración, renovando la

gestión de lo público, atendiendo principalmente a la relación coste-eficacia. En este sentido, se considera que a) los productos y servicios públicos son mensurables, b) se deben utilizar para ello instrumentos científicos, c) los actores deben rendir cuentas ante los gestores y los ciudadanos, y c) la organización pública debe regularse por los resultados y no exclusivamente por los recursos. Naturalmente la educación y los centros escolares deben atenerse a esta nueva filosofía.

Aunque impulsada por ellas, la centralidad de la evaluación como estrategia de mejora no es meramente producto de la aplicación de las tesis economicistas que se han expuesto anteriormente. Como suele ocurrir en otros campos, los planteamientos más generales no se adoptan sin más en el ámbito de la educación, son necesarias teorías regionales. Las fuentes de estas teorías pueden encontrarse en el que se conoce como movimiento de la *eficacia escolar* y en el también conocido como movimiento para la *mejora de la eficacia escolar*. Como es sabido, el primero de ellos, de carácter más bien teórico y centrado en la investigación, surge a partir del informe Coleman. Mientras que en ese informe se subrayaba la importancia del contexto socioeconómico y cultural como factor determinante del rendimiento de los alumnos, algunos investigadores llamaron la atención sobre el hecho de que en contextos similares se producía resultados diferentes, suponiendo que esto se debe a la escuela y a sus profesores. Desde ese punto de partida, la cuestión era determinar cuáles eran los factores que están presentes en las escuelas de éxito o eficaces.

No puede decirse, sin embargo, que la investigación sobre la de identificación de los factores de éxito haya contribuido a resolver las preguntas que puedan plantearse acerca de las escuelas eficaces. Por una parte encontramos enormes listas que se han ido aumentando unos estudios tras otros: «Así, en una revisión realizada recientemente por Scheerens y Bosker [...] sobre estudios modélicos, encontraron 719 elementos relacionados con la eficacia agrupados en 13 factores generales» (Muñoz-Repiso y Murillo, 2001, p. 7). Además, como ha señalado Creemers (2001) los estudios sobre factores de éxito en distintos países llegan a conclusiones contradictorias. Por otra parte, estamos hablando de un tipo de investigación que adolece de cierto rigor científico. Wrigley (2007, p. 28) afirma que es un error asumir que la correlación estadística sea lo mismo que la causalidad, o bien, citando a Scheerens, plantea que «no es fácil evaluar las bases empíricas exactas de la lista de factores [...] La mayoría de las revisiones no expresa la significación estadística ni el tamaño de los efectos de los diversos factores en términos de su asociación con resultados de logro ajustado» (citado en Wrigley, 2007, p. 29). En esta misma línea, Carrasco (2007) considera que el movimiento de la eficacia escolar sobrestima las posibilidades de la estadística para hacerse cargo unívocamente de los procesos de la vida social. Finalmente, también se ha señalado la ambigüedad del lenguaje que se utiliza para referirse a los factores de eficacia, de manera que, detrás de la apariencia de certidumbre que confiere la exactitud matemática, se oculta una imprecisión terminológica que se presta «a la manipulación y a la reinterpretación política» (Wrigley, 2007, p. 28).

Por su parte, el movimiento de la mejora de la eficacia escolar, amparándose en el supuesto de que existe una racionalidad científico-técnica capaz de conseguir escuelas eficaces mediante la aplicación de un instrumental apropiado, ha tratado de establecer relaciones entre los estudios de los factores de eficacia y las estrategias de mejora, dando a entender que existe un vínculo entre conocimiento y política educativa. Pero, salvo que se trate de una mera justificación, el vínculo es muy poco consistente. A este respecto, se ha destacado que los estudios sobre eficacia y mejora se centran en factores sobre los que se puede actuar (Scheerens, 1998), mientras que relega aquellos sobre los que no se puede actuar desde la escuela, operando como si realmente no intervinieran; sin embargo, 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; Carabaña, 2009). La mera eliminación estadística del factor más relevante, el contexto sociocultural, puede tener alguna utilidad informativa, pero en modo alguno permite pensar que en la práctica puede actuarse en los centros escolares como si no existiera. Es cierto que algunas escuelas de áreas más pobres alcanzan resultados iguales o incluso mejores que otras de clase media, pero ello no justifica la minusvaloración del efecto del contexto, ni cambia significativamente la tendencia general, ni, sobre todo, sirve como pretexto para culpar a las otras, pues son fruto de «altos niveles de dedicación, tiempo y esfuerzo» (Wrigley, 2007, p. 31), difícilmente sostenible en el tiempo y difícilmente exportable a todos los casos. Como afirman Mortimore y Whitty (2000, p. 21), «una crítica particular al trabajo de mejoramiento escolar, es que esta perspectiva ha tendido a exagerar el punto hasta el cual escuelas individuales pueden desafiar desigualdades estructurales».

La idea de que, una vez que conocemos los factores de éxito, la mejora consiste en reproducirlos en las escuelas ineficaces, adolece de 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» (Reynolds y Stoll, 2001, p. 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» (Reynolds y Stoll, 2001, p. 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» (Reynolds y Stoll, 2001, p. 114).

Lauker y Wikely (2001) sostienen que, a pesar de la complejidad que han alcanzado los estudios cuantitativos, entre el conocimiento de los factores de éxito y la formulación de recetas para mejorar quedan muchas cuestiones que no están resueltas. 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).

ALGUNOS INTERROGANTES SOBRE LA EVALUACIÓN COMO ESTRATEGIA DE MEJORA DE LA EDUCACIÓN

Aunque por razones políticas no gozan de suficiente visibilidad, son numerosos los estudios que ponen de manifiesto que el modelo de evaluación dominante no constituye, como se pretende, un recurso decisivo para la mejora de la educación. Desde luego en países como los Estados Unidos o Inglaterra, en los que la fórmula está aplicándose ya desde hace casi dos décadas, no puede decirse que haya mejorado significativamente la formación de niños y jóvenes; ni siquiera puede decirse que mejoren los resultados (vid. Merchán, 2010 y 2011, en prensa). En España, la Comunidad Andaluza fue la primera que puso marcha las llamadas Pruebas o Evaluaciones de diagnóstico, verdadero estandarte, junto con las pruebas PISA, de la dinámica de la evaluación en sistema educativo. En una nota de prensa fechada el 18 de octubre de 2006, en la que se reseñaba la intervención de la Consejera de Educación en el Parlamento andaluz, decía la citada responsable que «que las pruebas proporcionarán información ‘objetiva y rigurosa’ tanto para la mejora del rendimiento del alumnado como de la práctica educativa» (Consejería de Educación, 2006), sin embargo, cinco años después, los datos ponen de manifiesto que aquel vaticinio era meramente retórico, pues nada ha cambiado desde entonces. De manera que resulta algo sospechosa la importancia que las actuales políticas educativas conceden a la evaluación de centros y profesores. Cabe concluir que sus pretendidas bondades no se justifican por los resultados que se obtienen en la mejora de la educación, lo que nos hace pensar que la vigencia y publicidad del mecanismo se explica más bien por la rentabilidad política que se deriva de trasladar la responsabilidad fundamental de la mejora de la educación a los centros y a los profesores.

Pero, más allá del concepto imperante y de su instrumentalización por las políticas educativas dominantes, la evaluación de la educación plantea problemas conceptuales y prácticos que nos hacen ser precavidos sobre su potencialidad como estrategia de mejora. La tesis que, finalmente, quiero desarrollar es la de que la evaluación de la práctica de la enseñanza es un instrumento que, utilizado adecuadamente, ofrece posibilidades para desencadenar procesos de cambio, pero que, sin embargo, se trata de un recurso limitado que, por tanto, carece de fuerza suficiente para producirlos por sí sólo.

Si nos atenemos a la idea de la evaluación como análisis de resultados y mecanismo de palanca para alcanzar logros en el rendimiento de los alumnos, los hechos han demostrado su incapacidad. Pero, además, la dinámica de pruebas y evaluaciones genera toda una serie de consecuencias negativas a las que Sahron y Berliner (2007) se han referido como «efectos colaterales». En este sentido, siguiendo el principio de indeter-

minación de Heisenberg en física cuántica, se ha señalado que cuando una medida (el rendimiento de los alumnos) se convierte en objetivo, deja de ser una buena medida, pues toda medición de un sistema lo perturba (vid. Stobart, 2010, p. 135).

La evaluación de la práctica de la enseñanza, entendida como análisis de los procesos que se desarrollan en el aula ofrece, desde luego, mayores posibilidades para la mejora de la educación, si bien es una perspectiva que no deja de entrañar dificultades. Por una parte, dado que la evaluación implica valoración, se plantea el problema de disponer de referencias: ¿cómo sabemos si las cosas se hacen bien o mal, mejor o peor? No cabe duda de que los resultados que obtienen los alumnos en pruebas o exámenes constituyen un referente de los logros en su aprendizaje y, en este sentido, no voy a denostar la realización de este tipo de ejercicios. Sin embargo estos datos no pueden operar como referencias absolutas ya que, en primer lugar, la enseñanza produce a veces aprendizajes diferidos en el tiempo, aprendizajes que difícilmente pueden apreciarse en el plazo de realización de las pruebas, y, en segundo lugar, los exámenes y ejercicios de cualquier tipo, en el contexto escolar, no siempre son instrumentos apropiados para valorar los aprendizajes adquiridos. En este sentido, mi propuesta es la de diversificar lo más ampliamente posible los instrumentos de medición, no centrándonos exclusivamente en las pruebas y los exámenes, procurando, en todo caso, diversificar también los modelos de estas pruebas. Esto nos permitirá disponer a la vez de registros cuantitativos y cualitativos, y, en definitiva, de una comprensión más amplia de los problemas de aprendizaje.

Por otra parte, a la hora de valorar los procesos que se desarrollan en el aula, al margen de los resultados, nos encontramos con la dificultad de formular juicios sobre la práctica de la enseñanza. Sin afirmar que todo vale, a este respecto, Perrenoud (2000) sostiene que en el aula se producen una gran variedad de situaciones y que el estado de nuestro conocimiento sobre la enseñanza no permite un discernimiento claro sobre cuáles son las actuaciones más eficaces. Es posible que dispongamos de algunas ideas claras acerca de cómo se produce el aprendizaje en contextos cotidianos, pero el contexto escolar es un medio con muchas y significativas particularidades en el que quizás no se aprenda de la misma forma. En todo caso, sabiendo cómo se aprende no necesariamente sabemos cómo debemos enseñar.

Cuando tratamos de la evaluación de la práctica de la enseñanza en un sentido más amplio que el mero recuento de resultados, se persigue una mejor comprensión de los procesos que se desarrollan en el aula y de las dificultades de aprendizaje que en ese contexto se advierten en los alumnos. Es decir, siguiendo el lema que reza en la página web de la Agencia Andaluza de Evaluación Educativa, se evalúa para conocer, suponiendo que del conocimiento que se produce se derivan estrategias de mejora. La idea está tomada directamente del campo de la medicina, en el que se establece una relación más o menos directa entre el diagnóstico y la curación. Sin embargo, sin negar el interés y las posibilidades de este planteamiento, entiendo que subyace en el una

lógica aparente, de mero sentido común que, trasladada al campo de la práctica de la enseñanza y de la acción humana en general, adolece de cierta simplicidad.

La educación es un fenómeno complejo, de tal manera que, solamente una autorreflexión acerca de la realidad de los fenómenos difícilmente puede producir una adecuada comprensión sobre ella. En este sentido mi propuesta es que la evaluación, como análisis de los problemas de aprendizaje en el aula, requiere un marco teórico apropiado, así como disponer de mayor información de la que se deriva de la mera observación de los hechos. No es fácil que en el contexto escolar pueda desarrollarse estos recursos, ya que la evaluación está mediatizada por problemas burocráticos y organizativos, por las urgencias de la acción y, en definitiva, por el hecho de que en ese contexto priman otras finalidades y otro sentido distinto del de la producción de conocimiento. Naturalmente estos condicionantes no implican necesidad, pero habrá que convenir que limitan en la práctica sus posibilidades.

Finalmente, entendiendo la evaluación como producción de conocimiento para la mejora, se plantea otro problema al que debemos prestar atención. Me refiero a la relación entre conocimiento y acción, pues se supone que la mejora requiere nuevas pautas de actuación en el aula. Desde luego, de los resultados de las pruebas de diagnóstico y productos similares como PISA, no se desprende un conocimiento relevante sobre la realidad ni, sobre todo, ninguna pauta de actuación (Carabaña, 2009); en términos más generales Dewey sostenía que «los resultados científicos proporcionan una regla para conducir la observación y la indagación, no para la acción», es decir, el conocimiento no dice a los profesores lo que deben hacer (vid. Kliebard, 2000, p. 130). Por lo demás, aunque estableciéramos conclusiones científicas, basadas en el conocimiento sobre la realidad, acerca del modo en que los docentes deben actuar en el aula, esto no quiere decir que realmente pueda actuarse conforme a esas directrices, pues la actuación de las personas, como la actividad de los profesores en el aula, no se rige única ni fundamentalmente por lo que sabemos. No siempre hacemos lo que queremos, ni queremos lo que hacemos. Además, no conviene olvidar que entre enseñanza y aprendizaje no existe una relación causal directa pues en esa interacción intervienen muchos otros factores. Mi propuesta en este sentido es, desde luego, profundizar en el conocimiento de la práctica de la enseñanza mediante la evaluación entendida como análisis de los procesos y los logros, pues es evidente que el saber ayuda en la acción; pero, dada la complejidad de la vida en el aula, debemos relativizar las posibilidades del conocimiento para la mejora. Al fin y al cabo no estamos tratando de fenómenos físicos ni mecánicos, sino de fenómenos sociales.

REFERENCIAS BIBLIOGRÁFICAS

CARABAÑA, J. (2009). Las diferencias entre países y regiones en las pruebas PISA. Documento 2 del Colegio Libre de Eméritos, en http://www.colegiodeemeritos.es/docs/repositorio//es_ES//documentos/pisa_carabana_%28vf%29.pdf. Consultado el 10 de mayo de 2010.

- CARRASCO ROZAS, A. (2007). Mejoramiento escolar en contextos de pobreza. Problematizando supuestos desde la evidencia. *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*. Vol. 5 (5e), 291-299, <http://www.rinace.net/arts/vol5num5e/art40.pdf>. Consultado el 15 de mayo de 2010.
- CARRASCO ROZAS, A. (2008). Investigación en Efectividad y Mejora Escolar: ¿Nueva Agenda? *Revista Electrónica Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*. Vol. 6 (4), 6-23. <http://www.rinace.net/reice/numeros/arts/vol6num4/art1.pdf>. Consultado el 10 de junio de 2010.
- CONSEJERÍA DE EDUCACIÓN. Nota de prensa de 18 de octubre de 2006. En http://www.juntadeandalucia.es/educacion/nav/contenido.jsp?pag=/Contenidos/GabinetePrensa/Notas_de_prensa/2006/Octubre/nota_prensa_181006_pruebadiagnostico_parlamento&vismenu. Consultado el 18 de marzo de 2010.
- CREEMERS, B. (2001). La base de conocimientos de eficacia escolar. En D. Reynolds y otros (eds.). *Las escuelas eficaces. Claves para mejorar la enseñanza*. Madrid: Santillana, 51-70.
- KLIEBARD, H. M. (2002). *Changing Course. American Curriculum Reform in the 20th century*. New York: Teacher College Press.
- LAUDER, H., JAMIESON, I. y WIKELY, F. (2001). Los modelos de la escuela eficaz: limitaciones y capacidades. En R. G. Slee y otros (dirs.). *¿Eficacia para quién? crítica de los movimientos de las escuelas eficaces y de la mejora escolar*. Madrid: Akal, 69-95.
- MERCHÁN, F. J. (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. J. (2011). La evaluación de centros y profesores como instrumento de la política educativa. *Organización y Gestión Educativa* (en prensa).
- MONS, N. (2009). Effets théoriques et réels des politiques d'évaluation standardisée. *Revue française de pédagogie*, 169, 99-140. Disponible también en: http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/111FR.pdf.
- MORTIMORE, P. y WHITTY, G. (2000). *Can school improvement overcome the effects of disadvantage?* London: Institute of Education University of London.
- 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.
- PERRENOUD, P. (2000). Obligation de compétence et analyse du travail: rendre compte dans le métier d'enseignant, en: https://www.unige.ch/fapse/SSE/teachers/perrenoud/php_main/php_2000/2000_29.html. Consultado el 12 de mayo de 2010.
- REYNOLDS, D. y STOLL, L. (2001). La fusión de eficacia escolar y mejora de la escuela. La base de conocimientos. En D. Reynolds y otros (eds.). *Las escuelas eficaces. Claves para mejorar la enseñanza*. Madrid: Santillana.

- SCHEERENS, J. (1998). The school effectiveness knowledge base as a guide for school improvement. En A. Hargreaves y cols. (eds.). *International handbook of educational change*. Dordrecht: Kluwer.
- SHARON, N. y BERLINER, D. (2007). *Collateral Damage. How Highstakes testing corrupts America's schools*. Cambridge, Mass.: Harvard Education Press.
- STOBART, G. (2010). *Tiempos de pruebas: los usos y abusos de la evaluación*. Madrid: Morata.
- TRÖHLER, D. (2009). Armonizar el globo educativo. Política mundial, rasgos culturales y los desafíos a la investigación educativa. *Profesorado. Revista de currículo y formación del profesorado*, 13 (2), en <http://www.ugr.es/~recfpro/rev132FIRINV1.pdf>. Consultado el 10 de mayo de 2010.
- WRIGLEY, T. (2007). *Escuelas para la esperanza. Una nueva agenda hacia la renovación*. Madrid: Morata.