

## Calidad y evaluación como tendencias globales en política educativa: estudio comparado de agencias nacionales de evaluación en educación obligatoria en Europa

Gabriel Álvarez-López<sup>1</sup>, María Matarranz<sup>2</sup>

**Resumen.** La evaluación hoy en día es una de las principales preocupaciones a nivel mundial. Esta preocupación radica en la necesidad de mejorar la calidad de sus sistemas educativos y en la sostenibilidad de las políticas públicas. Sin embargo, y aunque no hemos llegado a un consenso con respecto a cómo se deben afrontar, la gran mayoría de los países desarrollados cuentan con instituciones evaluadoras y programas de evaluación. En este contexto, y desde la realidad europea en la que nos encontramos, es interesante detenerse a analizar si estas instituciones y estas políticas son sostenibles a nivel supranacional y en qué condiciones de deben llevar a cabo estas evaluaciones.

Este trabajo propone reflexionar acerca de estas cuestiones a través del estudio realizado sobre las agencias de evaluación de la educación obligatoria de Alemania, Italia y España. Para llevar a cabo esta investigación se ha seguido la metodología comparada en educación, pues los objetos de estudio tienen un grado de homogeneidad – heterogeneidad que nos permite estudiarlos desde esta perspectiva. A través de este análisis y de la reflexión previa en torno a la calidad como preocupación de los sistemas educativos contemporáneos, y la evaluación como tendencia política global, extraemos unas conclusiones que nos invitan a reflexionar en torno a dos preguntas: ¿qué entendemos por calidad educativa y cómo puede la evaluación contribuir a la mejora de la misma?, ¿pueden los actuales modelos dominantes de evaluación promover una mejora de la calidad? A partir de esta reflexión, cerramos las conclusiones proponiendo la creación de una *Agencia Europea de Evaluación* para la etapa de escolarización obligatoria, que nos permita seguir construyendo un Espacio Europeo de Educación.

**Palabras clave:** Evaluación, Calidad educativa, Educación obligatoria, Organizaciones Internacionales, Política educativa supranacional, Análisis comparado

## [en] Quality and evaluation as global educational political trends: comparative study of national evaluation agencies in compulsory education in Europe

**Abstract.** Assessment today is one of the main concerns of the developed countries. This concern lies on the need to improve the quality of education systems and the sustainability of public policies. However, and despite the lack on consensus on how to face the challenges posed by quality and evaluation in national education systems, the majority of developed states have assessment institutions and national evaluation programs. In this context, and taking into account the European reality, it is interesting to draw attention to whether these institutions and policies are sustainable at a supranational level and on what conditions these evaluations can be carried out.

This paper intends to reflect on these issues by means of a study carried out on the assessment agencies of basic education in Germany, Italy and Spain. To carry out this research, we have followed the comparative methodology in education, since the objects of study present a degree of homogeneity – heterogeneity that allows us to consider them from this perspective. Through the analysis and previous reflection on quality as a current concern of contemporary educational systems, and evaluation as a global political trend, we draw conclusions that lead us to reflect on two questions: what do we mean by quality and how can evaluation contribute to its improvement? Can existing dominant models improve quality? From this reflection, we close the conclusions proposing the creation of a European Evaluation Agency for compulsory education, which will allow us to continue building a Common European area for research and innovation.

**Keywords:** Assessment, Educational quality, Compulsory education, International Organizations, Supranational educational policy, Comparative analysis

<sup>1</sup> Universidad Autónoma de Madrid (UAM)  
E-mail: [g.alvarez@uam.es](mailto:g.alvarez@uam.es)

<sup>2</sup> Universidad a Distancia de Madrid (UDIMA)  
E-mail: [maria.matarranz@udima.es](mailto:maria.matarranz@udima.es)

**Sumario.** Introducción. 1. La calidad y la evaluación como tendencia política global. 1.1. Calidad y educación. 1.2. Evaluación y rendición de cuentas en educación. 2. Calidad y evaluación de la educación en la Unión Europea. 3. Resultados de la investigación. 4. Discusión y conclusiones. 5. Referencias bibliográficas.

**Cómo citar:** Álvarez-López, G.; Matarranz, M. (2020). Calidad y evaluación como tendencias globales en política educativa: estudio comparado de agencias nacionales de evaluación en educación obligatoria en Europa. *Revista Complutense de Educación*, 31(1), 85-93.

## Introducción

La búsqueda de la calidad en la enseñanza se ha convertido casi en sinónimo de control, rendición de cuentas, financiación selectiva o, en suma, eficiencia de los sistemas educativos (Báez, 2007). Esta búsqueda de la calidad ha estado vinculada permanentemente al éxito en las evaluaciones tanto nacionales como internacionales. De esta manera, la evaluación en la educación obligatoria ha pasado a ocupar un lugar central en las agendas políticas de los Estados. Parte del éxito innegable de este crecimiento exponencial se debe, por una parte, al trabajo de los organismos supranacionales que están en continua retroalimentación con gobiernos nacionales y regionales y, por otra parte, al impacto mediático de los programas internacionales de evaluación de competencias (Calero y Choi, 2012). El desarrollo de la evaluación en la etapa obligatoria (ahora a través de evaluaciones con un carácter más competencial) refleja, en todo caso, un énfasis en los “resultados” del sistema educativo y, en concreto, del logro de los estudiantes. No se trata de algo trivial, sino de un desplazamiento del foco de interés en los análisis e intervenciones sobre el sistema educativo, desde los inputs y procesos hacia los conocimientos (y en el mejor de los casos las competencias) adquiridos. Podríamos hablar de una “colonización” de los sistemas educativos (Calero y Choi, 2012) por parte de la evaluación, bajo la premisa de que, a mayor evaluación en los sistemas educativos, mayor será la calidad de los mismos.

La investigación que se presenta pretende reflexionar sobre la calidad y la evaluación como tendencias globales en política educativa, por medio de un estudio comparado en tres países de la Unión Europea que nos permita analizar las competencias de las agencias nacionales de evaluación existentes.

## 1. La calidad y la evaluación como tendencia política global

### 1.1. Calidad y educación

La calidad es un término complejo en el que se pueden identificar diversas concepciones: fenómeno excepcional (excelencia), perfección o mérito, adecuación a propósitos, producto económico/vector de eficiencia (relación calidad/precio) y como transformación (cambio cualitativo). Así pues, la calidad como concepto conlleva una alta dosis de ambigüedad (Martínez Rizo, 2009; Stake, 2006) lo que ha facilitado su rápida adopción por los distintos agentes sociales (Tiana, 2006).

En materia educativa, la calidad ha tenido una implantación paulatina en los sistemas educativos occidentales. El interés por la misma ha ido creciendo exponencialmente desde la década de los 90 y, según Tiana (2006) entre los motivos que contribuyen a explicar el creciente interés encontramos: (1) la creciente presión de la economía sobre la educación, (2) el renovado énfasis puesto en la educación como factor de desarrollo, (3) la reconsideración del papel del Estado en cuanto prestador de servicios, (4) el malestar creado en las últimas décadas a consecuencia de la expansión y masificación registradas por los sistemas educativos, (5) la evidente tensión experimentada entre la búsqueda de la excelencia y de la equidad, o (6) la desilusión provocada muchas veces por los procesos de reforma educativa.

Gargallo (2003) sostiene que la preocupación por la mejora de la calidad aumenta en momentos de crisis económica en los que se deben implantar reformas que suponen esfuerzos económicos relevantes, acompañada por un incremento de la necesidad de evaluar dicho sistema.

Por lo tanto, la calidad en el ámbito educativo no es un término novedoso, no obstante es difícil de delimitar dado que existen tantas concepciones sobre la calidad en educación como maneras de comprender la educación misma. El concepto de calidad educativa abarca varias dimensiones que lo articulan en una unidad verdaderamente integral (Seiblod, 2000). Para autores como Gargallo (2003) o Tiana (2006) la calidad educativa es un sistema de coherencias múltiples que puede adoptar tres aproximaciones: la calidad como eficacia, la calidad como eficiencia, o la calidad como la satisfacción de necesidades y expectativas. Gargallo (2003), desde una concepción integradora del término, asume que la calidad debería aglutinar “la consideración de los inputs del sistema, de los procesos y resultados, de la satisfacción de los «clientes» y de los implicados en el proceso, en un proceso de construcción continuo” (p. 35).

Muñoz-Repiso y Murillo (2010) afirman que para delimitar el concepto de calidad educativa habría que hacer una distinción dado que en el mismo concepto podemos localizar dos elementos, por un lado, los fines educativos, por otro lado, los medios. Los fines educativos pertenecen, según los autores, al mundo de los valores, mientras que los medios hacen referencia a la eficacia escolar y a la mejora de la escuela.

De la mano de la búsqueda de la calidad en educación, y siendo conocedores de la complejidad que entraña, nos encontramos con la evaluación como instrumento de medida de la calidad. Así, Pérez Juste (2005) advierte del riesgo de que la evaluación acabe siendo un fin en sí misma, y por ello propone “cuidar con esmero su concepción, su diseño, su ejercicio y el propio uso de la información” (p. 25). En definitiva, la búsqueda de la calidad en educación y su evaluación se han convertido en protagonistas, sea cual sea el sentido que le demos.

## 1.2. Evaluación y rendición de cuentas en educación

Rendir cuentas en educación significa “hacer responsable a un individuo, grupo, agencia o gobierno de sus acciones y de sus consecuencias conforme intenta alcanzar unas metas específicas” (Baker, 2009, p. 92). Por su parte, Linn (2008) sostiene que “el objetivo de un sistema de responsabilidades es distinguir entre escuelas en cuanto a la efectividad de la instrucción, para que las acciones se puedan tomar para mejorar el rendimiento de los estudiantes en las escuelas consideradas ineficaces” (p. 708). Así, se empieza a identificar la dimensión política de la evaluación. Dorn y Ydesen (2014) identifican dos rasgos fundamentales en este aspecto: 1) el papel de la rendición de cuentas como un fenómeno que no puede ser tratado de forma aislada en la sociedad, 2) las prácticas de rendición de cuentas siempre parecen abarcar una cierta configuración históricamente determinada de posiciones de las partes interesadas.

Las políticas de rendición de cuentas no pasan inadvertidas para el universo educativo. Las principales implicaciones de la rendición de cuentas que han sido objeto de investigación se han centrado en aspectos como: la gestión pública (Lai y Wang, 2014), la elección de centro (Adams y Hill, 2006) y la comunidad educativa (Holme, Richards, Jimerson y Cohen, 2010). Dentro de esta comunidad educativa, los efectos negativos como la inflación de la puntuación (Koretz, 2008), de enseñanza-al-test (Stecher, 2002), y el aumento de tensiones emocionales (Berryhill, Linney y Fromewick, 2009), superan claramente los resultados positivos. También encontramos estudios que analizan los efectos que estas pruebas tienen sobre las políticas inclusivas (Davies y Elliot, 2012), o sobre la brecha racial (Darling – Hammond, 2007).

Algunas de las valoraciones más importantes se pueden dividir en críticas a la ideología que sustenta estas prácticas y críticas a las pruebas en sí mismas. En el grupo de las críticas a la concepción ideológica, destacan aquellas que advierten de la deriva hacia la “sociedad de la auditoría” (Kostogriz y Doecke, 2011), hacia la privatización del sector educativo (Ball, 2014), hacia un reduccionismo o simplificación de la realidad educativa (Labaree, 2014), o aquellas que señalan que a pesar de que las intenciones eran buenas, los resultados han sido nefastos (Amrein y Berliner, 2003). Estas críticas tienen su versión más dura en quienes hacen una enmienda a la totalidad a este tipo de políticas sosteniendo que significan un asalto a la pedagogía de calidad y a los ideales democráticos de equidad y justicia social (Au y Tempel, 2012; Tuck, 2013). En el grupo de las críticas enfocadas a las pruebas en sí mismas, a su diseño, implementación y uso, se encuentran trabajos que se hacen eco de los límites de estos sistemas (Lee y Wong, 2004), los que sostienen que han sido claramente insuficientes (Koretz, 2008), trabajos que afirman que las pruebas “pueden dictar contenidos, pero no pueden dictar pedagogía” (Diamond, 2012), o, por último, trabajos que sostienen que estas pruebas adolecen de contexto (Hutt, 2014).

La rendición de cuentas, por lo tanto, se ha convertido en una cuestión de gran relevancia que, pese a ser una tendencia educativa al alza, no está exenta de cuestionamientos. Para Smith (2014) el *accountability* es una de las consecuencias de la globalización que desemboca en una tendencia convergente mundial de utilización de los exámenes como una herramienta legítima para la reforma educativa. Sin embargo, Sobe (2015) lo entiende como una parte constitutiva de lo global y no como una consecuencia. Sea como fuere, la rendición de cuentas es una de las tendencias más determinantes en el desarrollo de los programas de evaluación de los sistemas educativos nacionales.

## 2. Calidad y evaluación de la educación en la Unión Europea

La globalización económica y cultural ha sido clave en la preocupación de los organismos supranacionales por las políticas educativas. Esta actúa básicamente en dos sentidos (Kamens y McNeely, 2010): por un lado, sobre la cultura política mundial y, por otro lado, en las organizaciones internacionales (gubernamentales o no) y las entidades supranacionales. Ambas a su vez influyen en los Estados que se ven “obligados” a evaluar y a realizar estudios que analicen lo que estos movimientos globales les indican. Así pues, los organismos internacionales y supranacionales tienen un papel fundamental. No se entiende el fenómeno globalizador sin su existencia, y así lo comparten Burbules y Torres (2001) cuando sostienen que

la globalización significa fundamentalmente el surgimiento de instituciones supranacionales, cuyas decisiones determinan y restringen las opciones políticas de cualquier Estado – nación en particular (...), es el papel de los Estados y de las estructuras administrativas y ministeriales en la interrelación y en la provocación de organismos supranacionales; o las redes de universalización de procedimientos y el papel del Estado nacional y su influencia política y económica (p. 13).

Desde los Acuerdos de Bretton Woods en 1944 con el nacimiento de los organismos que se han constituido como los máximos exponentes de esta globalización económica y cultural, el Fondo Monetario Internacional (FMI, por sus siglas) y el Banco Mundial (BM), la primacía de la economía neoliberal ha impactado en la formulación de políticas educativas en la práctica totalidad de los Estados desarrollados y en vías de desarrollo (Kamens y Benavot, 2011); a veces por la vinculación de financiación a reformas y otras directamente por recomendaciones explícitas de dichas instituciones u otras como la Organización para la Cooperación y el Desarrollo Económico (OCDE) o la Organización Mundial del Comercio (OMC).

Estas instituciones y algunas otras han buscado la vinculación de la educación con el trabajo y la creación de estándares de excelencia (que han sido utilizados más como instrumentos de control político que como herramientas de mejora educativa). Además, advierten García Garrido y García Ruíz (2005), si estas directrices suponen un proceso uniformizante, “podrían engendrar un cierto empobrecimiento de la educación” (p. 38). Sin embargo, también se reconoce la labor de estos organismos a favor de la calidad de los sistemas.

En la Unión Europea (UE), es claro el posicionamiento de la calidad educativa entre sus prioridades (ET2010 y ET2020), considerando fundamental el fomento de la calidad en materia educativa para alcanzar el objetivo estratégico formulado en el Consejo Europeo de Lisboa del año 2000

Así, la principal preocupación de la evaluación no sería tanto la calidad de las instituciones educativas en términos de “productos”, como servir de instrumento para mejorar la calidad con una orientación formativa. Desde el Consejo Europeo de Lisboa del año 2000, donde se entiende que el pilar desde el que construir el progreso social y económico debe ser el conocimiento, hasta nuestros días, la UE ha desarrollado tímidos intentos de poner en marcha evaluaciones externas de competencias lingüísticas de los estudiantes (Bonnet, 2002). Estos intentos se han consolidado en el Estudio Europeo de la Competencia Lingüística (Comisión Europea, 2012) que es la primera evaluación sistemática de competencia lingüística (lengua extranjera) que se hace desde esta organización supranacional.

Por otra parte, donde sí resulta evidente la puesta en marcha de mecanismos de rendición de cuentas es en la etapa terciaria, la cual, constituye un buen ejemplo de la búsqueda de la armonización por parte de los países integrantes del Espacio Europeo de Educación Superior (EEES), y que se ve representada por la creación en 1999 de una agencia de calidad a nivel europeo, la Red Europea para el Aseguramiento de la Calidad en la Educación Superior (ENQA). Este organismo fue posible gracias a la preocupación compartida por parte de los Estados miembros, que se materializa en la declaración Bolonia del 19 de junio de ese mismo año. En este EEES se han llevado a cabo mecanismos de evaluación de estudiantes, docentes, programas formativos, e incluso de centros e instituciones educativas. Evidencias suficientes de que la UE camina con paso firme al establecimiento de mecanismos para el rendimiento de cuentas. Sin embargo, las etapas educativas obligatorias son aún un territorio inexplorado, aunque sí existen agencias de calidad y/o evaluación a nivel nacional, no existe (aún) ninguna agencia a nivel europeo que vele por estas cuestiones.

### **Estudio comparado sobre modelos de agencias de evaluación y sus competencias**

Tras la reflexión de los apartados anteriores, la evaluación se posiciona como una de las herramientas más potentes para el desarrollo de políticas educativas basadas en evidencias. Para llevar a cabo las evaluaciones de las etapas obligatorias, los diferentes países miembros de la Unión Europea se han dotado de distintos organismos o agencias que se encargan de diseñar, implementar y realizar los informes de sus evaluaciones (Eurydice, 2015), es decir, de realizar un diagnóstico y un análisis de los sistemas educativos. Cabe preguntarse, en consecuencia, si la puesta en marcha de estas evaluaciones encamina a los sistemas educativos hacia la mejora de la calidad educativa.

Dada la importancia de las agencias de evaluación en los países de la UE, su influencia en los sistemas educativos y su impacto a nivel mediático, resulta pertinente el estudio de las competencias que estas organizaciones tienen sobre la soberanía de los países en los sistemas educativos. Así bien, resulta pertinente también cuestionarse, a la luz de esta tendencia supranacional en evaluación de los sistemas educativos, si caminamos hacia un Espacio Europeo de Educación (Matarranz, 2017) y, en su caso, si resultaría acertada la creación de una Agencia Europea de Evaluación para las etapas obligatorias de la enseñanza formal.

Cabe señalar que, para dar respuesta a estas cuestiones, nos centraremos en la evaluación de estudiantes de las etapas de enseñanza obligatoria. Si bien para futuros estudios se plantea la posibilidad de analizar a fondo la evaluación del profesorado como parte esencial del proceso de enseñanza aprendizaje y como continuación de las iniciativas que ya se están llevando a cabo en el Espacio Europeo de Educación Superior.

Para la puesta en marcha de dicho estudio se utilizará el método comparado, para lo cual, es necesario realizar una distinción de los dos momentos claves de la investigación (Caballero, Manso, Matarranz y Valle, 2016): en primer lugar, el diseño de la investigación quedaría definido con las tres primeras fases del método comparado, esto es: selección y definición del problema, formulación de hipótesis o presupuestos de partida, y elección de la unidad de análisis. En segundo lugar, el desarrollo de la investigación, compuesto por las cuatro fases que constituyen el núcleo del método comparado: fase descriptiva, fase interpretativa, fase de yuxtaposición y fase comparativa.

Así bien, a partir de la experiencia que encontramos en las agencias de calidad en la Educación Superior, creemos que encontraremos convergencias en las competencias ejercidas por los organismos a nivel federal o estatal, así como en el desarrollo de los programas de evaluación de las competencias de los estudiantes. No así en el resto de programas de evaluación, pues consideramos que, en la evaluación de docentes y centros, los Estados van a divergir tanto en los organismos responsables como en el diseño e implementación de los programas de evaluación (Eurydice, 2010; 2015).

Tabla 1. Árbol de indicadores para el estudio comparado.

Dimensiones	Parámetros	Indicadores
Unidad de Evaluación	Sentido institucional	<ul style="list-style-type: none"> <li>- Organismo responsable</li> <li>- Año de creación</li> <li>- Adscripción</li> <li>- Naturaleza</li> <li>- Funciones</li> </ul>
Programas de evaluación	Sentido de la evaluación	<ul style="list-style-type: none"> <li>- Existencia del plan general de evaluación</li> <li>- Finalidades y Objetivos de las evaluaciones</li> </ul>
	Dimensiones de la evaluación	<ul style="list-style-type: none"> <li>- Estudiantes</li> <li>- Profesores</li> <li>- Proyectos</li> <li>- Dirección educativa</li> <li>- Administración</li> <li>- Centro educativo</li> </ul>
	Evaluación de estudiantes	<ul style="list-style-type: none"> <li>- Tipo de evaluación</li> <li>- Nivel educativo</li> <li>- Cursos</li> <li>- Competencias evaluadas</li> <li>- Instrumentos de evaluación</li> <li>- Personal a cargo de las pruebas</li> <li>- Personal a cargo de las correcciones</li> <li>- Objetivos de la evaluación</li> <li>- Uso de los resultados</li> </ul>

Fuente: Elaboración propia.

La selección de los países se ha realizado tomando como base el modelo de organización y administración en materia educativa, por considerarse que este modelo es determinante para establecer diferentes políticas de evaluación a nivel territorial (Calero y Choi, 2012). En este caso, dada la conveniencia de contar con Estados con políticas descentralizadas, hemos seleccionado tres Estados en los que el nivel de autoridad regional es elevado (Hoogue, Marks y Schakel, 2010): Alemania, España e Italia. Estos tres países, según estos autores, presentan altas cotas de autogobierno o *self-rule* (el grado en que un gobierno regional puede ejercer una autoridad independiente de injerencias del poder central) y niveles de gobierno compartido o *shared-rule* (el grado en que las regiones participan en la toma de decisiones a nivel nacional, en la medida en que forman parte de la voluntad política del poder central) diferente, siendo Alemania el Estado con mayor cotas de gobierno compartido, mientras que España tiene unas cotas más modestas e Italia apenas tiene gobierno compartido. Por lo tanto, siendo Estados descentralizados mantienen características que los diferencian en el desarrollo de esa descentralización.

De estos países hemos localizado aquellos organismos a nivel federal o estatal que tienen competencias en la evaluación del sistema educativo. El análisis de dichos organismos nos va a permitir realizar un estudio comparado entre ellos con el fin de extraer convergencias y divergencias en sus competencias en evaluación y en el desarrollo de sus políticas evaluativas. Los indicadores seleccionados (Tabla 1) se recogieron por ser considerados los idóneos para la realización de la investigación propuesta, siguiendo el trabajo de Álvarez – López (2017).

Enmarcamos el análisis en un estudio comparado de carácter estático, ya que la finalidad de este estudio es realizar un análisis de las competencias que las agencias de evaluación tienen en los países en la realidad actual. Para lo cual, profundizaremos en las agencias de los países seleccionados en el año 2018.

### 3. Resultados de la investigación

Presentamos los resultados siguiendo el árbol de dimensiones, parámetros e indicadores diseñado y atendiendo a una breve síntesis descriptiva de cada Estado, para profundizar en la yuxtaposición antes de cerrar con las conclusiones y la discusión que nazca a partir de los resultados.

Para el caso de Alemania el Instituto para la Mejora de la Calidad Educativa (IQB por sus siglas en alemán), es el responsable de realizar el seguimiento y aplicación de los estándares de la educación definidos por la Conferencia

Permanente de Ministros de Educación y Asuntos Culturales de los Länder. El IQB es también uno de los principales institutos que participan activamente en la investigación educativa empírica en Alemania. Las funciones que realiza el IQB son el diseño y desarrollo de: (1) las pruebas de rendimiento y modelos de nivel de competencia (Estándares Nacionales), (2) del Estudio Nacional de Evaluación (Ländervergleiche) con el fin de determinar el grado en que se están cumpliendo los estándares educativos en los dieciséis Estados federales alemanes (Länder), y (3) de las pruebas de comparación a nivel estatal (VERA).

Tabla 2. Sentido institucional

Indicador	Alemania	Italia	España
Organismo responsable	Instituto para la Mejora de la Calidad Educativa (IQB).	Instituto Nacional para la Evaluación del Sistema Educativo y la Formación (INVALSI).	Instituto Nacional De Evaluación Educativa (INEE).
Año de creación	2004	1970	1993
Adscripción	Dependiente de la <i>Conferencia Permanente de los Ministros de Educación y Asuntos Culturales de los Länder – Kultusministerkonferenz KMK.</i>	Supervisado por el <i>Ministero dell’ Istruzione, dell’ Università e della Ricerca, Miur.</i>	Dependiente del Ministerio de Educación y Formación Profesional
Naturaleza	Organismo público académico independiente.	Organismo con personalidad jurídica de derecho público.	Órgano colegiado de evaluación dependiente de la Dirección General de Evaluación y Cooperación Territorial del Ministerio.

Fuente: Elaboración propia.

Para el caso italiano, encontramos al Instituto Nacional para la evaluación del sistema educativo de la educación y la formación (INVALSI por sus siglas en italiano), perteneciente al Ministerio de Educación, Universidad e Investigación. Entre sus funciones encontramos algunas como: seguimiento y control del sistema educativo, la coordinación y preparación de las pruebas para la evaluación externa de los resultados de aprendizaje de todos los estudiantes; asesoramiento y formación para la autoevaluación de centros y la participación en proyectos de investigación nacionales e internacionales sobre evaluación.

En España, el organismo responsable de la evaluación general del sistema educativo es el Instituto Nacional de Evaluación Educativa (INEE, por sus siglas en español). Sus funciones, que se realizan en coordinación con las Administraciones educativas, son en esencia: (1) la coordinación de las políticas de evaluación general del sistema educativo y la puesta en marcha de programas evaluativos, (2) la coordinación en las evaluaciones internacionales y en la elaboración de los indicadores internacionales de la educación, así como el seguimiento de las actuaciones de la Unión Europea en este ámbito, y (3) la elaboración del Sistema Estatal de Indicadores de la Educación, la realización de investigaciones y estudios de evaluación del sistema educativo, así como la difusión de la información de ambas actuaciones.

Así pues, las tres agencias estudiadas (pese a la diferencia temporal en que fueron creadas<sup>3</sup>), tienen carácter público y mantienen una vinculación directa con los ministerios de educación nacionales.

Tabla 3. Funciones de los organismos evaluadores

Funciones de los organismos evaluadores	Alemania	Italia	España
Diseño y aplicación de pruebas	✓	✓	✓
Seguimiento y control del sistema educativo	✓	✓	✓
Asesoramiento en autoevaluación de centros		✓	
Participación en investigaciones/estudios/informes...	✓	✓	✓
Coordinación de políticas de evaluación			✓
Elaboración de indicadores de evaluación	✓	✓	✓

Fuente: Elaboración propia.

En la Tabla 3, se observa que, gran parte de las funciones de los organismos evaluadores son compartidas por las tres agencias nacionales. Las tres instituciones coinciden en funciones tales como el diseño y aplicación de

<sup>3</sup> Cuestión que no se aborda en el presente trabajo por no ser objeto de estudio de la investigación.

pruebas de rendimiento a nivel nacional; la participación en investigaciones nacionales e internacionales sobre evaluación; la elaboración de estudios/informes fruto de las pruebas llevadas a cabo en los distintos sistemas nacionales de educación; o la elaboración de los indicadores propios de la evaluación del rendimiento en educación.

Tabla 4. Sentido de la evaluación

Indicador	Alemania	Italia	España
Existencia del plan general de evaluación	✘	✓	✘
Finalidades y Objetivos de las evaluaciones	<ul style="list-style-type: none"> <li>- Seguimiento y control de la calidad educativa nacional.</li> <li>- La revisión central de la consecución de los estándares educativos y de la eficiencia de las escuelas.</li> <li>- La participación en estudios comparativos internacionales del rendimiento de los estudiantes</li> <li>- La presentación de informes la educación conjunta de la Federación y los Länder</li> </ul>	<ul style="list-style-type: none"> <li>- Mejorar y estandarizar la calidad del sistema educativo.</li> <li>- Seguimiento y control del sistema educativo.</li> <li>- Proporcionar a los directores y maestros de escuela un instrumento útil para la auto-evaluación y para la mejora de su didáctica.</li> </ul>	<ul style="list-style-type: none"> <li>- Contribuir a mejorar la calidad y la equidad de la educación.</li> <li>- Orientar las políticas educativas.</li> <li>- Aumentar la transparencia y eficacia del sistema educativo.</li> <li>- Proporcionar información sobre el grado de consecución de los objetivos educativos españoles y europeos</li> </ul>

Fuente: Elaboración propia.

Cabe señalar que, como podemos observar en la Tabla 4, de las tres instituciones, solo una de ellas, la italiana, cuenta con un plan general de evaluación, el PON (por sus siglas Programa Operativo Nacional) para el periodo de tiempo comprendido entre el 2014 y el 2020. Más allá de la existencia (o no) de estos planes generales, nos encontramos con un sentido de la evaluación común en los tres: el diseño y la aplicación de pruebas de rendimiento, la elaboración de informes o estudios, el establecimiento y revisión de estándares educativos, el seguimiento y el control del sistema educativo, orientación en políticas educativas nacionales y la participación en investigaciones bien nacionales, bien internacionales. En definitiva, las finalidades y los objetivos de todas ellas tienen como meta común contribuir a la mejora de la calidad educativa.

Tabla 5. Dimensiones de la evaluación

INDICADOR	Alemania	Italia	España
Estudiantes	✓	✓	✓
Profesores	✓		
Proyectos		✓	✓
Dirección educativa		✓	
Administración			
Centro educativo		✓	
Educación de adultos		✓	

Fuente: Elaboración propia.

La Tabla 5 evidencia la sintonía entre Estados a la hora de evaluar a los estudiantes. Sin embargo, no se vuelve a encontrar otra dimensión que sea compartida por las tres instituciones nacionales analizadas. Mientras únicamente Alemania evalúa a sus docentes, Italia y España se preocupan por los proyectos. Italia se desmarca con una serie de programas de evaluación de dimensiones como la función directiva o los centros educativos, aspectos que no son abordados por Alemania o España, al menos en el nivel administrativo central.

Tabla 6. Evaluación de estudiantes<sup>456</sup>

Indicador	Alemania	Italia	España
Tipo de evaluación <sup>4</sup>	Estandarizada Sumativa	Estandarizada Sumativa	Estandarizada Sumativa
Nivel educativo (CINE <sup>5</sup> )	- CINE 2	CINE 1 CINE 2	CINE 1 CINE 2
Cursos	Último curso de educación secundaria.	Muestras de estudiantes de 2º y 5º de Primaria 1º y 3º año Secundaria inferior (Obligatoria en 3º)	Muestras de estudiantes de 6º de Primaria 4º año Secundaria
Competencias evaluadas	Competencia matemática Competencia lingüística alemana. Competencia lingüística inglesa.	Competencia matemática Competencia lingüística italiana	Competencia matemática Competencia lingüística castellana Competencia científica
Instrumentos de evaluación	Cuestionario con preguntas de elección múltiple, de respuesta corta o desarrollo y abiertas.	Cuestionario con preguntas de elección múltiple.	Cuestionario con preguntas de elección múltiple, de respuesta corta o desarrollo y abiertas.
Personal a cargo de las pruebas	Profesores del centro escolar	Personas externas al centro <sup>3</sup>	Personas externas al centro
Personal a cargo de las correcciones	Profesores del centro escolar Personas externas al centro	- Personas externas al centro	- Personas externas al centro
Objetivos de la evaluación	<ul style="list-style-type: none"> <li>- Tomar decisiones sobre la trayectoria educativa de los estudiantes.</li> <li>- Garantizar la calidad de los centros y las prácticas educativas.</li> <li>- Realizar comparaciones entre los <i>Länder</i></li> </ul>	<ul style="list-style-type: none"> <li>- Seguimiento y control del sistema educativo.</li> <li>- Analizar el rendimiento del alumnado para la mejora de los niveles de aprendizaje.</li> <li>- Tomar decisiones sobre la trayectoria educativa de los estudiantes.</li> </ul>	<ul style="list-style-type: none"> <li>- Seguimiento y control del sistema educativo y de los centros escolares.</li> <li>- Obtención de resultados de los indicadores sobre los estudiantes y los centros.</li> <li>- Obtención del Graduado en ESO.</li> </ul>
Uso de los resultados	<ul style="list-style-type: none"> <li>- Muestra de los resultados de aprendizaje del alumnado.</li> <li>- Obtención de certificación.</li> <li>- Seguimiento a lo largo del tiempo de los resultados en los indicadores.</li> </ul>	<ul style="list-style-type: none"> <li>- Obtención de la certificación del nivel CINE 2.</li> <li>- Conocer el nivel de rendimiento de los estudiantes.</li> <li>- Realizar (indirectamente) evaluaciones a los directores de centro y los docentes.</li> <li>- Autonomía de los centros para difundir los resultados.</li> </ul>	<ul style="list-style-type: none"> <li>- Revisión y mejora de los sistemas educativos y de los centros escolares en función de los resultados obtenidos.</li> <li>- Obtención de certificación</li> </ul>

Fuente: Elaboración propia.

La yuxtaposición presentada en la Tabla 6 nos da información sobre la evaluación de estudiantes. Las sinergias entre Estados son constantes. Encontramos una convergencia bastante alta entre los programas de evaluación de estudiantes: el tipo de evaluación, los objetivos que persiguen, los niveles elegidos, las competencias analizadas, los instrumentos utilizados, el tipo de personal a cargo de las pruebas y de las correcciones, así como el uso de los resultados. Sin que el estudio se centre en ello, el grado de similitud entre Estados podemos explicarlos (al menos en parte) en la influencia de los programas internacionales de evaluación de estudiantes. Sin embargo, este es un punto que se escapa a las intenciones de este trabajo.

A modo de resumen, percibimos lineamientos en las tres agencias de evaluación educativa. Se observan analogías significativas en las funciones y tareas que desempeñan las tres agencias, en la descripción de la finalidad y los objetivos que persiguen, así como el sentido de la evaluación que cada una de ellas tiene. En lo que respecta a la evaluación de estudiantes (dada la discrepancia del alcance que tiene la agencia italiana con respecto a la alemana y la española) la convergencia es clara, pues las tres agencias comparten elementos en todos los indicadores analizados: cursos en que los estudiantes son evaluados, competencias que se evalúan, instrumentos utilizados para ellos, personal que interviene, objetivos y usos de los resultados.

Rescatando la premisa del presente trabajo, ya descrita en el apartado de metodología, resulta evidente que existen convergencias en las competencias ejercidas por los organismos a nivel federal o estatal, así como en el desarrollo de los programas de evaluación de las competencias de los estudiantes.

<sup>4</sup> Tres tipos de evaluación: sumativa, estandarizada o formativa. (Eurydice, 2010.)

<sup>5</sup> Clasificación Internacional Normalizada de Educación CINE 1: Educación primaria / CINE 2: Educación secundaria inferior.

<sup>6</sup> Personas designadas por los órganos nacionales responsables de la organización de las pruebas, o por empresas externas contratadas para este fin. (Fuente: Eurydice, 2010)



#### 4. Discusión y conclusiones

A la luz del trabajo realizado, y como ya se indicó previamente, cabe reflexionar sobre dos cuestiones fundamentales.

En primer lugar, respecto al modelo de evaluación que aplican las agencias estudiadas, y si estas evaluaciones mejoran la calidad de sus respectivos sistemas educativos, creemos conveniente reflexionar sobre las siguientes propuestas:

*Sería positivo (y necesario) repensar la calidad en educación, qué entendemos por esta, y cómo la evaluación puede contribuir a su mejora.*

Al ser la calidad educativa un término que tiene tantas concepciones como puntos de vista, (tal y como ya vimos en el marco teórico) definir los indicadores que permitan evaluarla resulta un tanto inverosímil, concluimos por ello que el término de calidad tiene, innegablemente, connotaciones políticas.

No obstante, bajo una perspectiva puramente pedagógica, se hace imprescindible contemplar la evaluación como una herramienta que debe permitir la toma de decisiones para la mejora, dando un mayor protagonismo a los participantes, y al mismo tiempo, permitiendo a los responsables identificar las necesidades que se deben atender y reconocer cuáles son sus debilidades y sus fortalezas. Cabe, por tanto, esperar que una mejora de la calidad educativa pase por una serie de evaluaciones enfocadas a la mejora. Sin embargo, nos encontramos con que el modelo de evaluación principal en las agencias de evaluación de los países estudiados ha sido estandarizado y sumativo.

En este punto, se hace imprescindible plantear la necesidad de analizar qué concepto de calidad educativa subyace a cada uno de los programas de evaluación puestos en marcha por las distintas agencias nacionales analizadas. Ciertamente los controles, las evaluaciones externas, la rendición de cuentas... tienen sentido si todos caminamos hacia la mejora en educación, a la optimización de los procesos de aprendizaje para crear ciudadanos preparados y con las competencias necesarias para desenvolverse de manera óptima a lo largo de su vida. Tener estas cuestiones claras no es baladí, nos ayuda a clarificar qué es aquello que queremos alcanzar en educación, y a definir (o redefinir en su caso) qué programas se necesitan para su logro.

*Cabe replantearse los modelos de evaluación dominantes*

Al hilo de esta cuestión, detectamos la necesidad de realizar un profundo análisis y una reflexión crítica y constructiva sobre el modelo de evaluación que estamos fomentando. Observamos un modelo de evaluación centrado en los resultados como único indicador de calidad. Olvidamos que la calidad educativa no es exclusivamente sinónimo de excelencia académica, entender así el concepto de calidad se nos antoja de un empobrecimiento tal, que seguir en estos procesos puede llegar a ser contraproducente para lo que se busca alcanzar. No debemos perder de vista el discurso de competitividad que también tiene cabida en la puesta en marcha de evaluaciones. Los famosos rankings de centros escolares según las pruebas de rendimiento de sus estudiantes, el informe PISA que evalúa competencias de estudiantes a nivel internacional, o incluso los rankings de universidades; cada uno con su sistema de indicadores definido “a la carta” y fomentando en cierto sentido la mercantilización de la educación. Nosotros creemos que sería más conveniente tender hacia una concepción sistémica de la evaluación (Álvarez – López, 2017) en la que se tengan en cuenta otras dimensiones, metodologías complementarias, estrategias de recogida de información y de análisis de corte cualitativo y un claro enfoque democrático, equitativo, formativo y emancipador de la evaluación de los sistemas educativos.

En segundo lugar, nos planteamos una cuestión ya abordada por otros autores (Valle, 2006; Matarranz, 2017): si caminamos hacia un Espacio Europeo de Educación y, dado el objeto de estudio de la presente investigación, si resultaría acertada la creación de una Agencia Europea de Evaluación para las etapas obligatorias de la enseñanza formal. Tras esta investigación, estamos en condiciones de afirmar que caminamos hacia un Espacio Europeo de Educación (Valle, 2006; Matarranz, 2017), dentro del cual la puesta en marcha de programas de evaluación a nivel nacional son pequeños pasos que (en compañía de la puesta en marcha de otros mecanismos) nos acercan a la mejora de los sistemas educativos.

Programas de evaluación que, como se ha demostrado, tienen grandes elementos convergentes entre Estados. Estas sinergias nos invitan a proponer la creación de una Agencia Europea de Evaluación, cuyo objetivo esencial sea la mejora de la educación y el desarrollo de una cultura de la calidad en la etapa de enseñanza obligatoria a nivel europeo; así como impulsar el desarrollo de la garantía de calidad a nivel supranacional dando apoyo y seguimiento a los sistemas educativos nacionales y proporcionando servicios integrales de redes y oportunidades.

La creación de esta agencia que proponemos creemos que tiene numerosas ventajas, entre las que podemos destacar (i) facilidad de la comparabilidad entre países miembros, (ii) coordinación de las políticas de evaluación educativa a nivel europeo, (iii) análisis y perspectiva supranacional de la educación obligatoria, (iv) desarrollo de programas de evaluación europeos, (v) posibilidad de reconocer títulos en la etapa de educación obligatoria entre países europeos.

Sin embargo, esta propuesta no está exenta de peligros. Por una parte, la evaluación tiene una vertiente de cercanía al objeto evaluador que no se debe perder de vista. Por otra parte, los Estados están obligados a informar a su ciudadanía sobre el estado de las políticas públicas de las que son garantes. Ambas ideas no son contradictorias si creemos que Europa es una oportunidad de crecimiento político y social.

La europeización de la educación, a la que contribuye la creación de una Agencia Europea de Evaluación, es una oportunidad que no debemos dejar escapar dados los retos que nos plantea la globalización económica y cultural en una perspectiva internacional. Construir Europa es afianzar los lazos que nos unen como ciudadanos de paz y de progreso.

## 5. Referencias bibliográficas

- Adams Jr., J. E., & Hill, P. T. (2006). Educational *accountability* in a regulated market. *Peabody Journal of Education*, 81(1), 217-235. DOI: 10.1207/S15327930pje8101\_10
- Álvarez – López, G. (2017). *La evaluación de la educación básica en el sistema educativo español: estudio comparado en el ámbito de las Comunidades Autónomas* (Tesis Doctoral). Universidad Autónoma de Madrid.
- Amrein, A. L. y Berliner, D. C. (2003). The effects of high-stakes testing on student motivation and learning. *Educational Leadership*, 60(5), 32-38.
- Au, W. y Tempel, M. B. (2012). *Pencils down: Rethinking high-stakes testing and accountability in public schools* (First ed.). Milwaukee: Rethinking Schools.
- Báez, B. (2007). Evaluación institucional: disciplina y retórica en la antepenúltima oleada de reformismo educativo a gran escala. *Tempora*, 10, 183-213.
- Baker, E. (2009). Consideraciones de validez prioritaria para la evaluación formativa y de rendición de cuentas. *Revista de Educación*, (348), 91-110.
- Ball, S. (2014). Globalización, mercantilización y privatización: tendencias internacionales en Educación y Política Educativa. *Education Policy Analysis Archives*, 22 (41). Recuperado de: <http://dx.doi.org/10.14507/epaa.v22n41.2014>
- Berryhill, J., Linney, J. A., & Fromewick, J. (2009). The effects of education *accountability* on teachers: Are policies too-stress provoking for their own good? *International Journal of Education Policy & Leadership*, 4(5), 1-14.
- Bonnet, G. (ed.) (2002). *The assessment of pupils' skills in English in eight european countries*. Brussels: European Network of Policy Makers for the Evaluation of Education Systems.
- Burbules, N. y Torres, C. A. (2001). Globalización y Educación. *Revista de Educación*, nº extraordinario, 13-29
- Caballero, Á., Manso, J., Matarranz, M. y Valle, J. M. (2016). Investigación en Educación Comparada: pistas para investigadores noveles. *Revista Latinoamericana de Educación Comparada*, 39, 39-56.
- Calero, J. y Choi, A. (2012). La evaluación como instrumento de política educativa. *Presupuesto y Gasto Público*, 67, 29-42.
- Comisión Europea (2012). *First European Survey on Language Competences. Final Report*. Bruselas: Comisión Europea.
- Consejo (2002). Programa de trabajo detallado para el seguimiento de los objetivos concretos de los sistemas de educación y formación en Europa. Diario Oficial, C(142), de 14 de junio de 2002.
- Consejo (2009). Conclusiones del Consejo de 12 de mayo de 2009 sobre un marco estratégico para la cooperación europea en el ámbito de la educación y la formación (“ET 2020”). Diario Oficial, C (119), de 28 de mayo de 2009.
- Darling – Hammond, L. (2007). Race, inequality and educational *accountability*: the irony of ‘No Child Left Behind’. *Race Ethnicity and Education*, 10 (3), 245-260.
- Davies, M., & Elliott, S. N. (2012). Inclusive assessment and *accountability*: Policy to evidence-based practices. *International Journal of Disability Development and Education*, 59(1), 1-6. doi:10.1080/1034912X.2012.654932
- Diamond, J. B. (2012). Accountability policy, school organization, and classroom practice: Partial recoupling and educational opportunity. *Education and Urban Society*, 44(2), 151-182. doi:10.1177/0013124511431569
- Dorn, S., & Ydesen, C. (2014). Towards a comparative and international history of school testing and accountability. *Education Policy Analysis Archives*, 22(115). <http://dx.doi.org/10.14507/epaa.v22.1913>
- Eurydice (2010). *Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de resultados*. Madrid: Ministerio de Educación – Secretaría General Técnica.
- Eurydice (2015). *Assuring Quality in Education. Policies and Approaches to School Evaluation in Europe*. Luxembourg: Publications Office of the European Union.
- García Garrido, J.L. y García Ruíz, M.J. (2005). *Temas candentes de la educación en el siglo XXI*. Madrid: Ediciones Académicas.
- Gargallo, B. (2003). Un modelo de evaluación del sistema educativo. *Revista Española De Pedagogía*, 61 (224), 33-59.
- Holme, J. J., Richards, M. P., Jimerson, J. B., & Cohen, R. W. (2010). Assessing the effects of high school exit examinations. *Review of Educational Research*, 80(4), 476-526.
- Hoogue, L., Marks, G. N. and Schakel, A.H. (2010). *The Rise of Regional Authority. A comparative study of 42 democracies*. London: Routledge.
- Hutt, E. (2014). The GED and the Rise of Contextless Accountability. *Teachers College Record*, 116 (9), 1-20.

- Kamens, D. H. y Benavot, A. (2011). National, regional and international learning assessments: Trends among developing countries, 1960-2009. *Globalisation, Societies and Education*, 9(2), 285-300
- Kamens D.H. y Mcneely, C. L. (2010). Globalisation and the growth of International Educational Testing and Nacional Assessment. *Comparative Education Review*, 54 (1), 5-25
- Koretz, D. (2008). Test-based educational *accountability*. *Zeitschrift Fur Padagogik*, 54(6), 777-790.
- Kostogriz, A., & Doecke, B. (2011). Standards-based *accountability*: Reification, responsibility and the ethical subject. *Teaching Education*, 22(4), 397-412.
- Labaree, D. (2014). Let's Measure What No One Teaches: PISA, NCLB, and the Shrinking Aims of Education. *Teachers College Record*, 116 (9),1-14.
- Lai, Y. & Wang, L. (2014). Educational *accountability* system in local government: Current conditions and impact. *Journal of Research in Education Sciences*, 59(3), 97-132. doi:10.6209/JORIES.2014.59(3).04
- Lee, J., y Wong, K. K. (2004). The impact of *accountability* on racial and socioeconomic equity: Considering both school resources and achievement outcomes. *American Educational Research Journal*, 41(4), 797-832. doi:10.3102/00028312041004797
- Linn, R. L. (2008). Methodological issues in achieving school *accountability*. *Journal of Curriculum Studies*, 40(6), 699-711. doi:10.1080/00220270802105729
- Martínez Rizo, F. (2009). Evaluación formativa en aula y evaluación a gran escala: Hacia un sistema más equilibrado. *Revista electrónica de investigación Educativa*, 11(2), 1-18
- Matarranz, M. (2017). *Análisis supranacional de la Política Educativa de la Unión Europea 2000-2015. Hacia un Espacio Europeo de Educación* (Tesis Doctoral). Universidad Autónoma de Madrid.
- Muñoz-Repiso, M. y Murillo, F.J. (2010). Un balance provisional sobre la calidad en educación: eficacia escolar y mejora de la escuela. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 8(2), pp. 177-186.
- Pérez Juste, R. (2005). Calidad de la educación, calidad en la educación: Hacia su necesaria integración. *Educación XXI*, (8), 11-33.
- Seibold, J. R. (2000). La calidad integral en educación. Reflexiones sobre un nuevo concepto de calidad educativa que integre valores y equidad educativa. *Revista Iberoamericana De Educación*, 23, 215-231.
- Smith, W.C. (2014). The global transformation toward testing for accountability. *Education Policy Analysis Archives*, 22 (116). <http://dx.doi.org/10.14507/epaa.v22.1571>
- Sobe, N. W. (2015). All that is global in not world culture: Accountability systems and educational apparatuses. *Globalisation, Societies and Education*, 13 (1), 135-148. <http://dx.doi.org/10.1080/14767724.2014.967501>
- Stake, R. E. (2006). *Evaluación comprensiva y evaluación basada en estándares*. Barcelona: Ed. Graó.
- Stecher, B. M. (2002). Consequences of large-scale, high-stakes testing on school and classroom practice. En L. S. Hamilton, B. M. Stecher, & S. P. Klein (Eds.), *Making sense of test-based accountability in education* (pp. 79-100). Santa Monica: Rand Corporation.
- Tiana, A. (2006). La evaluación de la calidad de la educación: Conceptos, modelos e instrumentos. *Transatlántica de Educación*, (1), 19-30.
- Tuck, E. (2013). Neoliberalism as nihilism? A commentary on educational *accountability*, teacher education, and school reform. *Journal for Critical Education Policy Studies*, 11 (2), 324-347.
- Valle, J. M. (2006). *La Unión Europea y su política educativa*. 2 tomos. Madrid: MEC