

Un modelo integrado de evaluación para el sistema educativo español

Dr. Francisco López Rupérez

Cátedra de Políticas Educativas de la UCJC

Índice

1. Introducción	3
2. El <i>feedback</i> y la evaluación formativa.....	6
2.1 Los antecedentes conceptuales de la noción de <i>feedback</i> en educación	6
2.2 <i>Feedback</i> y evaluación formativa	7
2.3 La evaluación formativa en diferentes niveles del sistema educativo	8
2.3.1 En el nivel <i>micro</i> : el aula	8
2.3.2 En el nivel <i>meso</i> : el centro	9
2.3.3 En el nivel <i>macro</i> : el sistema	11
3. Un modelo integrado de evaluación para el sistema educativo español	12
3.1 Evolución normativa de la evaluación del sistema educativo en España	12
3.1.1 LOGSE	12
3.1.2 LOPEGCE.....	13
3.1.3 LOCE	14
3.1.4 LOE.....	17
3.2 La situación actual	19
3.2.1 LOMCE	19
3.2.2 La rectificación del Real Decreto-ley.....	21
3.3 Bases para un modelo multinivel integrado de evaluación.....	23
3.3.1 Un enfoque formativo	23
3.3.2 Una función diagnóstica	23
3.3.3 Una dimensión de responsabilidad	24
3.3.4 Una aproximación multinivel	24
3.3.5 Un análisis competencial	24
3.3.6 Un criterio económico	25
3.4 El modelo propuesto	26
3.4.1 En el nivel del centro educativo.....	28
3.4.2 En el nivel de la Comunidad Autónoma	28
3.4.3 En el nivel del conjunto del Estado.....	29
4. Conclusiones y recomendaciones	31
4.1 Conclusiones	32
4.2 Recomendaciones	33
5. Referencias.....	34

Un modelo integrado de evaluación para el sistema educativo español

Dr. Francisco López Rupérez

Cátedra de Políticas Educativas de la UCJC

Resumen

Las evaluaciones nacionales e internacionales se han convertido en elementos centrales para asegurar la calidad y mejorar la educación. Aunque la evaluación genera una información valiosa solo se convierte en *feedback* cuando es utilizada para reducir la brecha existente entre el nivel actual y el nivel de referencia que se pretende alcanzar. Por ello, lo importante no es tanto la información que produce la evaluación como lo que se hace con esa información. El presente trabajo efectúa una revisión del tratamiento dado a la evaluación general del sistema educativo español en las leyes educativas. A partir de un análisis crítico, se propone y justifica un modelo evolucionado, con respecto a los anteriores, que se apoya en los siguientes criterios básicos: un enfoque formativo, una función diagnóstica, una dimensión de responsabilidad, una aproximación multinivel, un análisis competencial y un principio económico.

Abstract

National and international assessments have become very important tools to guarantee the quality of education and improve it. The assessment produces a valuable information but only is feedback when it is used to alter the gap between the actual level of performance and the desired reference level. For this reason, the relevant thing is not so much the information that the evaluation produces as what is done with that information. The present work makes an overview of the general assessments approach of the Spanish educational system in the education laws. From a critical analysis, it proposes and justifies an evolved model that is founded on the following basic criteria: a formative approach, a diagnostic function, a responsibility dimension, a multilevel way, an institutional competence analysis and an economic principle.

1. Introducción

La evaluación de los sistemas educativos ha ido adquiriendo una importancia creciente, a lo largo de las últimas décadas, en los países desarrollados. Este fenómeno se ha ido extendiendo a otros países de inferior nivel de desarrollo hasta ocupar, en el momento presente, un lugar central a nivel internacional en el ámbito de las políticas educativas; o, si se quiere, en el de las 'metapolíticas', entendidas aquí como políticas de segundo orden, vale decir, políticas para concebir y desarrollar otras políticas¹.

¹ LÓPEZ RUPÉREZ, F. (2001). *Preparar el futuro. La educación ante los desafíos de la globalización*. Cap. 5. Metapolíticas en Educación. Ed. La Muralla. Madrid.

Estaríamos, pues, ante un elemento sustantivo de lo que se ha dado en llamar ‘globalización de las políticas educativas’².

Según De Ketele *et al.* (1993), «Evaluar significa confrontar un conjunto de informaciones con un conjunto de criterios para tomar una decisión»³. El desarrollo progresivo de una sociedad y de una economía basadas en el conocimiento ha ido revalorizando el papel de la educación –en tanto que institución social que forma a las personas y trabaja con el conocimiento–, mucho más allá de los intereses puramente sectoriales. En este contexto, la búsqueda de políticas educativas que se traduzcan en actuaciones eficaces y eficientes se ha convertido en una meta compartida por gobiernos de numerosos países situados alrededor de todo el globo. Como han subrayado Baird *et al.*⁴, citando a Higham y Yeormans (2010), las evaluaciones nacionales e internacionales se han convertido en elementos centrales para asegurar la calidad y mejorar la educación; se consideran decisivas para la definición de las políticas pues actúan a modo de palancas y controles sobre los propios sistemas educativos. En este mismo orden de ideas, las políticas basadas en el conocimiento, en la evidencia empírica y en la investigación han venido a ser un elemento de amplio consenso como factor de calidad de la gobernanza educativa⁵.

Además de lo anterior, la revalorización del dato, que ha traído consigo la revolución digital con sus posibilidades amplificadas de computación, está convirtiéndose en un factor cultural característico de las sociedades desarrolladas del siglo XXI; lo que facilita la toma en consideración, por las ⁴ instancias políticas de decisión, de las evaluaciones cuantitativas a gran escala y atempera aquellas resistencias, ideológicamente fundadas, que han estado presentes en determinados sectores, tanto políticos como educativos. De otro modo, sería difícil explicar el hecho de que el número de países –o de economías– participantes en las pruebas de PISA se haya duplicado a lo largo de sus seis ediciones y alcance, en la actualidad, a más países en desarrollo que propiamente desarrollados⁶. O que, en lo que a España respecta, se haya producido –por primera vez en la historia de PISA– la participación, con muestra ampliada, de sus 17 Comunidades Autónomas en la edición 2015 de este programa de evaluación internacional.

Por otra parte, la implicación de la OCDE en este tipo de evaluaciones de los sistemas educativos ha superado, con creces, el impulso, la consideración y el impacto de que ha gozado la evaluación del rendimiento de los alumnos efectuada desde la IEA (International Association for the Evaluation of Educational Achievement), a través, entre otros, del TIMSS (Trends in International Mathematics and Science Study) y del PIRLS (Progress in International Reading Literacy), desde 1995 y 2001 respectiva-

2 PONT FERRER, B. (2017). *Reformas educativas: el caso del liderazgo escolar en perspectiva comparada*. Tesis Doctoral. Universidad Complutense de Madrid. Madrid.

3 DE KETELE, J. M.; ROEGIERS, X. (1995). *Metodología para la recogida de la información*. Ed. La Muralla. Madrid.

4 BAIRD, J., HOPTENBECK, T. N.; NEWTON, P. E.; STOBART, G.; STEEN-UTHEIM, A. T. (2014). *Assessment and Learning: State of the Field Review*. Norwegian Knowledge Centre for Education & University of Oxford. <https://www.researchgate.net/publication/263654863_Assessment_and_Learning_State_of_the_Field_Review>.

5 LÓPEZ RUPÉREZ, F.; GARCÍA GARCÍA, I.; EXPÓSITO CASAS, E. (2017). *La calidad de la gobernanza del sistema educativo español. Un análisis empírico*. Universidad Camilo José Cela. Madrid.

6 OECD (2016). *PISA 2015 Results. Policies and Practices for Successful Schools. Volume I*. OECD Publishing. Paris.

mente⁷. El hecho de que la OCDE sea un organismo intergubernamental de los países desarrollados –consolidado y de gran prestigio en ambientes tanto políticos como académicos–, con mecanismos sistemáticos y efectivos de consulta y de implicación de los países miembros, es otro de los factores explicativos del avance de las evaluaciones de rendimiento de los alumnos a gran escala de la mano de PISA. Más allá de las críticas fundadas que puedan efectuarse a este programa⁸, lo cierto es que ha contribuido, de un modo decisivo, a consolidar la evaluación cuantitativa y sus consecuencias de cara tanto a la orientación de las políticas educativas, como al desarrollo de la investigación con una base empírica.

España se ha incorporado, en lo indispensable, a ese movimiento mundial. No obstante, la falta de estabilidad en la visión ha generado, por razones políticas, modificaciones sustantivas en la concepción del modelo de evaluación de nuestro sistema educativo de la mano de cambios no menores en el marco legal. Se han sumado a ello las crecientes tensiones territoriales en materia educativa, para desembocar en un panorama, aunque reglado, francamente desordenado. Solo nuestra participación en las evaluaciones internacionales –principalmente PISA– ha sido capaz de compensar, al menos en parte, dicho panorama al suministrar un marco estable para la evaluación y para la comparación *intra* e *inter* nacional de sus resultados. Frente a ello, resulta ocioso insistir en la importancia creciente que, a nivel internacional, se otorga a los procedimientos de evaluación de los sistemas educativos como instrumentos fundamentales para orientar las políticas y promover la mejora de sus resultados. Por los motivos antes descritos, y debido, además, a las limitaciones en cuanto a la utilización de la información generada, el actual modelo de evaluación de nuestro sistema educativo constituye una de las actuales debilidades de la Educación española que, por su importancia en el plano estratégico, es necesario corregir cuanto antes.

5

La cuestión de la 'evaluación de la calidad del sistema' ha entrado a formar parte de la agenda de reformas pactada por los partidos políticos en el seno de la Subcomisión parlamentaria del Pacto Social y Político por la Educación. Justamente por ello, procede efectuar una reflexión fundada sobre las características de un modelo de evaluación del sistema educativo español que atienda los principios de pertinencia, eficacia y eficiencia, propios de una buena gestión pública, y tome en consideración las peculiaridades de nuestro modelo territorial.

Este es el propósito principal del presente trabajo que se ha organizado, junto con esta Introducción, en los siguientes epígrafes. En el segundo apartado se abordan algunos elementos conceptuales referidos a la evaluación formativa en una aproximación multinivel; tipo éste de evaluación que, por estar claramente vinculado con la mejora, posee un indiscutible valor estratégico. En el tercer apartado se efectúa un recorrido histórico sobre la posición que ha ocupado la evaluación general del sistema español en las diferentes leyes educativas de la democracia, se efectúa un análisis específico de la situación actual en el marco de la LOMCE y se propone y justifica un modelo multinivel integra-

7 BREAKSPEAR, S. (2012). *The Policy Impact of PISA: An Exploration of Normative Effects of International Benchmarking in School System Performance*, OECD Education Papers, No. 71, OECD Publishing. <<http://dx.doi.org/10.1787/5k9fdqffr28-en>>.

8 SJOBERG, S. (2017). «PISA Testing. A Global Educational Race?». *Europhysicsnews*, 48/4, 17-20.

do de evaluación para nuestro sistema educativo. Finalmente, en el apartado cuarto, se formula un conjunto de conclusiones y de recomendaciones que puede alimentar el debate al respecto e inspirar, en su caso, la acción política correspondiente.

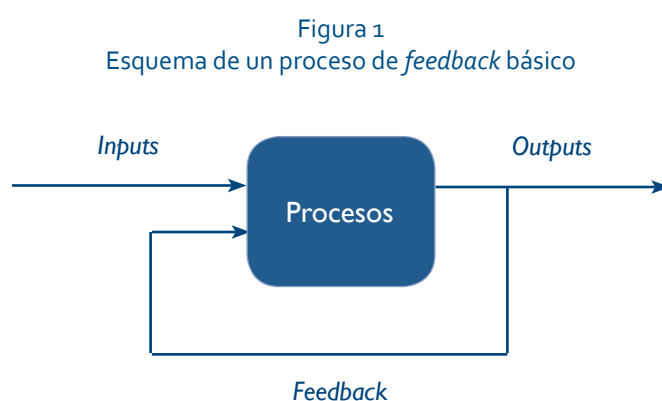
2. El *feedback* y la evaluación formativa

De acuerdo con Black y Wiliam (1998), «empleamos el término general de evaluación para referirnos a todas aquellas actividades realizadas por profesores –o por sus alumnos actuando sobre sí mismos– que proporcionan información para ser empleada en la enseñanza y el aprendizaje. Tal evaluación se convierte en formativa cuando la evidencia así generada es utilizada eficazmente para adaptar la enseñanza y hacerla confluir con las necesidades de los alumnos»⁹.

Por su parte, el *feedback* es considerado habitualmente como un elemento esencial de la evaluación formativa; pero se trata éste de un término con un amplio recorrido conceptual, más acá de la educación, al que merece la pena aproximarse, siquiera sea de un modo sintético.

2.1 Los antecedentes conceptuales de la noción de *feedback* en educación

- ⁶ Según el Diccionario Panhispánico de Dudas¹⁰, de la Real Academia Española de la Lengua, el término *feedback* es una «voz inglesa usada en distintas disciplinas científico-técnicas con los sentidos de 'retorno de parte de la energía o de la información de salida de un circuito o un sistema a su entrada', 'acción que el resultado de un proceso material ejerce sobre el sistema físico o biológico que lo



Fuente: elaboración propia.

⁹ BLACK, P.,; WILIAM, D. (1998). «Assessment and classroom learning». *Assessment in Education*, 5, 7-75

¹⁰ <<http://lema.rae.es/dpd/srv/search?key=feedback>>

origina' y, en general, 'modificación de la actitud o estrategia inicial en un proceso a partir del análisis de sus resultados'». La figura 1 ilustra, mediante un esquema, el concepto del *feedback* más simple.

Aun cuando dicho término tiene su origen en la segunda mitad del siglo XIX en el ámbito de la mecánica, y a principios del XX en el de la electrónica, lo cierto es que otras disciplinas –tales como la biología, la psicología, la cibernética o la teoría de la información– incorporaron el concepto, para alcanzar después a las ciencias sociales¹¹ y convertirse en algo de uso frecuente en el ámbito de la educación¹². Pero fue la teoría general de sistemas la que, asumiendo de forma explícita el principio de unidad del conocimiento, lo incorporó como uno de esos conceptos transversales, en el seno de un marco epistemológico de un elevado nivel de generalidad, capaz de ser aplicado a un abanico de problemas sorprendentemente amplio¹³.

Quizás uno de los aspectos más destacables, en el contexto del presente estudio, del concepto de *feedback* –o retroalimentación– sea su vinculación con los comportamientos inteligentes. Los teóricos de la inteligencia computacional han sido capaces de concluir que el *feedback* –entendido como «un proceso mediante el cual el resultado de una iteración precedente, del proceso de inteligencia, se combina con el *input* en la subsiguiente aplicación»– añade impulso ('momentum') a los procesos de aprendizaje artificial¹⁴. Algunos autores han llegado al extremo de considerar que «la esencia de la inteligencia es el *feedback*»¹⁵. Como señala Richardson en su obra antes citada, «La esencia del concepto es un círculo de interacciones, un bucle cerrado de información y de acción». Y sobre esa idea reposa, entre otros fenómenos, la adaptabilidad de los seres vivos a las exigencias del entorno, el aprendizaje de la experiencia, e incluso el 'ciclo percepción/acción'¹⁶ en el funcionamiento cerebral que, desde la neurociencia, explica, entre otras cosas, esa vinculación tan profundamente humana entre cognición y emoción^{17, 18}.

2.2 *Feedback* y evaluación formativa

Una de las áreas de la educación en donde el concepto de *feedback* se ha instalado de un modo preferente es la de la evaluación y, en particular, la de la llamada 'evaluación formativa'. Shute (2008) usa el término híbrido de '*feedback* formativo' y lo define como «información comunicada al alumno destinada a modificar su pensamiento o conducta con el propósito de mejorar el aprendizaje»¹⁹. Ha-

11 RICHARDSON, G. P. (1991). *Feedback Thought in Social Science and Systems Theory*. University of Pennsylvania Press. Philadelphia.

12 HATTIE, J.A.; TIMPERLEY, H. (2007). «The Power of Feedback». *Review of Educational Research*, 77, 81-112.

13 VON BERTALANFY, L. (1979). *Perspectivas de la teoría general de sistemas*. Alianza Editorial. Madrid.

14 KOVACH, D. (2017). «The Computational Theory of Intelligence: Feedback». *International Journal of Modern Nonlinear Theory and Application*, 6, 70-73. <<http://dx.doi.org/10.4236/ijmnta.2017.62006>>.

15 <<https://mindhacks.com/2013/03/05/the-essence-of-intelligence-is-feedback/>>.

16 FUSTER, J. (2014). *Cerebro y libertad. Los cimientos cerebrales de nuestra capacidad de elegir*. Cap. 4. El ciclo percepción/acción. Ariel. Barcelona.

17 FUSTER, J. (2012). «Entrevista». *Participación Educativa*, 2.ª época, vol. 1, n.º 1, 29-31. <ntic.educacion.es/cee/revista/pdfs/RPE_DIC2012_05Fuster.pdf>.

18 HINTON, CH.; MIYAMOTO, K.; DELLA CHIESA, B. (2008). «Brain Research, Learning and Emotions: implications for education research, policy and practice». *European Journal of Education*, 43(1), 2008, 87-103.

19 SHUTE, V. J. (2008). «Focus on formative feedback». *Review of Educational Evaluation*, 37, 3-14.

ttie y Timperley (2007)²⁰ sostienen un enfoque conceptual similar. Pero, como han advertido Havnes *et al.* (2012), hay un supuesto implícito en dicha conceptualización y es que «para que el *feedback* sea formativo y mejore el aprendizaje necesita ser formulado, suministrado y enmarcado de modo que invite a un compromiso activo con él por parte de los alumnos»²¹.

En este punto, y con vistas a mejor posicionarse en una posterior discusión, es preciso distinguir entre evaluación, información y *feedback*. La evaluación educativa es un proceso o actividad realizada por el profesor –y, en su caso, por el alumno– que proporciona información²². Pero, como ha advertido Ramaprasad (1983) en un contexto temático más general, «La información [...] es *feedback* únicamente cuando es utilizada para alterar la brecha existente ente el nivel actual y el nivel de referencia»²³. Desde esta perspectiva, la evaluación formativa solo será eficaz en su orientación hacia la mejora cuando vaya acompañada de una acción de retroalimentación efectiva sobre el objeto de la evaluación, esto es, por un *feedback* eficaz. De ahí que lo importante, en orden a promover la mejora educativa, no sea tanto la información que produce la evaluación como lo que se hace con esa información; es decir, las acciones de cambio que, de un modo deliberado y planificado, a partir de ella se desarrollen.

2.3 La evaluación formativa en diferentes niveles del sistema educativo

8

La naturaleza transversal de los mecanismos del *feedback* permite postular una 'invariancia de escala' para dicho fenómeno y sostener su validez en los diferentes niveles del sistema educativo: el *micro* o propio del aula, el *meso* o referido al centro escolar y el *macro* que concierne al sistema en su conjunto. En este sentido, y habida cuenta del papel central de dichos mecanismos en la conceptualización de la evaluación formativa, cabe aplicar esta noción a cada uno de ellos con un significado genérico similar, esto es, como una evaluación para la mejora.

2.3.1 En el nivel *micro*: el aula

En un sentido estricto, y de acuerdo con su significado original, la evaluación formativa se refiere a la evaluación efectuada sobre ese nivel micro del sistema definido por el espacio del aula. Según Bloom *et al.* (1971), la evaluación formativa concierne «al proceso de elaboración del currículo, a la enseñanza y al aprendizaje, y su propósito es la mejora de cualquiera de esos tres procesos»²⁴. Afecta, pues, al profesor, al alumno y a su interacción; y, a diferencia de la evaluación sumativa o de la evaluación

20 HATTIE, J. A.; TIMPERLEY, H (2007). *Op. cit.*

21 HAVNES, A.; SMITH, K.; DYSTHE, O.; LUDVIGSEN, K. (2012). «Formative assessment and feedback: Making learning visible». *Studies in Educational Evaluation*, 38, pp. 21-27.

22 BLACK, P.; WILLIAM, D. (1998). «Inside the Black Box: Raising Standards Through Classroom Assessment». *Phi Delta Kappa*, oct 1998, pp. 1-18.

23 RAMAPRASAD, A. (1983). «On the definition of feedback». *Behavioural Science*, 28, 4-13.

24 Bloom, B. S.; Madaus, G. F.; Hastings, J. T. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill.

de certificación, la evaluación formativa está expresamente orientada a la mejora, de ahí su conexión natural con el concepto de *feedback* y con la aplicación práctica de este concepto en el ámbito de la educación²⁵.

Existe una abundante evidencia empírica sobre la fuerte influencia que tiene el *feedback* sobre la eficacia de los aprendizajes^{26, 27, 28}; pero quizás sea el laborioso programa de investigación de John Hattie el que haya aportado los fundamentos más sólidos al respecto^{29, 30}. De acuerdo con las aportaciones de dicho autor, el *feedback* –propio de la evaluación formativa– es uno de los componentes de la enseñanza que mayor impacto tiene sobre los resultados de los alumnos. A partir de una síntesis metanalítica, encuentra un valor del tamaño del efecto $d = 0,90$, superior al obtenido en el trabajo precedente de Fuchs *et al.* ($d = 0,70$)³¹. Aunque existen en la literatura investigaciones que arrojan resultados negativos para el *feedback* como componente de la enseñanza³², ello podría ser debido a un escaso grado de pertinencia de las acciones que comporta el *feedback*, a un nivel de calidad insuficiente o, incluso, a una confusión conceptual entre *feedback* e información³³, cuyas diferencias han sido descritas más arriba.

2.3.2 En el nivel *meso*: el centro

Desde una aproximación general, el centro educativo puede ser considerado como un tipo específico de organización y, por tanto, es de aplicación a este caso la conceptualización del ciclo de mejora continua de las organizaciones conocido como 'ciclo de Deming'³⁴; en esencia, supone una aplicación, al ámbito de las organizaciones, de las ideas de *feedback* y autoregulación³⁵. Justamente por ello, aquellas escuelas que aplican, al conjunto de la organización, una evaluación formativa –entendida en su significado más amplio– pueden ser consideradas como 'organizaciones inteligentes' que son capaces de corregir los errores y aprender de la experiencia.

Las escuelas inteligentes son, de hecho, 'sistemas adaptativos'. Como hemos señalado en otro lugar, «Los sistemas adaptativos no responden pasivamente a los acontecimientos, a las tensiones internas o a las presiones del entorno, sino que desarrollan estrategias adaptativas que orientan la

25 HATTIE, J. (2017). '*Aprendizaje visible*' para profesores. *Maximizando el impacto del aprendizaje*. Cap.7. El flujo de la clase: el lugar del *feedback*. Paraninfo. Madrid.

26 SCHILDKAMP, K.; LAI, M. K. (2013). «Data-based Decision Making: An Overview». En Schildkamp K., Lai M., Earl L. (eds) *Data-based Decision Making in Education. Studies in Educational Leadership*, vol 17. Springer, Dordrecht. <https://doi.org/10.1007/978-94-007-4816-3_2>.

27 BAIRD J. A. *et al.* (2014). *Op. cit.*

28 FUCHS, L. S.; FUCHS, D. (1986). «Effects of systematic formative evaluation. A meta-analysis». *Exceptional Children*, 53(3), 199-208.

29 HATTIE, J. A.; TIMPERLEY, H. (2007). *Op. cit.*

30 HATTIE, J. A. (2009). *Visible Learning. A Synthesis of Over 800 Meta-analyses Relating to Achievement*. Routledge. Abington.

31 FUCHS, L. S.; FUCHS, D. (1986). *Op. cit.*

32 KLUGER, A. N.; DENISI, A. (1996). «The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory». *Psychological Bulletin*, 119, 254-284.

33 HAVNES, A.; SMITH, K.; DYSTHE, O.; LUDVIGSEN, K. (2012). *Op. cit.*

34 LÓPEZ RUPÉREZ, F. (1994). *La gestión de calidad en educación*. Ed. La Muralla. Madrid.

35 RICHARDSON, G.P. (1991). *Op. cit.*

actividad del sistema y dan lugar a una evolución de su organización que les permite sacar provecho de las situaciones nuevas»³⁶.

En un contexto social y económico tan dinámico como el del siglo XXI, fuertemente vinculado al impacto de la revolución digital y, en un sentido más amplio, al de la llamada 'cuarta revolución industrial'³⁷ —que concierne al desarrollo de las tecnologías basadas en sistemas físicos, en sistemas biológicos, en sistemas digitales y en sus interacciones—, la respuesta adaptativa de los centros educativos está en el núcleo mismo de la innovación y de la mejora. Y los mecanismos del *feedback* y de la autorregulación son los que, de un modo ordenado y sistemático, hacen aquéllas posibles. Como señala Hattie (2017), citando el trabajo de Ben Levin (2008), «[...] la clave de la mejora educativa reside en mejorar las prácticas de enseñanza y aprendizaje diarios en las escuelas, basadas en la noción de que la escuela es la unidad apropiada de evaluación»³⁸.

Por todo ello, la evaluación interna —o autoevaluación— de los centros constituye uno de los instrumentos estratégicos para la mejora, que ha de formar parte de sus proyectos educativos, contando con el debido apoyo técnico para que dicho instrumento sea efectivo. Su práctica sistemática contribuye a generar un cambio cultural en los centros favorable al uso de la evaluación y a la importancia del *feedback* formativo. Dicho cambio cultural introduce a las escuelas en un círculo virtuoso que les hace organizaciones inteligentes, alimenta la innovación e induce la mejora escolar³⁹.

10

Otro de los mecanismos asociados al *feedback* formativo en este nivel *meso*, propio de los centros, es el de la rendición de cuentas por los resultados, mecanismo éste que opera sobre los comportamientos y que se ha revelado imprescindible para que la autonomía escolar produzca efectos positivos, o de mejora, sobre el rendimiento de los alumnos⁴⁰. Por tal motivo, la evaluación interna ha de ser complementada con una evaluación externa que genere, en cada escuela, este tipo de '*feedback* de responsabilidad' en la administración de una mayor libertad. Ésta es una importante precaución a adoptar pues sin ella una mayor autonomía de los centros puede resultar contraproducente, al dar pie al desarrollo de comportamientos oportunistas ajenos a los intereses generales. Investigaciones relativamente recientes muestran que la autonomía afecta negativamente a los resultados de los alumnos en países en desarrollo y de bajo rendimiento y, sin embargo, lo hacen positivamente en países desarrollados y de alto rendimiento la mayor parte de los cuales disponen de sistemas de rendición de cuentas⁴¹.

36 LÓPEZ RUPÉREZ, F. (2001). *Preparar el futuro. La educación ante los desafíos de la globalización*. Cap. 6. Políticas educativas para el siglo XXI. Ed. La Muralla. Madrid.

37 SHWAB, K. (2016). <www.weforum.org/es/agenda/2016/10/cuatro-principios-de-liderazgo-de-la-cuarta-revolucion-industrial>

38 HATTIE, J. (2017). *Op. cit.*

39 ELMORE, R., F. (2004). *School reform from inside out: Policy, practice, and performance*. Harvard Education Press. Cambridge, MA.

40 WOESSMANN, L.; LUEDEMANN, E.; SCHUETZ, G.; WEST, M. R. (2009). *School Accountability, Autonomy, and Choice around the World*. Edward Elgar. Cheltenham.

41 HANUSHEK, E. A.; LINK, S.; WOESSMANN, L. (2012). *Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA*. ADB Economics Working Paper Series No. 296. Asian Development Bank.

2.3.3 En el nivel *macro*: el sistema

La evaluación formativa en el nivel *macro*, propio del sistema educativo en su conjunto, se apoya en los mismos mecanismos más arriba descritos con carácter general. No obstante, al tratarse en este caso de un elemento básico para una gobernanza de calidad, las acciones generadoras del *feedback* se escapan del ámbito del profesorado o de la dirección escolar para alcanzar a los responsables de las políticas educativas, de su concepción y de su implementación. Corresponde, pues, a éstos definir las metas, las estrategias y los planes de actuación para la mejora, a la luz de la información facilitada por la evaluación a gran escala.

Las evaluaciones del conjunto del sistema pueden ser, por su extensión, muestrales o censales. Las primeras suministran una información estadísticamente representativa del comportamiento del conjunto y, por tanto, solo sirven para orientar las políticas y producir *feedbacks* efectivos en el ámbito territorial en el que la muestra empleada posea esa representatividad. Tal es el caso de evaluaciones internacionales como las citadas: TIMMS, PIRLS o PISA. Particularmente en este último caso, el empleo de muestras representativas a nivel nacional, por un lado, y a nivel de regional, por otro, permite extraer consecuencias en esos dos ámbitos territoriales que resultan de interés en todos los supuestos y, en especial, cuando las regiones –como es el caso español con sus Comunidades Autónomas– disponen de amplias competencias en materia educativa.

Las segundas, al alcanzar a todos y cada uno de los alumnos, proporcionan una información exhaustiva que puede ser agrupada por centros, por distritos, por provincias, etc. A esta exhaustividad en la información le corresponde una exhaustividad potencial en los *feedbacks*, que pueden operar con un valor formativo, o de mejora, extraordinariamente preciso y efectivo. Pero, además, permite la rendición de cuentas en el nivel deseado –aula, centro, distrito, región, etc.– en función de los criterios de oportunidad que se establezcan. La acción formativa de este tipo de evaluaciones censales está sometida a criterios de proximidad y corresponde, por ello, su implementación a la instancia profesional, administrativa o política correspondiente, que disponga de la competencia para aplicar aquellos *feedbacks* que, en cada caso, proceda.

Existe un consenso amplio en el ámbito internacional sobre la importancia de la evaluación general de los sistemas educativos como elemento de calidad y factor de mejora. Así, el Gobierno federal alemán, ante el *shock* generado por los datos mediocres obtenidos en la primera edición de PISA (2000) impulsó, en colaboración con la OCDE, la realización de un análisis comparado de los seis países occidentales que mejores resultados obtuvieron en esas pruebas –Canadá, Finlandia, Francia, Reino Unido y Holanda– con el fin de identificar buenas prácticas en materia de definición e implementación de políticas educativas centradas en las escuelas⁴².

Uno de los rasgos comunes más relevantes que surgió de los análisis comparados fue la importancia que todos los países considerados, sin excepción, concedían a los resultados. Ello se reflejaba en dos tipos de políticas:

42 OECD (2004). *What Makes Schools Systems Perform? Seeing School Systems through the Prism of PISA*. OECD Publishing. Paris.

- La definición y aplicación de estándares de rendimiento –o soluciones análogas– como instrumentos normativos que permiten establecer, sin ambigüedad, los logros que deben alcanzar los alumnos al finalizar determinadas etapas educativas.
- La definición y aplicación de sistemas de seguimiento cuantitativo o de control (*monitoring*) para el pilotaje de los respectivos sistemas escolares.

Según el citado estudio, aun cuando los enfoques particulares pudieran diferir de unos países a otros, todos ellos consideraban el seguimiento y la comparación de sus resultados, en términos tanto nacionales como internacionales, como una parte fundamental de su estrategia para asegurar rendimientos escolares elevados.

3. Un modelo integrado de evaluación para el sistema educativo español

A modo de preámbulo a la exposición del modelo de evaluación y de sus bases conceptuales, que se proponen en este trabajo, procede efectuar un recorrido histórico por las diferentes concepciones de la evaluación general del sistema educativo español reflejadas en la secuencia de leyes de educación de la democracia en las que aquéllas se han establecido.

12

3.1 Evolución normativa de la evaluación del sistema educativo en España

Aun cuando la Ley General de Educación (1970)⁴³ contenía una referencia a una valoración del rendimiento de los centros (art. 11. Cinco), no puede afirmarse que dicha ley contemplara un modelo de evaluación general del sistema educativo. La evaluación periódica del rendimiento educativo de los centros se vinculaba, particularmente, a la modificación, en su caso, de la clasificación de los centros, de titularidad privada, de Bachillerato y de Formación Profesional de Segundo Grado –en Libres, Habilitados y Homologados– en función de los resultados obtenidos en dicha evaluación (art. 95. Dos). Es a partir de 1990, con la Ley orgánica de Ordenación General del Sistema Educativo (LOGSE) cuando la educación española comienza a conceder una importancia creciente a la evaluación del sistema en su conjunto. En lo que sigue, se efectuará una revisión histórica de los diferentes modelos de evaluación del sistema educativo español a partir del análisis comparado de las sucesivas leyes orgánicas.

3.1.1 LOGSE

La LOGSE⁴⁴ –en su título 4, artículo 55, apartado g)– sitúa de un modo explícito, por primera vez, la evaluación del sistema educativo como uno de los seis pilares de la calidad de la enseñanza.

43 <<https://www.boe.es/boe/dias/1970/08/06/pdfs/A12525-12546.pdf>>

44 Ley Orgánica 1/1990, de 3 de octubre, de Ordenación General del Sistema Educativo. <BOE-A-1990-24172 >.

En coherencia con lo anterior, en su preámbulo se describe con más detalle la intención de la Ley al respecto en los siguientes términos:

«(...) La ley atribuye una singular importancia a la Evaluación general del sistema educativo, creando para ello el Instituto Nacional de Calidad y Evaluación. La actividad evaluadora es fundamental para analizar en qué medida los distintos elementos del sistema educativo están contribuyendo a la consecución de los objetivos previamente establecidos. Por ello, ha de extenderse a la actividad educativa en todos sus niveles, alcanzando a todos los sectores que en ella participan. Con una estructura descentralizada, en la que los distintos ámbitos territoriales gozan de una importante autonomía, es aún más fundamental contar con un instrumento que sirva para reconstruir una visión de conjunto y para proporcionar a todas y cada una de las instancias la información relevante y el apoyo preciso para el mejor ejercicio de sus funciones. En coherencia con ello, el Instituto Nacional de Calidad y Evaluación contará con la participación de las Comunidades Autónomas».

Aunque, de un modo escueto, se hace referencia en el citado preámbulo al significado conceptual de la evaluación del sistema –en tanto que instrumento generador de información para el *feedback*–, a su sentido práctico en un sistema educativo descentralizado, a una orientación multinivel dirigida 'a cada una de las instancias', y a la creación de una institución específica: el Instituto Nacional de Calidad y Evaluación.

La tabla 1 muestra, de un modo sintético, los rasgos característicos del modelo de evaluación del sistema educativo español introducido en la LOGSE, de acuerdo con una aproximación multinivel que facilitará, en lo sucesivo, un análisis comparado, o de evolución con el tiempo, de los diferentes componentes del modelo. En dicha tabla, por su importancia como instrumento para una evaluación descriptiva, se ha añadido la creación de un sistema estatal de indicadores, pues aun cuando no figura expresamente en la Ley sí lo hace en su desarrollo, mediante el Real Decreto 928/1993, de 18 de junio, por el que se regula el Instituto Nacional de Calidad y Evaluación.

Se trata, en su conjunto, de un modelo inicial poco definido, en el que no se especifica la naturaleza muestral o censal de los procesos de evaluación considerados, ni si las evaluaciones de los centros son o no prescriptivas, internas o externas. A la luz del texto legal, parece más bien que se pretendiera con él hacer arrancar el sistema dejando para más adelante una mayor concreción.

3.1.2 LOPEGCE

La Ley Orgánica 9/1995, de 20 de noviembre, de la Participación, la Evaluación y el Gobierno de los Centros Docentes (LOPEGCE)⁴⁵, avanzó un poco más hacia esa mayor concreción esperada respecto del modelo de evaluación y complementó, en cierta medida, el anterior. La tabla 2 recoge, de acuerdo con el mismo patrón de análisis de la tabla precedente, las aportaciones de la nueva ley educativa a

⁴⁵ Ley Orgánica 9/1995, de 20 de noviembre, de la participación, la evaluación y el gobierno de los centros docentes. <BOE-A-1995-25202>.

Tabla 1
Rasgos característicos del modelo de evaluación del sistema educativo español en la LOGSE, de acuerdo con una aproximación multinivel

Órgano competente	Ámbito territorial		
	Centro educativo	Comunidad Autónoma	Conjunto del Estado
Centro educativo	*	*	*
Administración educativa	*	– Evaluación del sistema educativo en el ámbito de sus competencias. (Art. 62.1).	*
Ministerio de Educación	– Elaboración de sistemas de evaluación para las diferentes enseñanzas y los correspondientes centros. (Art. 62.4.a).	*	– Evaluaciones del sistema educativo. (Art. 62.4.c). – Sistema estatal de indicadores. (RD 928/1993).

14

Fuente: elaboración propia.

la LOGSE en materia de evaluación. Así, en la tabla 2, se rellena la casilla correspondiente a las evaluaciones internas de los centros educativos, que han de obedecer a lo preceptuado por su correspondiente Administración educativa competente. Y también se ocupa, en dicha tabla, la casilla de las evaluaciones externas realizadas sobre los centros docentes por las Administraciones educativas mediante planes periódicos. Además, se prevé el apoyo, por parte del Estado, a las Administraciones educativas para la elaboración de sus propios planes de evaluación.

3.1.3 LOCE

La Ley Orgánica 10/2002, de 23 de diciembre de Calidad de la Educación (LOCE)⁴⁶, supuso un avance en cuanto a la concreción del modelo y su consolidación. Así, se advierte en el preámbulo de la referida Ley una especificación mucho más clara y precisa de la finalidad de la evaluación del sistema, de sus diferentes ámbitos, de su justificación y de sus mecanismos. Cabe, en este sentido, destacar los párrafos siguientes:

46 Ley Orgánica 10/2002, de 23 de diciembre, de Calidad de la Educación. < BOE-A-2002-25037 >.

Tabla 2
Rasgos característicos del modelo de evaluación del sistema educativo español en la LOPEGCE, de acuerdo con una aproximación multinivel

Órgano competente	Ámbito territorial		
	Centro educativo	Comunidad Autónoma	Conjunto del Estado
Centro educativo	<ul style="list-style-type: none"> Los centros evaluarán su propio funcionamiento al final de cada curso de acuerdo con lo preceptuado por su Administración educativa. (Art. 29.3). 	*	*
Administración educativa	<ul style="list-style-type: none"> Planes de evaluación externa periódicos de los centros docentes sostenidos con fondos públicos. (Art. 29.1). 	*	*
Ministerio de Educación	*	*	<ul style="list-style-type: none"> Evaluación general del sistema educativo. (Art. 28.1). Apoyo a las Administraciones educativas para la elaboración de sus planes de evaluación. (Art. 28.2).

Fuente: elaboración propia.

«[...] El segundo eje de medidas de la Ley consiste en orientar más abiertamente el sistema educativo hacia los resultados, pues la consolidación de la cultura del esfuerzo y la mejora de la calidad están vinculadas a la intensificación de los procesos de evaluación de los alumnos, de los profesores, de los centros y del sistema en su conjunto, de modo que unos y otros puedan orientar convenientemente los procesos de mejora. Esta acentuación de la importancia de los resultados no supone, en modo alguno, ignorar el papel de los procesos que conducen a aquéllos, ni de los recursos en los que unos y otros se apoyan. La evaluación, es decir, la identificación de los errores y de los aciertos no sólo es un factor básico de calidad; constituye, además, un instrumento ineludible para hacer inteligentes políticas educativas a todos los niveles y para incrementar, progresivamente, su oportunidad y su adecuación a los cambios».

«[...] El refuerzo de la autonomía de los centros se basa, igualmente, en la confianza mutua y en la responsabilidad; en el acuerdo entre centro y Administración, que deben considerarse como socios principales en la tarea de hacer avanzar la educación en el plano local; y, a la vez, en la necesidad de responder de los resultados mediante procedimientos de evaluación que

Tabla 3
Rasgos característicos del modelo de evaluación del sistema educativo español en la LOCE, de acuerdo con una aproximación multinivel

Órgano competente	Ámbito territorial		
	Centro educativo	Comunidad Autónoma	Conjunto del Estado
Centro educativo	<ul style="list-style-type: none"> Los centros evaluarán su propio funcionamiento al final de cada curso, de acuerdo con lo preceptuado por su Administración educativa. (Art. 101.4). 	*	*
Administración educativa	<ul style="list-style-type: none"> Planes de evaluación externa periódicos de los centros docentes sostenidos con fondos públicos. (Art. 101.1). Apoyo a los centros para resolver los problemas detectados. (Art.101.3). 	<ul style="list-style-type: none"> Evaluación que las Administraciones educativas de las Comunidades Autónomas puedan realizar de acuerdo con sus competencias. (Art. 96.1). 	*
Ministerio de Educación	*	*	<ul style="list-style-type: none"> Plan de Evaluación general del sistema educativo periódico (Art. 99) Evaluaciones generales de diagnóstico en Primaria y ESO de las competencias básicas de esos niveles educativos. (Art. 17, 30 y 97) Sistema estatal de indicadores (Art. 98).

16

Fuente: elaboración propia.

faciliten la mejora y permitan orientar y modular las acciones conjuntas de cada centro educativo y de cada Administración».

«[...] El Título VI está referido a la evaluación del sistema educativo que, en su dimensión general, se sitúa en el ámbito de las competencias estatales, sin perjuicio de las competencias y obligaciones que en esta materia corresponde a las Administraciones educativas en sus respectivos ámbitos territoriales».

Del análisis comparativo entre las tablas 1, 2 y 3 se advierte la progresividad de los planteamientos de la nueva Ley. Así, se profundiza en los logros de las leyes anteriores en cuanto al modelo de evaluación

general del sistema, se introduce un título propio –el Título VI– para la evaluación del sistema educativo y se aporta una mayor concreción, como corresponde todo ello a esa importancia que la LOCE confiere a los resultados de los alumnos como factor decisivo de calidad del sistema educativo. Se integra la evaluación general del sistema en el marco coherente de un Plan periódico y se introducen, por primera vez, las evaluaciones generales de diagnóstico en Primaria y en la ESO, sobre el grado de adquisición de las competencias básicas propias de cada nivel; preservándose, en todo caso, los ámbitos competenciales del Estado y de las Administraciones educativas referidos en las leyes anteriores.

3.1.4 LOE

La Ley Orgánica 2/2006, de 3 de mayo, de Educación (LOE)⁴⁷ prosiguió esa senda progresiva de consolidación y afinamiento que se ha destacado en los análisis anteriores. En el ámbito de la evaluación del sistema se aprecia, con respecto a la LOCE, incluso una cierta continuidad en la justificación de sus fundamentos, tal y como parece deducirse de la lectura de su preámbulo:

«[...] La existencia de un marco legislativo capaz de combinar objetivos y normas comunes con la necesaria autonomía pedagógica y de gestión de los centros docentes obliga, por otra parte, a establecer mecanismos de evaluación y de rendición de cuentas. La importancia de los desafíos que afronta el sistema educativo demanda como contrapartida una información pública y transparente acerca del uso que se hace de los medios y los recursos puestos a su disposición, así como una valoración de los resultados que con ellos se alcanzan. La evaluación se ha convertido en un valioso instrumento de seguimiento y de valoración de los resultados obtenidos y de mejora de los procesos que permiten obtenerlos. Por ese motivo, resulta imprescindible establecer procedimientos de evaluación de los distintos ámbitos y agentes de la actividad educativa, alumnado, profesorado, centros, currículo, Administraciones, y comprometer a las autoridades correspondientes a rendir cuentas de la situación existente y el desarrollo experimentado en materia de educación».

«[...] Las enseñanzas que tienen carácter obligatorio son la educación primaria y la educación secundaria obligatoria. En la etapa primaria se pone el énfasis en la atención a la diversidad del alumnado y en la prevención de las dificultades de aprendizaje, actuando tan pronto como éstas se detecten. Una de las novedades de la Ley consiste en la realización de una evaluación de diagnóstico de las competencias básicas alcanzadas por el alumnado al finalizar el segundo ciclo de esta etapa, que tendrá carácter formativo y orientador, proporcionará información sobre la situación del alumnado, de los centros y del propio sistema educativo y permitirá adoptar las medidas pertinentes para mejorar las posibles deficiencias. Otra evaluación similar se llevará a cabo al finalizar el segundo curso de la educación secundaria obligatoria».

La tabla 4 recoge lo esencial de las aportaciones de la LOE. Al comparar su contenido con el de la tabla 3, referida a la LOCE, se advierte la más arriba citada continuidad que va acompañada de un

⁴⁷ Ley Orgánica 2/2006, de 3 de mayo, de Educación. < BOE-A-2006-7899 >.

Tabla 4
Rasgos característicos del modelo de evaluación del sistema educativo español en la LOE, de acuerdo con una aproximación multinivel

Órgano competente	Ámbito territorial		
	Centro educativo	Comunidad Autónoma	Conjunto del Estado
Centro educativo	<ul style="list-style-type: none"> – Autoevaluación de los centros educativos. (Art. 145.2). 	*	*
Administración educativa	<ul style="list-style-type: none"> – Planes de evaluación de los centros educativos. (Art. 145.1). – Apoyo a la autoevaluación de los centros educativos. (Art. 145.2). – Evaluaciones de diagnóstico de las competencias básicas al finalizar el segundo ciclo de Primaria, con carácter formativo y orientador para todos los centros, y teniendo como referencia las evaluaciones generales de diagnóstico. (Art. 21 y 144.1). – Evaluaciones de diagnóstico análogas a las anteriores al finalizar el segundo de la ESO (Art. 29 y 144.1). – Estas evaluaciones de diagnóstico se realizarán con criterios de homogeneidad en todo el Estado. (Art. 144.1). 	<ul style="list-style-type: none"> – Evaluaciones de diagnóstico de las competencias básicas al finalizar el segundo ciclo de Primaria, con carácter formativo, interno y orientador para los centros, y teniendo como referencia las evaluaciones generales de diagnóstico. (Art. 21 y 144.1, y 144.2). – Evaluaciones de diagnóstico análogas a las anteriores al finalizar el segundo de la ESO. (Art. 29, 144.1 y 144.2). – Estas evaluaciones de diagnóstico se realizarán con criterios de homogeneidad (Art. 144.1) y bajo el control y desarrollo de las Administraciones educativas (Art. 144.2). 	*
Ministerio de Educación	<ul style="list-style-type: none"> – Evaluaciones generales de diagnóstico con datos representativos del alumnado y de los centros sobre competencias básicas en Primaria y ESO. (Art. 144.1). 	<ul style="list-style-type: none"> – Evaluaciones generales de diagnóstico con datos representativos del alumnado y de los centros de cada Comunidad Autónoma sobre competencias básicas en Primaria y ESO. (Art. 144.1). 	<ul style="list-style-type: none"> – Evaluaciones generales de diagnóstico con datos representativos del alumnado y de los centros del conjunto del Estado sobre competencias básicas en Primaria y ESO. (Art. 144.1). – Sistema estatal de indicadores (Art. 143.3).

Fuente: elaboración propia.

afinamiento que le añade precisión. Así, por ejemplo, a propósito de las evaluaciones generales de diagnóstico se especifica su naturaleza muestral, al aludir a la representatividad de los datos; se fija el ámbito multinivel al referirse expresamente al alumnado, a los centros, a las Comunidades Autónomas y al conjunto del Estado; se precisa el momento de las dos evaluaciones de diagnóstico: al finalizar el segundo ciclo de Educación Primaria, por un lado, y el segundo curso de la ESO, por otro; se establece el enfoque formativo e interno de estas evaluaciones para los centros, bajo el control y desarrollo de las Administraciones educativas correspondientes; y se determina su carácter censal, cuando menos, en el nivel de los centros educativos.

Conforme el modelo de evaluación del sistema ha ido desarrollándose, ha ido ganado en precisión y también en robustez, con el riesgo de caer en una redundancia poco productiva, en una inflación del proceso evaluador. Sobre esta circunstancia volveremos más adelante con ocasión de la propuesta del modelo integrado que culminará el desarrollo del presente trabajo.

3.2 La situación actual

La situación presente es tributaria de dos componentes que se abordarán en lo que sigue de forma separada: por un lado, la contribución al modelo actual de la Ley Orgánica 8/2013, de 9 de diciembre, para la Mejora de la Calidad Educativa (LOMCE); y, por otro, la procedente de la modificación transitoria introducida en la propia Ley por el Real Decreto-ley 5/2016, de 9 de diciembre, de medidas urgentes para la ampliación del calendario de implantación de la Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa.

3.2.1 LOMCE

Como se reconoce en el preámbulo de la LOMCE⁴⁸, una de las mayores aportaciones de la Ley en relación con la LOE radica en las novedades que se introducen en el modelo de evaluación general del sistema:

«[...] Las evaluaciones externas de fin de etapa constituyen una de las principales novedades de la LOMCE con respecto al marco anterior y una de las medidas llamadas a mejorar de manera más directa la calidad del sistema educativo. Veinte países de la OCDE realizan a sus alumnos y alumnas pruebas de esta naturaleza y las evidencias indican que su implantación tiene un impacto de al menos dieciséis puntos de mejora de acuerdo con los criterios de PISA».

Más adelante, y refiriéndose a dichas pruebas, el citado preámbulo especifica:

«[...] Estas pruebas tendrán un carácter formativo y de diagnóstico. Por un lado, deben servir para garantizar que todos los alumnos y alumnas alcancen los niveles de aprendizaje adecuados para el normal desenvolvimiento de la vida personal y profesional conforme el título

48 Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. <BOE-A-2013-12886 >.

pretendido, y además deben permitir orientar a los alumnos y alumnas en sus decisiones escolares de acuerdo con los conocimientos y competencias que realmente posean. Por otro lado, estas pruebas normalizan los estándares de titulación en toda España, indicando de forma clara al conjunto de la comunidad educativa cuáles son los niveles de exigencia requeridos e introduciendo elementos de certeza, objetividad y comparabilidad de resultados. Además, proporcionan a los padres, a los centros y a las Administraciones educativas una valiosa información de cara a futuras decisiones. El objetivo de esta evaluación es la mejora del aprendizaje del alumno o alumna, de las medidas de gestión de los centros y de las políticas de las Administraciones».

La tabla 5 resume lo esencial del sistema de evaluación general contenido en la LOMCE. Así, la comparación con el contenido de la tabla 4 permite identificar aquello que se conserva del modelo anterior; esto es, la autoevaluación de los centros educativos con el apoyo de las Administraciones educativas; planes de evaluación de los centros por parte de las Administraciones educativas, en el marco de sus competencias; planes plurianuales de evaluación general del sistema educativo con carácter periódico y a cargo del Instituto Nacional de Evaluación Educativa. Asimismo, dicha comparación hace posible distinguir lo que se introduce como novedad: dos evaluaciones individualizadas en Educación Primaria –una en tercero y otra de final de etapa en sexto– sin valor académico, otras dos de final de etapa en ESO y en Bachillerato, con valor académico.

20

Se advierte en el modelo de la LOMCE los siguientes inconvenientes:

- Una cierta redundancia en el nivel de centro entre los planes de autoevaluación –que constituyen un residuo de la LOE cuyo correspondiente artículo no se ve modificado en la LOMCE– y la evaluación individualizada en tercero de Educación Primaria, toda vez que, según el enfoque adoptado por la Ley, se tratan ambas de evaluaciones internas de los centros y de finalidad formativa.
- Una reiteración de evaluaciones individualizadas y de carácter censal en la Educación Primaria.
- Una cierta confusión conceptual entre las evaluaciones formativas y las 'evaluaciones de certificación' que son aquellas conducentes a un diploma o a un título⁴⁹.
- La asignación de efectos académicos a la evaluación final individualizada de la ESO que, al tratarse de una etapa obligatoria, constituye una anomalía en el panorama europeo. La OCDE hace tiempo que ha hecho desaparecer de su sistema internacional de indicadores de la educación las tasas de graduación en educación secundaria inferior⁵⁰.
- Una cierta ambigüedad en cuanto al reparto competencial entre el Gobierno y las Administraciones educativas de las Comunidades Autónomas.
- Una franca discontinuidad con los modelos anteriores que introduce una ruptura en las series cronológicas de los datos correspondientes.

49 DE LANDSHEERE, G. (1996). *El pilotaje de los sistemas educativos*. Ed. La Muralla. Madrid.

50 OECD (2017). *Education at a Glance. OECD Indicators*. OECD Publishing. Paris.

Tabla 5
Rasgos característicos del modelo de evaluación del sistema educativo español en la LOMCE, de acuerdo con una aproximación multinivel

Órgano competente	Ámbito territorial		
	Centro educativo	Comunidad Autónoma	Conjunto del Estado
Centro educativo	<ul style="list-style-type: none"> – Autoevaluación de los centros educativos con el apoyo de las Administraciones educativas. (Art. 145). 	*	*
Administración educativa	<ul style="list-style-type: none"> – Planes de evaluación de los centros educativos en el marco de sus competencias. (Art. 145.1). 	<ul style="list-style-type: none"> – Evaluación individualizada a todos los alumnos al finalizar sexto de Primaria con criterios y características comunes en todo el sistema educativo español, con planes de mejora potestativos. (Art. 21.1 y 3). – Otras evaluaciones con fines de diagnóstico. (Art. 144.2). 	*
Ministerio de Educación	*	*	<ul style="list-style-type: none"> – Planes plurianuales de evaluación general del sistema educativo periódicamente. (Art. 143.1). – Sistema estatal de indicadores de la Educación (Art. 143.) – Evaluaciones finales individualizadas en ESO y Bachillerato (Art. 29, 36 bis y 144.1).

Fuente: elaboración propia.

3.2.2 La rectificación del Real Decreto-ley

El Real Decreto-ley más arriba citado⁵¹, además de modificar algunos aspectos del calendario de implantación de la LOMCE, introdujo un cambio sustantivo en su modelo de evaluación general, al derogar la regulación de las evaluaciones finales individualizadas y sustituirla transitoriamente por otra,

51 Real Decreto-ley 5/2016, de 9 de diciembre, de medidas urgentes para la ampliación del calendario de implantación de la Ley Orgánica 8/2013, de 9 de diciembre, para la mejora de la calidad educativa. < BOE-A-2016-11733 >.

hasta la entrada en vigor de la normativa resultante del Pacto de Estado Social y Político por la Educación.

Así, en lo que respecta a la evaluación final de la ESO, se establece literalmente lo siguiente:

«[...] a) La evaluación no tendrá efectos académicos.

b) La evaluación será considerada muestral y tendrá una finalidad diagnóstica, participando únicamente en ella el alumnado matriculado en cuarto curso que haya sido seleccionado por la Administración educativa, exclusivamente por la opción cursada. La selección de alumnos y centros será suficiente para obtener datos representativos. Las Administraciones Educativas podrán elevar el número de centros participantes por encima de las necesidades muestrales o hacerla con carácter censal.

c) Evaluará el grado de adquisición de la competencia matemática, la competencia lingüística, la competencia social y cívica y tendrá como referencia principal las materias generales del bloque de las asignaturas troncales cursadas en cuarto de Educación Secundaria Obligatoria».

En lo relativo a la prueba final de Bachillerato, se establece literalmente lo siguiente:

«[...] a) La evaluación únicamente se tendrá en cuenta para el acceso a la Universidad, pero su superación no será necesaria para obtener el título de Bachiller.

b) Podrá presentarse a la evaluación el alumnado que esté en posesión del título de Bachiller, así como los alumnos que se encuentren en alguna de las situaciones contempladas en la disposición adicional tercera del Real Decreto 310/2016, de 29 de julio.

c) Las administraciones educativas, en colaboración con las Universidades, que asumirán las mismas funciones y responsabilidades que tenían en relación con las Pruebas de Acceso a la Universidad, organizarán la realización material de la evaluación final de Bachillerato para el acceso a la Universidad. No obstante, cada administración educativa podrá delimitar el alcance de la colaboración de sus universidades en la realización de la prueba. Dicha evaluación tendrá validez para el acceso a las distintas titulaciones de las universidades españolas.

d) La adquisición de las competencias se evaluará a través de las materias generales cursadas del bloque de las asignaturas troncales de segundo curso y, en su caso, de la materia Lengua Cooficial y Literatura. Los alumnos que quieran mejorar su nota de admisión podrán examinarse de, al menos, dos materias de opción del bloque de las asignaturas troncales de segundo curso».

Esta rectificación de la Ley —severa a la luz de la importancia atribuida en su preámbulo a este tipo de evaluaciones externas— es de naturaleza política y, debido a su provisionalidad, no procede efectuar aquí un análisis rigurosamente técnico de la normativa correspondiente. Cabe, sin embargo, reflexionar en lo que sigue sobre un modelo de evaluación del sistema que integre, de un modo coherente, el enfoque multinivel; que sea pertinente con relación a la finalidad formativa que se pretende; que no eluda la función diagnóstica; que sea respetuoso con el reparto competencial de las instancias

intervinientes; y que sea económico, no sólo en el sentido de eficiente sino también en el de evitar evaluaciones redundantes que además de encarecer el modelo, lo devalúen.

3.3 Bases para un modelo multinivel integrado de evaluación

La propuesta de un modelo multinivel integrado para la evaluación del sistema educativo español, objeto último del presente trabajo, reposa en un conjunto de criterios, de bases o de opciones conceptuales y normativas que supone, de hecho, una justificación por adelantado del modelo que se especificará después.

3.3.1 Un enfoque formativo

Como ha sido mencionado en el anterior análisis conceptual, la finalidad formativa de la evaluación educativa es una componente estratégica de cualquier modelo de evaluación que pretenda contribuir a la mejora. Pero para que pueda operar de forma eficaz y transformar el sistema en menos de una década, no solo ha de ir acompañada de las acciones propias de un *feedback* efectivo, sino que ha de operar sobre todos y cada uno de los elementos constitutivos del nivel correspondiente^{52, 53}. Por tales motivos el enfoque formativo, en primer lugar, ha de ser suscrito por aquella instancia del sistema que es competente en cuanto a las actuaciones que requiere ese *feedback* efectivo individualizado. En segundo lugar, la correspondiente evaluación ha de ser censal por razones tanto de eficacia como de equidad. Ningún niño, ningún adolescente debería quedarse atrás y ello solo se conseguirá si las acciones proactivas, correctoras o de remediación alcanzan eficazmente a todos aquellos que las necesiten, identificados sobre la base de esa evaluación propia del nivel de que se trate: el aula, el centro, el distrito, la Comunidad Autónoma o el Estado en el caso este último, en particular, de sistemas centralizados.

En España, con un nivel notable de descentralización en materia educativa, son las Administraciones educativas quienes disponen de la competencia y de los recursos de gestión para actuar a nivel de cada alumno, de cada aula o de cada centro; de ahí que, por razones obvias, una evaluación censal a nivel del conjunto del Estado que pretenda ser formativa actuando directamente sobre los niveles inferiores resulte, con toda probabilidad, políticamente disfuncional.

3.3.2 Una función diagnóstica

La función diagnóstica de la evaluación puede operar sobre una muestra de sujetos o de elementos estadísticamente representativa de la correspondiente población. Ello permite definir las políticas, evaluar su impacto o corregirlas de un modo acorde con el nivel competencial y siempre de forma mucho más económica. Tal es el caso, por ejemplo, de las pruebas de PISA cuando se aplican, sea sobre una muestra de alumnos y de centros representativa de la población a nivel del conjunto del

52 LOPEZ RUPEREZ, F. (2001). *Op. cit.*

53 LEVIN, B. (2008). *Op. cit.*

Estado, sea sobre muestras ampliadas representativas de regiones, o en el caso español, de Comunidades Autónomas.

3.3.3 Una dimensión de responsabilidad

Otra de las dimensiones de la evaluación del sistema educativo –o de algunos de sus niveles– que resulta ineludible a la hora de concebir un modelo efectivo es la dimensión de responsabilidad o, si se quiere, de rendición de cuentas por los resultados producidos al gestionar el correspondiente ámbito competencial. En ausencia de esta dimensión es muy probable que intereses oportunistas de los actores de un nivel dado predominen sobre el impulso de mejora, de acuerdo con un balance en el que los intereses generales terminan perdiendo⁵⁴. Cada nivel habrá de rendir cuentas ante el nivel superior en un sistema en escalera cuyo último peldaño es la propia sociedad en su conjunto.

3.3.4 Una aproximación multinivel

Como cabe advertir con facilidad, el carácter externo o interno de la evaluación, su naturaleza censal o muestral, formativa, sumativa o de certificación se cruzan con los diferentes niveles de realidad de los sistemas educativos para generar una amplia colección de posibilidades, condicionadas no solo por aspectos conceptuales como los descritos más arriba, sino también por factores políticos tales como la distribución competencial por niveles. Este conjunto amplio de criterios y de interacciones habrá de ser tomado en consideración cuando se trate de concebir un modelo de evaluación del sistema educativo, integrado y coherente.

24

3.3.5 Un análisis competencial

El artículo 149 de la Constitución española⁵⁵ define las competencias exclusivas del Estado en diferentes ámbitos y materias. En su punto trigésimo, y en lo concerniente a la educación, establece como competencias exclusivas del Estado la «Regulación de las condiciones de obtención, expedición y homologación de títulos académicos y profesionales y normas básicas para el desarrollo del artículo 27 de la Constitución, a fin de garantizar el cumplimiento de las obligaciones de los poderes públicos en esta materia». La primera parte de este mandato constitucional legitima al Estado para ordenar las llamadas 'evaluaciones de certificación' que conducen a títulos. La segunda, mucho más genérica, da soporte a la ordenación básica de aquellas evaluaciones generales orientadas a promover el derecho a la educación de los ciudadanos españoles en condiciones de calidad equivalente. Ese ha sido el criterio aplicado en la regulación de las evaluaciones del sistema, y aceptado a la postre hasta la LOMCE, por los diferentes actores políticos, o por su mayoría. Por otra parte, y apelando principalmente a lo establecido y aceptado a este respecto en la LOCE y en la LOE, es posible definir el siguiente esquema

54 WOESSMANN *et al.* (2009). *Op. cit.*

55 <<https://www.boe.es/legislacion/documentos/ConstitucionCASTELLANO.pdf>>.

–basado en la Constitución española– de reparto competencial en materia de evaluación del sistema educativo:

- Corresponde al Estado la evaluación general del sistema educativo, lo que comporta la concepción y elaboración del correspondiente modelo, la definición de los ámbitos de evaluación y la elaboración de las pruebas correspondientes a ese nivel, así como la aprobación de la normativa básica que regula dicha evaluación.
- Corresponde al Estado, en colaboración con las Administraciones educativas, garantizar que la aplicación del modelo de evaluación general del sistema se produzca con criterios de homogeneidad, de manera que sea posible la comparabilidad de los datos entre sujetos, entre centros o entre Comunidades Autónomas, la obtención a partir de ellos de conclusiones fiables y la implementación efectiva de las mismas.
- Corresponde al Estado la dirección y coordinación de la participación de España en las pruebas internacionales.
- Corresponde al Estado la regulación de las evaluaciones de certificación conducentes a diplomas o títulos.
- Corresponde a las Administraciones educativas colaborar con el Estado en el desarrollo del modelo de evaluación general del sistema y ser consultadas sobre la normativa básica elaborada al efecto.
- Corresponde a las Administraciones educativas, en el ámbito de sus competencias y en colaboración con los órganos correspondientes del Estado, ejecutar las evaluaciones que, en el marco de la evaluación general del sistema, se establezcan, así como controlar su correcto desarrollo.
- Corresponde a las Administraciones educativas colaborar con el Estado en la participación de España en pruebas internacionales.
- Corresponde a las Administraciones educativas, en el marco de sus competencias, elaborar y realizar planes de evaluación periódica de los centros educativos, compartiendo unas directrices básicas acordadas en el seno de la Conferencia sectorial.
- Corresponde a las Administraciones educativas regular la forma en que los resultados de las diferentes evaluaciones que se realizan en el nivel de los centros, se traduzcan en planes de mejora, con la orientación y el apoyo necesarios.

3.3.6 Un criterio económico

Una acepción genérica de lo económico, a la que se aludirá en lo que sigue, se basa en la famosa 'navaja de Ockham': «*Pluralitas non est ponenda necessitate*» (La pluralidad no se debe multiplicar sin necesidad). Aplicado ese principio de economía, o de parsimonia, al caso que nos ocupa, se expresaría en la forma siguiente: «No hay que multiplicar las evaluaciones sin necesidad». Y no hay que hacerlo, en primer lugar, porque suelen ser caras, sobre todo cuando operan sobre los niveles más complejos del sistema. En ocasiones, una falta de afinación de las evaluaciones educativas previstas

en las normas, de forma que no se respete el criterio económico, hace inviable su realización, por lo que no resulta inusual que se recurra a una forma degradada de las mismas que las haga asumibles desde un punto de vista presupuestario. Con ello se corre el riesgo cierto de convertirlas en insuficientes en el plano educativo y ajenas a la concepción inicial.

Pero en el caso que nos ocupa, esa versión del principio de parsimonia aplicado a la evaluación ayuda, además, a preservar la efectividad de la herramienta. Una inflación de evaluaciones, regionales, nacionales e internacionales, puede alterar la fiabilidad del instrumento e incrementar el error de la medida por simple hartazgo de los protagonistas. Esa disponibilidad disminuida por parte de los actores de base se traduciría, probablemente, en una percepción defectuosa de la realidad evaluada y podría inducir la adopción de decisiones erróneas y de actuaciones inconvenientes desde su propia concepción.

La participación de España en evaluaciones internacionales a gran escala, hace posible y necesario integrar éstas, de forma efectiva, en el modelo de evaluación del sistema educativo. El caso paradigmático lo constituye PISA que, como es sabido, se refiere a las competencias adquiridas por los alumnos al final de la 'educación secundaria inferior', que en la mayor parte de los países de la OCDE se produce a los 15 años, en tres áreas básicas: Lectura, Matemáticas y Ciencias.

26

En su última edición (2015), las 17 Comunidades Autónomas españolas participaron con muestra ampliada en las pruebas. Ello haría innecesaria cualquier evaluación muestral referida a la ESO, sea en el nivel del conjunto del Estado, sea en el nivel de las Comunidades Autónomas. Además de beneficiarse de las ventajas que comporta la aplicación de ese principio de parsimonia, la riqueza de los cuestionarios que acompañan a las pruebas de PISA abre unas posibilidades de análisis secundarios muy superiores a las que permitirían evaluaciones de ámbito nacional o autonómico. Tales análisis secundarios, convenientemente planificados, permiten potenciar las funciones de diagnóstico de dicha evaluación internacional y la aplicación de la terapia correspondiente. Aunque en menor medida, los anteriores razonamientos son asimismo de aplicación a la Educación Primaria a propósito de las pruebas TIMMS y PIRLS.

3.4 El modelo propuesto

El modelo de evaluación del sistema educativo español que se propone contempla tres ámbitos territoriales o niveles de realidad del sistema: el centro escolar, la Comunidad Autónoma y el conjunto del Estado. Combina, de un modo funcional y coherente, las evaluaciones externas con las internas, las censales con la muestrales, las formativas con las diagnósticas; e incorpora, de forma efectiva, las evaluaciones internacionales en un marco integrado y eficiente, acorde con el principio de parsimonia. Además, considera la evaluación de certificación de Bachiller de forma análoga a como lo vienen haciendo, desde el siglo pasado, los países de nuestro entorno cultural.

La estructura básica del modelo y sus componentes se muestran, de un modo sintético, en la tabla 6 y se desarrollan a continuación.

Tabla 6
Rasgos característicos del modelo de evaluación del sistema educativo español que, de acuerdo con una aproximación multinivel, se propone

Órgano competente	Ámbito territorial		
	Centro educativo	Comunidad Autónoma	Conjunto del Estado
Centro educativo	<ul style="list-style-type: none"> – Planes anuales de autoevaluación con el apoyo y orientación de las Administraciones educativas. 	*	*
Administración educativa	<ul style="list-style-type: none"> – Planes de evaluación formativa censal de los centros. – Planes de mejora a la vista del análisis objetivo de los datos resultantes de los Planes de evaluación formativa. 	<ul style="list-style-type: none"> – Planes trienales de evaluación del sistema a nivel autonómico que integren: <ul style="list-style-type: none"> – Los Planes de evaluación formativa censal que, en un periodo de tres años, alcancen a todos los alumnos, a todas las aulas y a todos los centros. – Las evaluaciones internacionales con información desagregada y representativa a nivel de Comunidad Autónoma. – Otras evaluaciones específicas de ámbito autonómico que las circunstancias en cada caso aconsejen. 	*
Ministerio de Educación	*	*	<ul style="list-style-type: none"> – Planes trienales de evaluación general del sistema educativo que integren: <ul style="list-style-type: none"> – Las evaluaciones formativas censales de los centros realizadas por las Administraciones educativas. – Evaluaciones descriptivas basadas en el Sistema estatal de indicadores de la Educación. – Las evaluaciones internacionales y los análisis secundarios pertinentes. – Otras evaluaciones específicas muestrales que las circunstancias, en cada caso, aconsejen. – Evaluaciones de certificación conducentes al título de Bachiller.

3.4.1 En el nivel del centro educativo

El modelo contempla, en este nivel, el desarrollo por parte de los centros educativos sostenidos con fondos públicos de PLANES ANUALES DE AUTOEVALUACIÓN, de carácter preceptivo, elaborados por los propios centros con la orientación y el apoyo de las Administraciones educativas. Es considerado como un instrumento fundamental para generar *feedbacks* formativos y hacer inteligente la gestión de los centros, es decir, para corregir los errores y aprender de la experiencia.

Sin perjuicio de esta acción evaluadora –desarrollada al amparo del principio de autonomía de los centros–, correspondería a las Administraciones educativas, en el ámbito de sus competencias, la concepción e implementación de PLANES DE EVALUACIÓN FORMATIVA DE LOS CENTROS DOCENTES. Contendrían evaluaciones realizadas con una periodicidad trienal, a razón de una evaluación para cada una de las dos etapas que definen las enseñanzas básicas –Educación Primaria y Educación Secundaria Obligatoria– y aplicadas sobre cursos iniciales, o relativamente iniciales, de las correspondientes etapas, con el fin de hacer más efectiva su función formativa a nivel del alumno. De carácter censal en este ámbito –realizadas sobre cada uno de los alumnos–, incorporarían entre sus finalidades principales, por un lado, la mejora de los resultados de cada alumno, tomando en consideración los modelos de 'logro contextualizado' o, en su caso, de 'valor añadido contextualizado'⁵⁶; y, por otro, la aplicación del principio de responsabilidad o de rendición de cuentas ante las Administraciones educativas por los resultados obtenidos, tras corregir la influencia del nivel socioeconómico y cultural de los estudiantes y de sus familias sobre el rendimiento escolar.

Dicha acción evaluadora se complementaría, en los casos oportunos, con PLANES DE MEJORA a la vista del análisis objetivo de los datos resultantes de los PLANES DE EVALUACIÓN FORMATIVA DE LOS CENTROS DOCENTES. Estos PLANES DE MEJORA –preceptivos aunque acordados con los centros– se extenderían a lo largo de un trienio, tiempo éste que diferentes experiencias⁵⁷ han revelado como suficiente para corregir el rumbo de un centro educativo, con efectos medibles sobre sus resultados.

3.4.2 En el nivel de la Comunidad Autónoma

Las Administraciones educativas elaborarían PLANES TRIENALES DE EVALUACIÓN DEL SISTEMA en el ámbito autonómico que integrarían, de un modo coherente, los siguientes elementos y sus resultados:

- Los PLANES DE EVALUACIÓN FORMATIVA DE LOS CENTROS que, mediante una adecuada programación, permitirían alcanzar, en un periodo de tres años, a todos los alumnos de los cursos objeto de evaluación, a todas sus aulas y a todos los centros. Para que todos los alumnos pudieran ser evaluados, a lo largo de su escolaridad obligatoria, una vez en la Educación Primaria y otra en la ESO,

56 OECD (2008). *Measuring Improvements in Learning Outcomes. Best Practices to Assess the value-Added of Schools*. OECD Publishing. Paris.

57 LÓPEZ RUPÉREZ, F. (2008). «La Educación Pública Prioritaria en la Comunidad de Madrid y el reto de la integración de la población socialmente desfavorecida» en *Políticas educativas para la cohesión social*. Comunidad de Madrid-Fundación Europea Sociedad y Educación. Madrid.

las correspondientes pruebas deberían realizarse el primer curso del trienio para la de Primaria y el último para la de Secundaria.

- Aquellas evaluaciones internacionales –las más importantes son también de carácter trienal– que aporten una información desagregada y representativa, a nivel de Comunidad Autónoma, incluyendo una planificación de los análisis secundarios que fueran pertinentes, de acuerdo con la visión propia de cada Plan.
- Otras evaluaciones específicas sobre aspectos concretos del sistema o relativos al impacto de las políticas en el ámbito autonómico, que, en cada caso, las circunstancias aconsejen.

Los citados planes trienales, no solo tendrían que contener la planificación de las actuaciones evaluadoras que deberían realizar las correspondientes Administraciones educativas, sino además habría que prever en ellos el uso que dar a la información producida por las distintas evaluaciones previstas en cada plan, con el propósito de generar *feedbacks*; es decir, con la finalidad de concebir políticas y desarrollar acciones efectivas para impulsar la mejora.

3.4.3 En el nivel del conjunto del Estado

El Gobierno de España, a través de su Ministerio de Educación, elaboraría en su ámbito de competencia, y en coherencia con la concreción del modelo en el nivel inferior, PLANES TRIENALES DE EVALUACIÓN GENERAL DEL SISTEMA EDUCATIVO. Dichos Planes integrarían los siguientes elementos, junto con los correspondientes análisis de datos, con finalidad diagnóstica y de mejora de las políticas en el nivel del Estado:

- Las evaluaciones formativas censales de los centros realizadas por las Administraciones educativas, que serían elaboradas con criterios homogéneos en todo el territorio nacional acordados en el seno de la Conferencia sectorial de educación.
- Evaluaciones efectuadas a partir de la explotación del Sistema estatal de indicadores de la Educación.
- Las evaluaciones internacionales y los análisis secundarios que fueran pertinentes de acuerdo con la visión propia de cada Plan.
- Otras evaluaciones específicas de carácter muestral sobre aspectos concretos del sistema, o relativos al impacto de las políticas en el ámbito del Estado, que en cada caso las circunstancias aconsejen.

Al igual que en el ámbito correspondiente a las Administraciones educativas, estos planes trienales deberían especificar las acciones que desarrollar con la información producida. Así, como contraejemplo, cabe destacar que después de 15 años de vida efectiva de PISA, de seis ediciones y dos ciclos completos de esta evaluación a gran escala, en los que España ha participado, ninguna previsión sobre qué hacer con esa muy rica información producida se ha formulado por parte del Ministerio de Educación, ni ninguna reforma específica orientada por esos datos ha sido pensada,

formulada e implementada en esa década y media, aunque en alguna ocasión se hayan invocado superficialmente.

Por otra parte, y de acuerdo con sus competencias exclusivas, corresponde al Estado la concepción y elaboración de EVALUACIONES DE CERTIFICACIÓN conducentes al título de Bachiller, cuya aplicación se produciría con la colaboración leal de las Administraciones educativas en sus respectivos territorios. Este tipo de evaluaciones finales de la 'educación secundaria superior' constituye, desde hace mucho tiempo, la regla en los países de nuestro entorno cultural. Tal es el caso del BAC francés, del BAC belga, del BAC de las Escuela Europeas, de la MATURITÀ italiana o del ABITUR alemán, entre otros. Sin perjuicio de sus ineludibles efectos académicos, el carácter censal y homogéneo de estas pruebas permitiría el análisis de sus resultados con utilidades diagnósticas y formativas, tanto para el conjunto del Estado como en el nivel propio de las Comunidades Autónomas.

La síntesis del modelo propuesto que se muestra en la tabla 6, al adoptar el mismo patrón de análisis que el de las tablas anteriores, facilita la comparación con los modelos correspondientes a los marcos normativos de la democracia en materia de evaluación del sistema educativo. De dicha comparación emergen las siguientes ventajas del modelo que se propone:

- Resulta más claro que los precedentes en cuanto a la terminología y más sencillo en cuanto a sus componentes.
- Se evidencia deliberadamente en él la estructura multinivel.
- Las finalidades de sus diferentes elementos se hacen más diáfanas y orientan, de un modo explícito y coherente, las distintas modalidades de evaluación consideradas: internas o externas; censales o muestrales; formativas, diagnósticas o de certificación.
- Una cierta homología del modelo recorre, en cierta medida, los tres niveles considerados, lo que resulta palmario en el segundo y tercer nivel. Esa suerte de 'invariancia de escala' dota de coherencia al modelo y facilita su integración.
- Compromete una utilización de la información contenida en el SISTEMA ESTATAL DE INDICADORES y no, como ha sucedido hasta el presente, de una mera enumeración de datos organizados de un modo sistémico.
- Integra como elemento capital las evaluaciones internacionales, lo que contribuye a hacerlo más parsimonioso que los anteriores, evitándose redundancias superfluas o no convenientes.

Cabe señalar que el modelo aportaría información sobre el rendimiento de los alumnos en materias fundamentales, de carácter censal en 2.º de Primaria y en 1.º de la ESO mediante la evaluación formativa a nivel de cada alumno, de cada centro y de cada Comunidad Autónoma; y en 2.º de Bachillerato mediante la evaluación de certificación conducente al título de Bachiller. Y de carácter muestral en 4.º de Primaria y 3.º de ESO (15 años) a través de las pruebas del TIMSS y el PIRLS, y de PISA respectivamente.

La organización por las Administraciones educativas de las pruebas censales de evaluación formativa se efectuaría de modo que en los tres años de duración de cada plan todos los alumnos de los referidos cursos –2.º de Primaria y 1.º de la ESO– recibirían una evaluación externa con efectos sobre la mejora escolar a través tanto del *feedback* formativo como del impacto positivo de la rendición de cuentas. La periodicidad de la prueba censal de Bachillerato, por razones obvias, sería anual (bianaual si se incluye la convocatoria extraordinaria).

Por su parte, la información muestral, procedente de las pruebas internacionales, se produciría con una periodicidad de 4 y 5 años respectivamente para el TIMSS y el PIRLS y de 3 para PISA, disponiendo en este último caso de datos representativos, no solo a nivel del conjunto del Estado, sino también de sus diecisiete Comunidades Autónomas. La tabla 7 resume los detalles del conjunto de las evaluaciones externas contempladas en el modelo.

Tabla 7
Conjunto de evaluaciones externas periódicas contempladas en el modelo

Cácter	Curso	Prueba	Periodicidad
Evaluación Muestral	4.º curso de Ed. Primaria	PIRLS	Cada 5 años
		TIMSS	Cada 4 años
Evaluación Muestral	3.º curso de ESO	PISA	Cada 3 años
Evaluación Censal	2.º curso de Ed. Primaria	De evaluación formativa censal	Cada 3 años
Evaluación Censal	1.º curso de ESO	De evaluación formativa censal	Cada 3 años
Evaluación Censal	2.º curso de Bachillerato	De evaluación de certificación	Anual

Fuente: elaboración propia.

4. Conclusiones y recomendaciones

La evaluación del sistema educativo español constituye un factor estratégico para mejorar la calidad de su gobernanza. Como anticipáramos en una publicación anterior⁵⁸:

58 LÓPEZ RUPÉREZ, F.; GARCÍA GARCÍA, I.; EXPÓSITO CASAS, E. (2017). *Op. cit.*

«La evaluación de un sistema educativo concierne, por igual, a la rendición de cuentas y a la dirección o pilotaje del sistema. Es, por lo tanto, un aspecto de la gobernanza verdaderamente crítico. Un conocimiento claro sobre el qué, el para qué, el por qué y el cómo de la evaluación a gran escala resulta imprescindible como elemento previo a la acción. Pero sólo la integración en el proceso de los diferentes niveles de realidad (individuo, instituciones escolares y sociedad), la reflexión sobre los efectos indeseados, la consideración del conocimiento técnico y profesional y la movilización de las mejores prácticas propiamente políticas permitirá lograr el éxito en esta dimensión crucial de la gobernanza».

El presente trabajo ha pretendido contribuir a la mejora de las políticas de evaluación desde la clarificación de sus bases conceptuales, el análisis de los modelos presentes en las diferentes leyes educativas de nuestra democracia y la conceptualización de un modelo de evaluación evolucionado para el sistema educativo español, multinivel, integrado, coherente y parsimonioso. En lo que sigue, se muestran un conjunto de conclusiones y se aportan algunas recomendaciones destinadas a contribuir a una mayor dosis de acierto en la elaboración del nuevo marco regulatorio que nos espera.

4.1 Conclusiones

Del conjunto de reflexiones y de análisis efectuado en el presente estudio se derivan, a modo de síntesis, las siguientes conclusiones:

32

- a) De conformidad con el consenso global dominante, las evaluaciones nacionales e internacionales se han convertido en elementos centrales para asegurar la calidad y mejorar la educación.
- b) La esencia de la inteligencia es el *feedback* –o retroalimentación– en tanto que bucle circular de información y de acción.
- c) La evaluación genera una información valiosa que solo se convierte en *feedback* cuando es utilizada para alterar la brecha existente ente el nivel actual y el nivel de referencia que se pretende alcanzar.
- d) La evaluación formativa es aquella expresamente orientada a la mejora, de ahí su conexión natural con el concepto de *feedback* y con la aplicación transversal de este concepto en el ámbito de la educación.
- e) Lo importante, en orden a promover la mejora educativa, no es tanto la información que produce la evaluación como lo que se hace con esa información.
- f) El centro educativo es la unidad clave del sistema desde el punto de vista tanto de la mejora como de la evaluación; evaluación que ha de combinar su finalidad formativa con la de la responsabilización, o rendición de cuentas, por los resultados derivados de administrar la autonomía escolar.
- g) No obstante lo anterior, la concepción de un modelo efectivo de evaluación del sistema educativo español ha de integrar, de un modo claro y coherente, el nivel de centro, el nivel de la Comunidad Autónoma y el nivel del conjunto del Estado.
- h) La evolución del modelo de evaluación del sistema educativo en España, establecido en las leyes

de la democracia –desde la LOGSE hasta la LOE, pasando por la LOCE– se ha caracterizado por una clara progresividad en cuanto a su alcance y planteamientos.

- i) La aportación de la LOMCE ha supuesto una cierta ruptura con esa línea evolutiva, en el plano tanto conceptual como metodológico; ruptura que ha sido posteriormente derogada mediante el correspondiente Real Decreto-ley y sustituida por una fórmula provisional e incompleta.

4.2 Recomendaciones

El modelo de evaluación del sistema educativo español, que a partir de las anteriores consideraciones y análisis comparativos en el presente trabajo se propone, puede articularse en las recomendaciones siguientes:

- a) La concepción de un modelo moderno de evaluación general para nuestro sistema educativo debería tomar explícitamente en consideración, de forma integrada, los siguientes criterios básicos: un enfoque formativo, una función diagnóstica, una dimensión de responsabilidad, una aproximación multinivel, un análisis competencial y un principio económico.
- b) En el nivel de centro, el modelo debería contemplar el desarrollo de:
- Planes anuales de autoevaluación, preceptivos y realizados por los centros escolares con el apoyo y orientación de las Administraciones educativas.
 - Planes de evaluación formativa censal de los centros, realizados por las Administraciones educativas.
 - Planes de mejora, acordados con los centros sostenidos con fondos públicos, en los casos oportunos y a la vista del análisis objetivo de los datos resultantes de los Planes de evaluación formativa.
- c) En el nivel de Comunidad Autónoma, el modelo debería contemplar el desarrollo de Planes trienales de evaluación del sistema, a nivel autonómico, que integrarían:
- Los Planes de evaluación formativa censal que, en un periodo de tres años, alcancen a todos los alumnos, a todas las aulas y a todos los centros en las etapas y cursos de la enseñanza básica objeto de evaluación.
 - Las evaluaciones internacionales que contengan información desagregada y representativa a nivel de Comunidad Autónoma.
 - Otras evaluaciones específicas de ámbito autonómico que las circunstancias, en cada caso, aconsejen.
- d) En el nivel del conjunto el Estado el modelo debería contemplar el desarrollo de Planes trienales de evaluación general del sistema educativo que integrarían:
- Las evaluaciones formativas censales realizadas por las Administraciones educativas pero elaboradas con criterios homogéneos, en todo el territorio nacional, acordados en el seno de la Conferencia sectorial de educación.

- Evaluaciones efectuadas a partir de la explotación del Sistema estatal de indicadores de la Educación.
 - Las evaluaciones internacionales y los correspondientes análisis secundarios que fueran pertinentes, de acuerdo con la visión propia de cada Plan.
 - Otras evaluaciones específicas de carácter muestral sobre aspectos concretos del sistema que, en cada caso, las circunstancias aconsejen.
 - Por otra parte, y de acuerdo con sus competencias exclusivas, corresponde al Estado la concepción y elaboración de Evaluaciones de certificación conducentes al título de Bachiller, cuya aplicación se produciría con la colaboración leal de las Administraciones educativas en sus respectivos territorios.
- e) Tanto los planes trienales de evaluación de las Administraciones educativas como los del Estado no solo tendrían que contener la planificación de las actuaciones evaluadoras que debería realizar la correspondiente Administración, sino además habría que prever en ellos el uso que dar a la información producida por las distintas evaluaciones previstas en cada plan, con el fin de, posteriormente, generar *feedbacks*, concebir políticas y desarrollar acciones efectivas para impulsar la mejora.
- f) Aprender de las insuficiencias y de los errores de los modelos anteriores y, en particular, aplicar un 'principio de parsimonia' consistente en no multiplicar las evaluaciones sin necesidad.

5. Referencias

BAIRD, J.; HOPTENBECK, T. N.; NEWTON, P. E.; STOBART, G.; STEEN-UTHEIM, A. T. (2014). *Assessment and Learning: State of the Field Review*. Norwegian Knowledge Centre for Education & University of Oxford.

BLACK, P.; WILIAM, D. (1998). «Inside the Black Box: Raising Standards Through Classroom Assessment». *Phi Delta Kappa*, oct 1998, pp. 1-18.

BLACK, P.; WILIAM, D. (1998). «Assessment and classroom learning». *Assessment in Education*, 5, 7-75

BLOOM, B. S.; MADAUS, G. F.; HASTINGS, J. T. (1971). *Handbook on Formative and Summative Evaluation of Student Learning*. New York: McGraw-Hill.

BREAKSPEAR, S. (2012). *The Policy Impact of PISA: An Exploration of Normative Effects of International Benchmarking in School System Performance*. OECD Education Papers, No. 71, OECD Publishing. <http://dx.doi.org/10.1787/5k9fdfqffr28-en>

DE KETELE, J. M.; ROEGIERS, X. (1995). *Metodología para la recogida de la información*. Ed. La Muralla. Madrid.

DE LANDSHEERE, G. (1996). *El pilotaje de los sistemas educativos*. Ed. La Muralla. Madrid.

- ELMORE, R. F. (2004). *School reform from inside out: Policy, practice, and performance*. Harvard Education Press. Cambridge, MA.
- FUCHS, L. S.; FUCHS, D. (1986). «Effects of systematic formative evaluation. A meta-analysis». *Exceptional Children*, 53(3), 199-208.
- FUSTER, J. (2012). «Entrevista». *Participación Educativa*, 2.ª época, vol. 1, n.º 1, 29-31.<ntic.educacion.es/cee/revista/pdfs/RPE_DIC2012_05Fuster.pdf>
- FUSTER, J. (2014). *Cerebro y libertad. Los cimientos cerebrales de nuestra capacidad de elegir*. Cap. 4. El ciclo percepción/acción. Ariel. Barcelona.
- HANUSHEK, E. A.; LINK, S.; WOESSMANN, L. (2012). *Does School Autonomy Make Sense Everywhere? Panel Estimates from PISA*. ADB Economics Working Paper Series No. 296. Asian Development Bank.
- HATTIE, J. A. (2009). *Visible Learning. A Synthesis of Over 800 Meta-analyses Relating to Achievement*. Routledge. Abington.
- HATTIE, J. A. (2017). «Aprendizaje visible» para profesores. *Maximizando el impacto del aprendizaje*. Cap.7. El flujo de la clase: el lugar del feedback. Paraninfo. Madrid.
- HATTIE, J. A.; TIMPERLEY, H. (2007). «The Power of Feedback». *Review of Educational Research*, 77, 81-112.
- HAVNES, A.; SMITH, K.; DYSTHE, O.; LUDVIGSEN, K. (2012). «Formative assessment and feedback: Making learning visible». *Studies in Educational Evaluation*, 38, pp.21-27.
- HINTON, CH.; MIYAMOTO, K.; DELLA CHIESA, B. (2008). «Brain Research, Learning and Emotions: implications for education research, policy and practice». *European Journal of Education*, 43(1), 2008, 87-103.
- KLUGER, A. N.; DENISI, A. (1996). «The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory». *Psychological Bulletin*, 119, 254-284.
- KOVACH, D. (2017). «The Computational Theory of Intelligence: Feedback». *International Journal of Modern Nonlinear Theory and Application*, 6, 70-73.
- HIGHAM, J.; YEOMANS, D. (2010). «Working together? Partnership approaches to 14-19 education in England». *British Educational Research Journal*, 36 (3), pp. 379–401.
- LEVIN, B. (2008). *How to change 5000 schools*. Harvard Education Press. Cambridge MA.
- LÓPEZ RUPÉREZ, F. (1994). *La gestión de calidad en educación*. Ed. La Muralla. Madrid.
- LÓPEZ RUPÉREZ, F. (2001). *Preparar el futuro. La educación ante los desafíos de la globalización*. Ed. La Muralla. Madrid.

LÓPEZ RUPÉREZ, F. (2008). «La Educación Pública Prioritaria en la Comunidad de Madrid y el reto de la integración de la población socialmente desfavorecida» en *Políticas educativas para la cohesión social*. Comunidad de Madrid- Fundación Europea Sociedad y Educación. Madrid.

LÓPEZ RUPÉREZ, F.; GARCÍA GARCÍA, I.; EXPÓSITO CASAS, E. (2017). *La calidad de la gobernanza del sistema educativo español. Un análisis empírico*. Universidad Camilo José Cela. Madrid.

OECD (2004). *What Makes Schools Systems Perform? Seeing School Systems through the Prism of PISA*. OECD Publishing. Paris.

OECD (2008). *Measuring Improvements in Learning Outcomes. Best Practices to Assess the value-Added of Schools*. OECD Publishing. Paris.

OECD (2016). *PISA 2015 Results. Policies and Practices for Successful Schools. Volume I*. OECD Publishing. Paris.

OECD (2017). *Education at a Glance. OECD Indicators*. OECD Publishing. Paris.

PONT FERRER, B. (2017). *Reformas educativas: el caso del liderazgo escolar en perspectiva comparada*. Tesis Doctoral. Universidad Complutense de Madrid. Madrid.

RAMAPRASAD, A. (1983). «On the definition of feedback». *Behavioural Science*, 28, 4-13.

36 RICHARDSON, G. P. (1991). *Feedback Thought in Social Science and Systems Theory*. University of Pennsylvania Press. Philadelphia.

SCHILDKAMP, K.; LAI, M. K. (2013). «Data-based Decision Making: An Overview». En Schildkamp K., Lai M., Earl L. (eds). *Data-based Decision Making in Education. Studies in Educational Leadership*, vol 17. Springer, Dordrecht. <https://doi.org/10.1007/978-94-007-4816-3_2>.

SHUTE, V. J. (2008). «Focus on formative feedback». *Review of Educational Evaluation*, 37, 3-14.

SJOBERG, S. (2017). «PISA Testing. A Global Educational Race?». *Europhysicsnews*, 48/4, 17-20.

VON BERTALANFY, L. (1979). *Perspectivas de la teoría general de sistemas*. Alianza Editorial. Madrid.

WOESSMANN, L.; LUEDEMANN, E.; SCHUETZ, G.; WEST, M. R. (2009). *School Accountability, Autonomy, and Choice around the World*. Edward Elgar. Cheltenham.

Agradecimientos

Agradezco vivamente a la profesora Isabel García García su lectura crítica del texto, sus pertinentes observaciones y su colaboración decisiva en la edición del original.



Cátedra de Políticas Educativas

www.ucjc.edu/la-universidad/estructura-academica/catedras