

Evaluación, Responsabilidad y Mejora educativa

Red por el Diálogo Educativo



“De vez en cuando, el labrador se detiene, al llegar al extremo del surco, para resoplar; pero también para contemplar la labor realizada, para sopesar lo que no salió como estaba previsto e intentar corregirlo y, al mismo tiempo, retomar coraje”.

Célèstin Freinet

I.	Introducción.....	2
II.	Diagnóstico de la situación de la evaluación educativa en España	3
2.1.	La evaluación interna en España	3
	La evaluación interna de centro	3
	La evaluación interna del docente	4
	La evaluación interna de los aprendizajes	5
2.2.	La evaluación externa en España	8
III.	Una visión de un modelo de evaluación de la educación: objeto y objetivos.	11
IV.	Hacia un modelo de evaluación del sistema educativo	13
4.1.	Claves para un modelo de evaluación interna	14
4.2.	Claves para un modelo flexible y compartido de evaluación externa	17
V.	La transición a un nuevo modelo de evaluación	19
VI.	Glosario de evaluación de los aprendizajes.....	21
VII.	Bibliografía.....	23

I. Introducción

Los sistemas educativos tienen como finalidad lograr que cualquier persona pueda hacer un uso efectivo de su derecho a la educación, partiendo de la base de que todos pueden aprender y lograr el éxito educativo. En su vertiente más reciente de los Objetivos del Desarrollo Sostenible se trata de lograr una *educación equitativa, inclusiva y de calidad* para todos. Para ello, la capacidad del sistema educativo debe contribuir de forma decisiva a lograr estos objetivos, que solo pueden garantizarse mediante la información generada por la dinámica de evaluación del propio sistema educativo, ya sean procesos o resultados de aprendizaje. Son esta capacidad evaluativa y el modelo de responsabilidad que lleva aparejado las que sirven a la vez como motor y guía para la mejora educativa de todos los centros educativos.

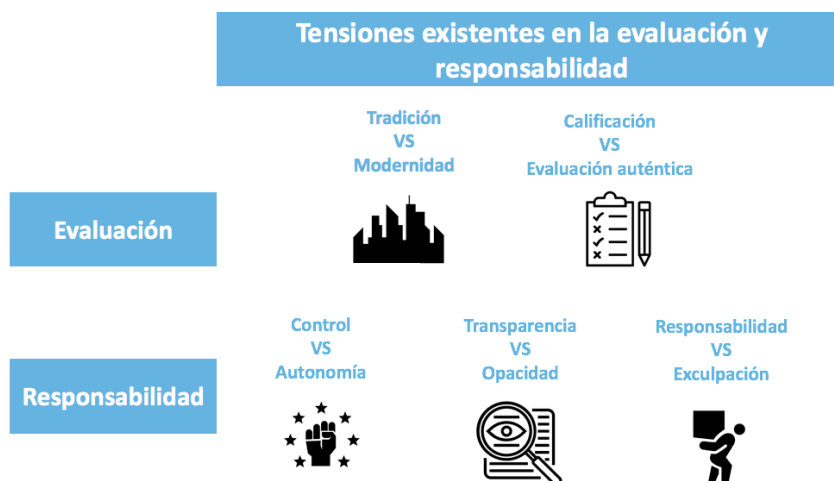
En los últimos años se han producido numerosos debates asociados a la evaluación de la educación fruto de la aparición de evaluaciones externas (autonómicas, nacionales e internacionales) que han ocupado un espacio central, en ocasiones excesivo, en la esfera pública. La falta de coherencia, unida a decisiones superpuestas de política educativa poco o nada consensuadas con la comunidad educativa, han dañado la institucionalización de la evaluación educativa como un pilar fundamental del sistema. Es por ello que la controversia en torno a la existencia de estas evaluaciones ha ido en aumento y a día de hoy son varias las posturas abiertamente contrarias a la existencia de evaluaciones externas, sea cual sea su formato, uso o finalidad. De hecho, estas actitudes y cuestionamientos podrían estar permeando en otros ámbitos del modelo educativo, lo cuál puede llevar al rechazo o inmovilismo frente al cambio. Por ello, resulta una tarea ineludible definir y clarificar los distintos formatos de la evaluación educativa, no solo externa sino también interna al centro, y argumentar de qué manera pueden (y de qué manera es probable que no logren) contribuir decisivamente a la mejora educativa.

Por la historia reciente, la evaluación educativa es quizás uno de los pilares del sistema que más reflexión y discusión requiere para poder lograr un consenso mínimo suficiente que permita al sistema avanzar hacia sus objetivos. En política educativa, es tan relevante señalar el objeto como el objetivo, a la vez que la conexión que tiene con los otros elementos que configuran la arquitectura general del sistema. Por ejemplo, el papel que juega la evaluación interna de los centros educativos es muy relevante en el proceso de aprendizaje, en el desarrollo y mejora de la competencia profesional del docente y del conjunto de programas (educativos y de organización y gestión) de los centros educativos. Pero la realidad educativa es compleja, y por tanto no existe un único instrumento de evaluación (ya sea interna o externa) capaz de generar una responsabilidad hacia la mejora y una información equilibrada y holística, cuantitativa y cualitativa.

Con todo ello, el objetivo de este documento es presentar y profundizar en esta cuestión, realizar un breve diagnóstico de la situación de la evaluación educativa en España en su vertiente interna y externa. Tras eso, se plantea una visión compartida que permita proyectar y construir un nuevo modelo ambicioso pero alcanzable en un periodo razonable de tiempo, respetando la autonomía de las administraciones y los centros educativos para construir su camino hacia la mejora educativa. Esta tarea no está exenta de transitar tensiones existentes en torno a la temática en cuestión y donde es deseable encontrar un equilibrio razonable: estas van desde el control a la autonomía, desde la transparencia hacia la opacidad en cuanto a la evaluación,

desde la tradición a la modernidad en cuanto a la evaluación, desde la calificación a evaluación por competencias, o desde la responsabilidad a la exculpación en la evaluación.

Gráfico 1. Las tensiones en la evaluación y la responsabilidad en la educación



II. Diagnóstico de la situación de la evaluación educativa en España

2.1. La evaluación interna en España

En el seno de todo centro educativo existen tres grandes dimensiones en torno a las cuales giran los planes, proyectos y programas emprendidos: a) la **gestión y organización** del centro educativo, b) el desarrollo de la **capacidad profesional docente** y c) el **aprendizaje del alumnado**. Hasta ahora¹, las iniciativas en torno a una evaluación interna de centro (o de los programas del centro), del profesorado o del alumnado, o bien no se han desarrollado lo suficiente o bien no se han adaptado a los últimos cambios y necesidades del sistema. Es en estas tres dimensiones (centro, profesorado, alumnado) donde se realiza un breve diagnóstico de la evaluación interna, en línea con los marcos de análisis e investigación educativa nacionales e internacionales dirigidos a la consecución de una educación inclusiva, equitativa y de calidad.

La evaluación interna de centro

De los estudios realizados por la Unión Europea (2004, 2015) se advierte que España define la evaluación interna como actividad obligatoria y de periodicidad anual. A esta actividad se le presupone la implicación de toda la comunidad educativa, movilizandolos numerosos recursos de apoyo (incluida la inspección educativa) y siempre con el objeto de mejorar la Programación General Anual (PGA) del centro educativo. Aún así, salvo alguna excepción, la repercusión de estas evaluaciones es tímida y a día de hoy, no existe un modelo para la evaluación interna de centro (ya sea prescriptivo u orientativo) en el ámbito nacional derivado de su trayectoria normativa.

¹ Se reconocen los intentos realizados con la LOPEGCE (1995) y acciones antecesoras como el Plan EVA iniciado en 1992.

Con respecto a la evaluación interna de centro, se pueden destacar las siguientes fortalezas del sistema educativo español:

- Tiene estructuras y órganos institucionalizados (centros educativos, profesión docente) para su realización efectiva, y dispone del apoyo externo de recursos personales (inspección educativa), económicos (en algún caso) y formación (dirigida especialmente a equipos directivos) para su desarrollo efectivo.
- Está vinculada al desarrollo de la autonomía del centro y a los procesos de innovación. Por tanto, tiene cabida y sentido en cuanto a su concreción en una programación anual y un proyecto de centro, aunque pueda resultar compleja.

A su vez, se identifican las siguientes debilidades del modelo de evaluación interna del centro en el sistema educativo español:

- No se dispone de indicadores o criterios que permitan a la evaluación interna (siga un enfoque cuantitativo, cualitativo o mixto) dejar de ser una valoración no guiada. Este es uno de los principales rasgos que lo diferencia de otros modelos europeos de evaluación interna de centros. Dicho de otro modo, se carece de marcos o dimensiones previas donde centrar la evaluación, salvo en alguna comunidad autónoma². En resumidas cuentas, no existe un modelo y un marco de evaluación interna de centro que cumpla el papel de auto-evaluación que haya permeado, ni normativa ni institucionalmente.
- Se enfrenta a dificultades prácticas como la delimitación de espacios y tiempos para la realización efectiva del proceso de evaluación, que suele ser generalmente postergado por otras tareas de gestión y organización del centro.
- Como consecuencia de una falta de un modelo, marco, procedimientos e instrumentos de evaluación comunes facilitados por la administración educativa, no se ve reforzada por el papel de la comunidad educativa en el proceso de evaluación (profesorado, alumnado y familias) a través de sus órganos colegiados o individualmente.
- No existe una clara vinculación con modelos de evaluación externa en un sentido asociado a la mejora (Bolívar, 2016), mediante el cuál podría profundizarse en el papel de la autonomía de centro, y cómo ésta podría articularse con la responsabilidad pública y social de los centros educativos.

La evaluación interna del docente

Los procesos de evaluación interna del docente suelen estar vinculados al desarrollo de su capacidad profesional en el centro educativo (esto es, de qué manera el docente contribuye al desarrollo de los objetivos del centro) y la facilitación de información sobre su rendimiento individual, con una orientación a varias finalidades:

- (i) Por un lado, a la mejora de su acción en la gestión, organización y participación;

² Ver Guía de autoevaluación y mejora (2007) para Asturias o Guía orientativa del proceso de autoevaluación y mejora de los centros educativos (2015) para Andalucía.

- (ii) Por otro, a los procesos formativos que repercuten en la mejora de su acción pedagógica y práctica docente, dimensión que cobra especial importancia en el desarrollo de los aprendizajes del alumnado.

En esta dimensión de la evaluación interna **todo queda por hacer**³. Según el informe de Eurydice (2018) *La profesión docente en Europa*, España carece de un sistema de evaluación interna del docente con frecuencia regular y no existen marcos generales que demanden su aplicación. Esto implica que la evaluación interna, por extensión, quede al albur de los centros, sin referencias ni guías que incluyan finalidades que puede cobrar la evaluación, la cultura evaluativa de la propia práctica docente y sus funciones, los recursos, los agentes de participación o los procedimientos e instrumentos.

Esta situación contrasta con el avance de muchos modelos a nivel internacional, como Reino Unido o Suecia, donde se hace una evaluación exclusivamente interna por parte de los directores escolares, tutores y/o miembros del consejo escolar (Eurydice, 2018). O por ejemplo, el *Teacher Performance Appraisal System* de Ontario (Canadá) para docentes con experiencia, donde la evaluación en el primer caso se centra en la evaluación que hace el director en base a un marco de 16 competencias, indicadores de desempeño y criterios de evaluación (OECD, 2009). En alguno de estos ejemplos, como el caso de Suecia, se cuenta obligatoriamente con programas de formación dirigidos a los directores de centro que fortalecen la consistencia, justicia y utilidad del sistema de evaluación.

La evaluación interna de los aprendizajes

En todo sistema educativo, la evaluación de los aprendizajes está fuertemente vinculada al tipo de aprendizajes que se promueven en la escuela, esto es, al currículum. Paralelamente, desde hace varias décadas, los aprendizajes transitan de un conocimiento enciclopédico y acumulativo a la irrupción de la era de la información y la comunicación en todos los ámbitos de la vida, también el escolar. Esto ha supuesto una transformación de la organización curricular de buena parte de los sistemas educativos avanzados. En algunos países, los cambios curriculares propiciaron avances en el campo de la evaluación de los aprendizajes, por ejemplo con el desarrollo de instrumentos como las rúbricas dentro del enfoque de **evaluación por competencias**. Estos planteaban una transición de un **currículum centrado en contenidos y competencias básicas o mínimas** (vinculadas a materias y evaluadas con pruebas objetivas) a un **currículum centrado en competencias a lo largo de la vida** (que conllevara aplicación de conocimientos académicos, toma de decisiones, resolución de problemas, capacidad de comunicación, aprender a aprender, trabajo cooperativo, autogestión) de carácter más transversal, donde cobra importancia en la evaluación el proceso, el contexto y la aplicación a contextos auténticos (Chatterji, 2003).

Durante las últimas décadas, en España, se han desarrollado cambios en la estructura curricular y, de manera más acusada, en las metodologías de enseñanza. Dicho esto, en muchos casos, los cambios curriculares siguen en ocasiones atrapados por su propia incapacidad de superar un enfoque de contenidos vinculados a compartimentos estancos de conocimiento (esto es, a las asignaturas). Como consecuencia, o quizás como causa, estas dificultades de transición curricular se topan habitualmente con usos exclusivos de la evaluación sumativa, basada en la falsa idea de las medidas de aprendizaje orientadas a la calificación. Estas medidas, formuladas a partir de medias aritméticas o ponderadas de notas parciales, actitud (una dimensión competencial desconocida en criterios de evaluación) y otros, además de pervertir el propósito del aprendizaje, carecen de consistencia o fiabilidad. Es por tanto necesario reconocer que, a

³ También en lo que se refiere al enfoque de evaluación externa.

día de hoy, **la evaluación interna de los aprendizajes en España tiene como debilidad principal** un fuerte arraigo a una tradición cultural basada en prácticas y usos específicos asociados a la calificación de modelos curriculares anteriores que no acaban de desaparecer.

En resumen, se destacan las siguientes fortalezas de la evaluación de los aprendizajes en el centro:

- Un cambio metodológico importante desde el punto de vista pedagógico, propiciado por la introducción de otros elementos a considerar en el aprendizaje más allá del conocimiento y los conceptos.
- El deseo del profesorado de adoptar otras formas de evaluación adaptadas a los nuevos aprendizajes, como la participación del alumnado en la evaluación de los aprendizajes, los procesos de co-evaluación y auto-evaluación y el uso de instrumentos como las rúbricas o escalas de valoración. Con esto se quiere enfatizar, no que haya cambiado la finalidad de la evaluación dominante (sumativa), sino que se incorpora poco a poco la finalidad formativa.
- El uso de las tecnologías para la evaluación está cada vez más incorporados al proceso de enseñanza y aprendizaje.

Como debilidades de la evaluación interna, se destacan:

- Un enfoque de evaluación de larga tradición, ajustado a un tipo de metodología más transmisiva y unidireccional, y de difícil encaje en otros modelos ya incorporados. A pesar de la progresiva utilización de instrumentos, finalidad o agentes nuevos al proceso de evaluación de los aprendizajes, sigue pesando la finalidad sumativa y los procedimientos para asignar calificaciones y la utilización de pruebas escritas como fuente de información más importante del aprendizaje. Estas están casi siempre vinculadas sólo a un aspecto de éste: el conocimiento de hechos, conceptos y procedimientos. Más en detalle:
 - Se adolece del **enfoque criterial** como base para diseñar instrumentos de evaluación, sobre el cual hay gran conocimiento científico no sólo aplicado al campo de la psicometría sino también al caso escolar⁴.
 - Se siguen utilizando pruebas descontextualizadas (generalmente escritas) como aspecto comprobable del aprendizaje y se dejan de evaluar o al menos no con el mismo “peso” proyectos, tareas integradas, servicios a la comunidad, trabajo colaborativo, o cualquier forma de demostrar aprendizaje que implique **no sólo conocimientos**, sino **también destrezas** (saber cómo), **actitudes** y **valores** (saber ser y estar).
 - Sigue sin incorporarse, o se hace de manera muy tímida, al alumnado en los propios procesos de evaluación: construcción de instrumentos como las rúbricas, discusión de los criterios utilizados para valorar un producto o desempeño. La **evaluación se convierte así en un instrumento de poder que se usa para otros fines no acordados**, por ejemplo, el control de la disciplina o el cumplimiento de tareas.

⁴ Ejemplos de pruebas escritas (o por ordenador) con este enfoque son las evaluaciones de diagnóstico realizadas en las distintas comunidades autónomas o las elaboradas por el Ministerio de Educación.

- Como consecuencia, y principalmente en secundaria, existe **un abuso del suspenso y la no superación** de asignaturas que tienen como peor consecuencia la elevada repetición de curso⁵, que daña notablemente la eficiencia y la igualdad de oportunidades del sistema educativo.
- Un modelo de prueba de acceso a la Universidad (EBAU) que tiene efectos nocivos sobre las dinámicas de aprendizaje previas y posteriores. Podría argumentarse que el Bachillerato es a día de hoy una etapa sin un claro fin en sí misma y donde el entrenamiento para la prueba (el conocido efecto *teaching to the test*) es un punto central, la identidad, de la etapa. No se disponen de estudios científicos que pueden sustentar alternativas al modelo de prueba de acceso a la Universidad, que documenten sus efectos sobre el Bachillerato y la ESO, y su relación con los estudios universitarios, que permiten un cuestionamiento y propuestas de alternativas tanto del sistema de acceso a la Universidad, como de los propios instrumentos utilizados y el uso de calificaciones y evaluaciones⁶.
- Un sistema de doble foco, donde se propone evaluar asignaturas y competencias clave, pero con desigual estructuración. Se incide en la configuración del dominio de evaluación, de la organización escolar, el sistema de promoción/repetición o los instrumentos por asignaturas, pero no en las competencias clave, que quedan en un segundo plano. Esto implica que las competencias clave se presentan como un elemento prescriptivo del currículum que no es funcional respecto a **la organización escolar** (todo está pensado para ser abordado desde un enfoque disciplinar), **la calificación y sus usos** (no sirven para la promoción o titulación, las calificaciones se dan por asignaturas y el tránsito a etapas no obligatorias hace uso exclusivo de éstas) o **el desarrollo de un dominio evaluable** (con la especificación de criterios, tareas de evaluación e instrumentos). Todo ello ha propiciado una desvalorización de las competencias clave y el papel que juegan en la evaluación de los aprendizajes. Los intentos de abordar la evaluación simultánea de asignaturas y competencias clave se han topado con dificultades técnicas⁷.

⁵ En España, la tasa anual de repetición de curso se dispara en la transición de primaria (donde no supera el 3% anual) a ESO (llegando casi al 10% anual⁵) de una manera desproporcionada y no acorde con la mayor parte de los sistemas educativos avanzados. Es por esto que España es uno de los países de la OCDE con mayor porcentaje de repetidores al final del periodo obligatorio (más del 30%), muy lejos de las medias de UE (11,2%) o OCDE (12,0%)⁵. Esto además tiene consecuencias negativas sobre la equidad: España es el país de la OCDE donde la relación entre el origen del alumnado y la repetición de curso es más fuerte, a igualdad de competencias. Diversos trabajos que analizan diferencias entre evaluación interna y externa en España y otros países de nuestro entorno muestran, por un lado, la existencia de patrones de evaluación en forma de curva (calificatoria) del alumnado y, por otro, cómo el alumnado de menor renta es sistemáticamente penalizado en la evaluación interna respecto a la externa. Calsamiglia y Loviglio (2018) muestran la existencia de un fuerte efecto curva para el caso de Cataluña. Por otro lado, existe evidencia sólida de que la evaluación interna (en comparación con la externa) perjudica a: (i) los chicos respecto a las chicas en Portugal (Angelo, 2016), Cataluña (Calsamiglia y Loviglio, 2018), Andalucía (Marcenaro y Vignoles, 2014) o Italia (Di Liberto et al, 2016); (ii) al alumnado de nivel socioeconómico más bajo respecto al de nivel más alto en Francia (Cosnefroy et al, 2005) y España (ver el caso de Cataluña (Calsamiglia y Loviglio, 2018) o Andalucía (Marcenaro y Vignoles, 2014).

⁶ Gauna, Dávila, Etxebarria y Sarasua (2013) establecen distintos campos de estudio sobre la selectividad que pueden aportar un corpus de conocimiento científico suficiente: (i) resultados en distintas universidades; (ii) análisis comparado; (iii) coherencia interna de la prueba; (iv) concurrencia y capacidad predictiva; (v) materias o aspectos concretos; (vi) formas de acceso paralelas; (vii) tribunal paralelo.

⁷ Procedimientos de evaluación y uso de descriptores generalmente, y otras cuestiones de índole conceptual, como asumir que las competencias clave, por su relación con las asignaturas, se pueden “calificar” haciendo uso de fórmulas basadas en la calificación de asignaturas.

2.2. La evaluación externa en España

En nuestro país falta tradición de evaluación global del sistema educativo: se ha evaluado **mucho a los alumnos, poco a profesores, centros educativos y programas, y nada a la aplicación de las políticas, leyes y normas por parte de las administraciones educativas**. Salvo excepciones, la evaluación externa de los aprendizajes por parte del alumnado aterriza en España hace aproximadamente 20 o 25 años. Con la llegada del nuevo siglo, se ha puesto el foco en la evaluación externa de los aprendizajes, primero con la participación de España en las pruebas TIMSS o PISA (a partir de 1995 y 2000 respectivamente) y después a través del desarrollo normativo de la LOE, que introduce evaluaciones de diagnóstico a mitad de etapa a nivel nacional (mediante la Evaluación General de Diagnóstico) o autonómico (mediante evaluaciones de diagnóstico autonómicas), en 4º de primaria y 2º de ESO. Por su parte, la LOMCE introduce inicialmente en 2013 nuevas evaluaciones (basadas en los nuevos estándares de competencias) de carácter sumativas y censales de fin de etapa⁸ combinadas con evaluaciones de diagnóstico de mitad de etapa (3º de primaria), y a partir de 2016 las reconfigura hacia un conjunto de evaluaciones de mitad y final de etapa, todas ellas de carácter muestral, formativo y de diagnóstico⁹.

Bajo este contexto, la aparición de evaluaciones externas ha contribuido de forma positiva en varios aspectos del sistema educativo:

- Poco a poco, aunque con problemas en algunos momentos, se ha ido construyendo **una cierta cultura de la evaluación** en los centros y en el sistema educativo, a través de la institucionalización de la evaluación externa.
- Ya sea a través de pruebas diagnóstico o PISA, si están correctamente configuradas y diseñadas, generan un criterio de **igualdad o mérito** en relación a la valoración de los niveles de logro o éxito de los alumnos en cada una de las competencias adquiridas o desarrolladas. Esto representa una fortaleza en especial al contraponerse con las diferentes valoraciones que se producen en la evaluación interna y las diferencias sistemáticas que se producen por tipo de centro o alumnado.
- La evaluación externa ha logrado incorporar modelos fundamentalmente ligados al enfoque competencial, planteando los ítems de evaluación como la **combinación** entre **contenido** (o saberes), **aplicación** y **contexto**. En algunos casos, se ha avanzado en la evaluación más allá de las competencias más habituales (como lectura o matemáticas), evaluando competencias como la **competencia social y ciudadana** o la **competencia de aprender a aprender**.
- El florecimiento de **nuevos debates educativos, paradigmas y formas** de investigar la educación que no hacen si no enriquecer el debate público y el conocimiento del mismo. Un ejemplo de idea que contribuye al debate educativo puede el resultado empírico a nivel internacional de que la calidad y la equidad son dos reversos de la misma moneda en lo que se refiere a un sistema educativo avanzado.
- En cuanto al **Ministerio de Educación (y el INEE)**, se han producido **importantes avances** (la mayor parte duraderos) permitiendo un fuerte desarrollo institucional: desarrollo de evaluaciones diagnóstico de carácter nacional (Evaluación General de Diagnóstico),

⁸ Ley Orgánica 8/2013

⁹ Real Decreto Ley 5/2016.

desarrollo del marco evaluativo de evaluaciones LOMCE, consolidación del Sistema Estatal de Indicadores de la Educación, promoción de análisis e investigaciones relacionadas con las evaluaciones externas (EGD, PISA, TIMSS, TALIS) y coordinación y cooperación, intensa durante algunos periodos, con las Comunidades Autónomas en materia de evaluación externa.

- En algunas **Comunidades Autónomas**, se ha avanzado notablemente en términos de **institucionalización y normalización de la evaluación externa** en la administración pública. Han aparecido agencias o servicios de evaluación educativa en todas las Comunidades Autónomas, las cuáles se han responsabilizado directamente en la implementación de evaluaciones externas, generando conocimiento específico en torno a la administración de pruebas, diseño de ítems, análisis de resultados, implicación de la comunidad investigadora y pilotajes experimentales de nuevos instrumentos de evaluación. En algunos casos, esta institucionalización ha facilitado y promovido un diálogo educativo constructivo entre centros educativos, administración y comunidad educativa (mediante informes de centro o de familia a partir de los resultados de las evaluaciones). Además, se han facilitado las **primeras investigaciones que analizan y describen el aprendizaje del alumnado**, así como en algunos casos las causas de ciertos resultados, y consecuencias de ciertas políticas y programas experimentales. Cabe destacar las contribuciones realizadas por el ISEI-IVEI en Euskadi, el Consell Superior d'Avaluació del Sistema Educatiu en Cataluña, la Agencia Andaluza de Evaluación Educativa (AGAEVE) o el servicio de evaluación educativa de la Comunidad de Madrid (ahora en la Subdirección General de Evaluación y Análisis).

A su vez, la evaluación externa presenta a día de hoy las siguientes debilidades institucionales y pedagógicas:

- A día de hoy, conviven evaluaciones externas de los modelos LOE y LOMCE (con propósitos, diseños y usos diferentes), lo cuál significa que en **2020 el sistema educativo español sigue sin disponer de un modelo de evaluación externa integral y consensuado**. La ausencia de modelo puede apreciarse en lo que se refiere al diseño institucional (el papel que juegan el Ministerio y las Comunidades Autónomas), al enfoque pedagógico-curricular (los marcos evaluativos de las evaluaciones LOE y LOMCE presentan algunas similitudes y diferencias), el tiempo educativo (el momento de la evaluación, donde conviven mitad y final de etapa), el propósito explícito de política educativa (que va desde la transparencia en sí misma – primera etapa LOMCE- hasta la mejora educativa de los centros) ni por supuesto, el tipo de evaluación (diagnóstica, formativa o sumativa).
- El hecho de que no exista un modelo de evaluación externa nacional hace que, **por incomparecencia**, las **evaluaciones internacionales** como **PISA** tengan un **papel excesivamente relevante** en la conversación pública, algo que no necesariamente se produce en otros países de nuestro entorno. Son evaluaciones útiles y reveladoras pero que en todo caso no pueden orientar la mejora educativa en los centros en la medida en la que son muestrales, no han sido construidas con esa finalidad y tampoco disponen de un marco evaluativo alineado con el currículum español.
- La falta de consistencia de un modelo de evaluación consensuado **ha minado la confianza de la comunidad educativa, familias y el público en general** acerca del papel que puede jugar la evaluación en la mejora educativa. Salvo excepciones, no se han llegado a consolidar iniciativas y proyectos útiles para la mejora de los centros educativos a partir de resultados de evaluaciones externas

- Las evaluaciones externas en España presentan limitaciones o riesgos que no deben ser ignoradas. En concreto: (i) el riesgo de evaluar para clasificar y no para mejorar; (ii) se producen en un único momento puntual, dando una foto fija que no siempre coincide con la realidad, y por tanto no valoran de manera continua el aprendizaje; (iii) conforme se acercan al final de cada etapa, plantean riesgos de pervertir el propósito de aprendizaje, donde la prueba puede convertirse en fin en sí mismo y no en medio para la mejora; (iv) la poca innovación de la evaluación externa en cuanto a competencias evaluadas y nuevos formatos de evaluación y; (v) es poco recordado que casi el 50% del alumnado del sistema educativo en España es bilingüe y esto condiciona las comparativas entre comunidades autónomas.
- Muchas pruebas externas de evaluación desarrolladas **carecen todavía del rigor y la fiabilidad necesaria** y exigible en este tipo de procesos. La comparabilidad entre pruebas que se desarrollan cada edición es dudosa y la estandarización de muchos procesos y fases es precaria.
- Es relevante además recordar el **papel en ocasiones negativo que ha jugado la evaluación externa en el debate público** en educación en España. Desde que aparecen los primeros resultados de las evaluaciones externas de PISA o TIMSS, se ha producido la paradoja de un acercamiento de los medios de comunicación al debate educativo (siendo España uno de los países donde más ocurre) acompañado de la proliferación de enfoques simplificadores y resultadistas en el mejor de los casos, y descontextualizados en el peor de ellos.
- Hasta el momento, es llamativa la **ausencia de una cultura de la responsabilidad sobre la evaluación externa** por parte de los centros en un sentido asociado a la mejora (Bolívar, 2016) conjugada con procesos de auto-evaluación o procesos de evaluación interna de centro.
- La evaluación externa **se ha centrado únicamente en resultados**, dejando de lado los procesos de aprendizaje y por tanto la mejora de las competencias profesionales docentes. Lamentablemente, en este sentido, apenas se han realizado evaluaciones externas (basadas un marco estandarizado y estadísticamente fiable) de observaciones de aula mediante instrumentos que permitan comprender en mayor profundidad para guiar la práctica docente. Estas evaluaciones (como CLASS, FFT, MQI o PLATO) tienen un gran potencial en lo que se refiere a mejora docente en el aula, con procesos de acompañamiento y tutorización.
- Las **estructuras administrativas que apoyan la evaluación externa han disfrutado de escasa estabilidad y poco apoyo técnico necesario para su actuación**. Por un lado, para el caso del INEE se han dado demasiado cambios de nombre, de funciones, de composición, de personas y de estructura, lo que ha afectado a su eficacia y a su credibilidad. Por otro, son muy pocas las comunidades autónomas las que tienen un instituto o una institución específica dedicada a este ámbito, lo que conlleva que muchas tareas se contraten externamente y no se cree conocimiento en la comunidad educativa. Muchas de las personas que forman parte de estos organismos están en comisión de servicio, sin estructuras estables y sin procesos de formación claros, algo que ocurre también en el propio INEE.

III. Una visión de un modelo de evaluación de la educación: objeto y objetivos.

Nuestro sistema educativo se encuentra en un momento crucial, lleno de complejidad y con unas grandes necesidades de progresar y mejorar. En este contexto, consideramos que la evaluación educativa debe jugar un papel central de un nuevo modelo que permita alcanzar las capacidades del sistema para el éxito escolar de todos los alumnos.

Nuestra visión de la evaluación¹⁰ es la de una evaluación como sistema de información y responsabilidad educativa efectiva que pueda:

- (i) Contribuir de manera decisiva a lograr **mejoras sustanciales en el aprendizaje** del alumnado hacia una educación inclusiva, equitativa y de calidad, definiendo aprendizajes de manera holística y participativa;
- (ii) **Fortalecer la profesión docente:** a nivel formativo, la evaluación tiene la capacidad de acompañar a los centros educativos en el refuerzo de su capacidad profesional, facilitando la reflexión orientada a la mejora, ya sea individual o colectiva;
- (iii) Ofrecer **información pública de calidad** a la sociedad de los procesos, objetivos, proyectos y resultados de aprendizaje de los centros educativos. La evaluación es por tanto un bien público que opera como pilar fundamental en los procesos de aprendizaje en la escuela y la incorporación de conocimiento científico para las políticas educativas.



Nuestra propuesta se construye sobre seis pilares básicos que permiten definir una visión compartida de la evaluación educativa. Esta propuesta entiende que la evaluación educativa debe construirse sobre un **marco de evaluación consensuado**, estar **basada en conocimiento científico (y generarlo)**, promover una **cultura evaluativa** de todos los agentes, **vincular y conectar sus distintas categorías** (finalidades, focos y agentes de evaluación), fomentar más

¹⁰ Basada en Fullan, Rincon-Gallardo y Hargreaves (2015)

responsabilidad por parte de los agentes y promover un **diálogo abierto y constructivo** sobre la educación que esté basado en la **información como bien público**.

- 1. La evaluación debe construirse sobre un marco de evaluación consensuado.** La historia reciente de nuestro país muestra una concatenación de decisiones solapadas y erráticas en la evaluación (sobre todo externa): en muchos casos, se deben a la falta de consenso básico sobre un marco de evaluación, algo que ya se ha señalado como una de las debilidades fundamentales del sistema. Apostamos por un proceso participativo plural de reflexión y discusión entre teoría y práctica que trabaje en la definición precisa de marcos de evaluación para cada uno de los objetos de evaluación del sistema educativo, consensuados con los grupos de interés. Este proceso debería delimitar la práctica evaluativa en cuanto a sus fines, criterios o estándares, agentes implicados, procedimientos, instrumentos, sistemas de retroalimentación y responsabilidades asumidas.
- 2. La evaluación educativa debe estar basada en conocimiento científico.** Toda medida propuesta acerca de la evaluación interna o externa debería fundamentarse, siempre que sea posible, en la investigación existente sobre el tema en cuestión: informes comparados, artículos científicos sobre aspectos concretos de la evaluación, meta-análisis y obras generales. No se trata de un planteamiento de importar modelos de otros países, de adoptar sin más las conclusiones derivadas de los diferentes estudios, ni por supuesto de sustituir el juicio humano o el marco teórico por un empirismo mal entendido. Más bien se trata de adaptar el conocimiento a nuestro contexto y siempre de manera conjunta y colegiada con los pilares fundamentales del sistema educativo: modelo de currículum, modelo de organización de centros o modelo de profesión y profesionalidad docente. Asimismo, la propia evaluación ha de retroalimentar a todo el sistema educativo, en sus múltiples niveles y objetos, con el fin de progresar hacia la consecución de los objetivos propuestos y generar a su vez nuevo conocimiento científico.
- 3. La evaluación educativa debe promover una mayor y mejor formación y cultura evaluativa.** La formación en todos los niveles del sistema educativo se ha de extender de forma que se maneje un lenguaje común sobre los conceptos básicos de la evaluación de los aprendizajes. Un salto en la capacidad de un sistema de evaluación educativa implica un salto en las capacidades de la comunidad educativa sobre los conceptos elementales de la evaluación, su significado teórico, su sentido global y sus implicaciones para la práctica. Asimismo, la formación puede servir de catalizador para un cambio en la cultura evaluativa dentro de procesos de reflexión participativos. Esto puede ocurrir mediante el cuestionamiento de prácticas tradicionales, el abordaje de resistencias al cambio, el fomento de la participación y la discusión sobre el papel que diferentes agentes pueden jugar.
- 4. La evaluación debe vincular y conectar finalidades, focos y agentes de evaluación.** Las finalidades de todos los objetos de la evaluación han de estar alineadas en sentido ascendente y descendente: los objetivos de calidad, inclusión y equidad del sistema educativo han de depender de la mejora y eficacia de los programas llevados a cabo en los centros educativos. A su vez, los centros educativos quedan condicionados por el fortalecimiento de los aprendizajes y por el desarrollo de la competencia profesional de los docentes. En sentido inverso, los conceptos de inclusión, equidad y calidad han de guiar la definición del resto de objetivos de los otros niveles de evaluación. Todo ello requiere de un diseño cuidadoso, tecnicado y operativo con la coordinación de distintos organismos (Ministerio y CC.AA) a través de sus agencias e institutos de evaluación, colectivos profesionales, asociaciones de madres y padres, y alumnado.

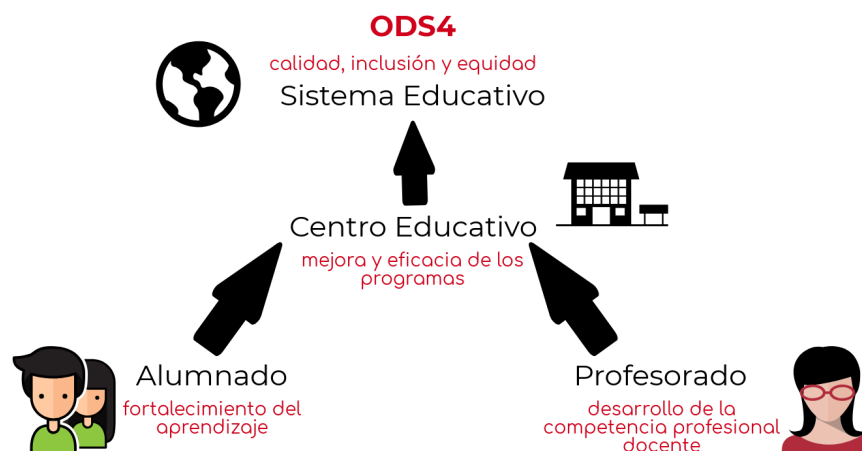
5. **La evaluación debe estar acompañada de responsabilidad por parte de todos los agentes del sistema.** No hay mejora del aprendizaje sin proceso de reflexión, ni proceso de reflexión sin marco de responsabilidad en los centros educativos y como parte de la competencia profesional docente. Por ello, es necesario profundizar en el desarrollo de marcos de responsabilidades, para lograr que incorporen procesos de evaluación (interna y externa) ligados al desarrollo profesional y los planes de mejora de los centros educativo. La evaluación de los centros educativos es además el primer paso hacia la evaluación de políticas del sistema educativo, de las cuales son también responsables la administración y los servicios de apoyo: estas deben responder y responsabilizarse ante las necesidades educativas y socioeconómicas de cada centro, ante la efectividad de los programas e inversiones que se impulsan y ante la consecución de grandes objetivos del sistema educativo.

6. **La evaluación debe promover un diálogo abierto, profundo y constructivo basado en la información como bien público.** En la era de la información, la transparencia y los datos abiertos, la escuela financiada con fondos públicos no puede quedarse de espaldas. Por muchos riesgos que la transparencia conlleve, la opacidad de un sistema educativo supone asumir que la mejora es un proceso interno, cerrado, sin certezas y en algunos casos, regresivo. Por ello, se debe regular y garantizar un sistema de información para los centros, alumnos y familias desde una perspectiva de educación inclusiva y equitativa y de calidad, que marque y guíe la conversación pública en torno al sistema educativo. Se trataría de un sistema de indicadores que refleje la realidad objetiva del proyecto y el desarrollo del centro educativo, descartando cualquier posibilidad de ranking o datos de composición del centro que pueda generar estigmatización o simplificación del proceso educativo. Resulta clave la cooperación entre servicios y/o agencias de evaluación e inspección educativa en el diseño de dicho sistema y recogida de información.

IV. Hacia un modelo de evaluación del sistema educativo

En la siguiente gráfica se presentan a la vez los objetos y objetivos del modelo de evaluación planteado. La evaluación educativa, tanto externa como interna, se fija en cuatro ámbitos con el objetivo fundamental de aumentar la calidad, la inclusión y la equidad: **alumnado, profesorado, centros educativos y sistema educativo.**

Evaluación Educativa objeto y objetivos



Para el alumnado, la evaluación aspira a convertirse en un instrumento de aprendizaje a lo largo de la vida y su fortalecimiento. Esto requiere de la adopción de un enfoque en el sistema educativo que esté alineado con la evaluación y validación de las competencias clave. Esto es, en un marco de resultados de aprendizaje, combinando la evaluación de estos aprendizajes desde el centro educativo con evaluaciones de diagnóstico, formativas y sumativas.

Para el profesorado, supone convertirse necesariamente en un instrumento para el desarrollo de la competencia profesional docente, entendida ésta como combinación de competencias individuales asociadas a la práctica y desempeño con competencias profesionales a nivel centro (Moya, 2018). La competencia profesional docente permitiría al docente, dentro de la visión planteada, disponer de una herramienta que le permita su progreso profesional.

Para el centro educativo, se ha de enmarcar en los procesos de mejora de los programas educativos que se llevan a cabo en el centro. La evaluación responde como consecuencia de la consecución de determinados factores de eficacia escolar, donde la participación y el papel de la comunidad escolar son elementos clave, junto con el papel del liderazgo pedagógico, el clima escolar, o el currículum.

A nivel sistema, las dimensiones de la evaluación anteriores permiten avanzar y en la evaluación del sistema educativo en su conjunto, cuya responsabilidad habría de enmarcarse en la mejora de la inclusión, equidad y calidad del sistema.

4.1. Claves para un modelo de evaluación interna

Cualquier proceso de evaluación ha de responder a una serie de demandas que se plantean aquí a modo de claves o estrategias. Por un lado, estrategias generales, y por otro estrategias para cada uno de los cuatro objetos: alumnado, profesorado, centro educativo y sistema educativo. Estas estrategias nacen a partir del diagnóstico previamente realizado de la situación de la evaluación interna en España y la visión planteada para un modelo de evaluación educativa.

i) Estrategias generales para la evaluación interna

E.1. Basar la evaluación interna en el conocimiento científico generado, fortaleciendo los puentes entre la política educativa, la investigación en evaluación y la práctica evaluativa¹¹.

E.2. Diseñar los procesos de evaluación para aumentar su validez, definiendo sus fines, los elementos a evaluar, los procedimientos e instrumentos, los recursos económicos y humanos, y los mecanismos de difusión y asunción de responsabilidades. Construir dichos procesos sobre un marco operativo dotado de participación abierta a todos los grupos de interés del sistema educativo (gobiernos, profesionales de la educación e instituciones afines y familias). Esta estrategia debería facilitar una evaluación planificada para los fines para los cuáles ha sido creada con las acciones a realizar, los procedimientos e instrumentos y el feedback que proporciona a los interesados.

E.3. Formar a los distintos participantes en los procedimientos e instrumentos de evaluación, así como fomentar una cultura de la evaluación orientada a la responsabilidad, transparencia y el bien público.

ii) Estrategias para la evaluación de los fines, organización y funcionamiento de los centros educativos

Le evaluación ha de convertirse en una herramienta de ayuda para comprender la realidad del centro escolar, mediante un proceso sistemático de recogida de información relevante, fiable, contrastada y válida que facilite la toma de decisiones en la perspectiva de una mejor adaptación a las necesidades del alumnado y a las demandas de la comunidad educativa. Para ello se plantean las siguientes estrategias:

E.4. Basar la evaluación en los pilares que sustentan y dan sentido a la autonomía del centro educativo, concretados y con alcances temporales diferenciados: proyecto educativo (largo plazo), proyecto de dirección y planificación estratégica (medio plazo) y PGA y programas concretos (corto plazo).

E.5. Ofrecer un marco de evaluación de centros con cabida para la vinculación de indicadores a los objetivos del sistema educativo y otros objetivos que el centro educativo defina en el ejercicio de su autonomía. Para ello es necesario consensuar las áreas de evaluación de un centro, entre las cuáles también están los programas específicos, proporcionar instrumentos y procedimientos sencillos y eficaces¹².

E.6. Establecer los tiempos mediante una planificación cuidadosa y dotar de recursos necesarios para la realización efectiva y realista de la evaluación del centro y sus programas contando con la participación de los diferentes órganos colegiados de la comunidad educativa y distribuyendo responsabilidades en el proceso, donde ha de destacarse la figura del equipo directivo. Se requiere

iii) Estrategias para la evaluación del desarrollo profesional docente

¹¹ Por ejemplo: a) Los estudios centrados en el movimiento de escuelas eficaces definen los factores que favorecen el éxito en la escuela (clima escolar y aula, dirección escolar, desarrollo profesional docente, currículo de calidad...); (b) Investigaciones como la de Hattie, permiten establecer, los factores de la mejora del aprendizaje y rendimiento del alumnado.

¹² En lo referido a las áreas, se apuntan entre las más relevantes: a) organización y funcionamiento escolar, b) currículo y proceso de enseñanza-aprendizaje, c) clima y convivencia escolar, d) procesos de formación e innovación educativa y e) relaciones con el entorno y las familias.

La evaluación del docente debería seguir los principios de la evaluación formativa. Esto es, ha de permitir fundamentalmente el desarrollo profesional. Para ello se debería:

E.7. Definir operativamente la evaluación del docente, con la participación de los profesionales, como un proceso de auto-evaluación, evaluación por pares o evaluaciones con instrumentos de observación de aula que generen compromisos y participación orientados a la mejora capacidad profesional docente. Un ejemplo claro se puede encontrar en el modelo de Capacidad Profesional Docente del Proyecto Atlántida. En concreto se propone al menos la evaluación de las dimensiones y estrategias de la docencia: didáctica, expertise...

E.8. Vincular la evaluación al desarrollo de la carrera docente y a las mejoras de la capacidad profesional docente en cuanto a su práctica individual y en cuanto a su competencia social y decisoria en el centro educativo, evitando la vinculación a los resultados del alumnado

iv) Estrategias para la evaluación del aprendizaje

E.9. Acompañar los cambios curriculares y metodológicos con un enfoque de evaluación para el aprendizaje (evaluación formativa): centrado en la potencialidad del feedback, la transparencia y participación del alumnado en el propio proceso de evaluación, el uso de resultados de aprendizaje en coherencia con esos referentes curriculares y metodologías y centrada en la evaluación de desempeños.

E.10. Proporcionar una alternativa, o mínimo complementaria, al uso de las calificaciones y la idea de pseudo-medida del aprendizaje que suele plasmarse en las notas numéricas, promoviendo el uso de niveles de desarrollo competencial descriptivos en rúbricas y otros instrumentos. Proporcionando especialmente al alumnado información personalizada y de proceso.

E.11. Promover el uso de tecnologías digitales para captar los avances en múltiples dimensiones de las competencias (saber, saber cómo y saber ser y estar) y en distintos contextos de aprendizaje (también los no formales o informales) disponibles para todas las personas.

E.12. Reducir de forma masiva el uso de la repetición de curso, principalmente en ESO, para mejorar las oportunidades del alumnado. Se trataría de combinar distintas iniciativas: (i) promover estrategias y medidas para visibilizar y difundir el peso de la de la dimensión cultural de la evaluación interna y su impacto en la repetición de curso; (ii) limitar a una sola vez la repetición de curso por nivel educativo (primaria y ESO) y revisar el sistema de promoción y titulación; (iii) facilitar a los centros educativos y el profesorado de estrategias, recursos y alternativas, a poder ser, en edades tempranas y enfatizando especialmente la figura del tutor/a; (iv) adoptar un currículo flexible en cuanto a su duración y contenidos, fomentando experiencias personalizadas de aprendizaje.

E.13. Profundizar en una evaluación facilitadora del acceso a distintas vías formativas en función de las competencias adquiridas entre etapas, especialmente entre la educación obligatoria y post-obligatoria. Esto pasa por:

- i. Desarrollar un sistema de certificación de los aprendizajes y las competencias al final del periodo obligatorio que sustituya al título de ESO, que combine evaluación interna y externa. Esta estrategia debe permitir orientar el

seguimiento de itinerarios formativos académicos y profesionales entre los 16 y los 18 años. El objetivo último es que todos los alumnos tengan oportunidades formativas.

- ii. Modificar la Evaluación del Bachillerato para el acceso a la Universidad (EBAU), primero desarrollando una reflexión acerca de sus consecuencias pedagógicas en etapas anteriores y posteriores, y después planteando un modelo alternativo basado en el marco de competencias.

4.2. Claves para un modelo flexible y compartido de evaluación externa

Se plantea un modelo articulado, coherente y consensuado de evaluación externa que por primera vez permita visualizar un horizonte de trabajo de medio plazo por parte de todas las administraciones. La falta de consenso y diálogo educativo entre los distintos modelos de evaluación externa ha minado la confianza de los actores del sistema, y es preciso reestablecer esta confianza planteando unas claves o estrategias ambiciosas y a la vez realizables en términos de consenso. Estas estrategias nacen a partir del diagnóstico previamente realizado de la situación de la evaluación externa en España y la visión planteada para un modelo de evaluación educativa.

i) **Claves para el diseño las pruebas de evaluación externa**

Se plantea y persigue identificar un marco de evaluación externa con las siguientes características:

- *Los fines y objetivos* de la evaluación: es necesario clarificar los objetivos de las evaluaciones externas (diagnóstico, formativa, sumativa), priorizando el diagnóstico y la evaluación formativa sobre la sumativa durante las etapas obligatorias y la sumativa al final del periodo obligatorio.
- *El tipo de prueba*, pudiendo ser estandarizada y cuantitativa o cualitativa, en especial si se considera la evaluación del desarrollo en la etapa infantil.
- *El objeto evaluado* debe definirse de manera amplia y acorde al currículum. Esto implica decidir qué competencias se evalúan, en qué áreas o ámbitos del conocimiento. Se plantea un modelo que trate de ampliar el foco más allá de las competencias habitualmente evaluadas como la lectura, matemáticas o ciencias.
- *La población evaluada* es clave para profundizar en los fines y objetivos de la evaluación (esto es, toda la población de alumnos, o solo una muestra). Se plantea un modelo combinado de evaluaciones muestrales (orientadas a la la finalidad formativa de todo el sistema, tanto autonómico como nacional) y evaluaciones censales (orientadas a la finalidad de diagnóstico para todo el alumnado).
- *Las responsabilidades de las instituciones* en el diseño de las pruebas (central o autonómica). Se plantea un modelo que comparta un marco común entre MECD y CCAA, sencillo y que dote de flexibilidad, autonomía y adaptabilidad a las necesidades de cada CCAA: un modelo que disponga de ítems comunes e ítems específicos que faciliten la operatividad de pruebas nacionales y autonómicas. Por otro lado, es relevante determinar los responsables de la gestión de las pruebas, que en el caso muestral

pueden ser las CCAA y en el caso censal puede descentralizarse incluso hasta el centro educativo.

- *Los tiempos de la evaluación* requieren de una definición clara de la etapa evaluada (primaria o ESO, e incluso infantil), momento de la etapa (mitad o final de etapa) y la periodicidad de la evaluación (anual, bianual o trianual). Se plantea un modelo que priorice la mitad de etapa, al menos en el caso censal, ya que es ahí donde los planes de mejora pueden incidir a tiempo sobre el alumnado; además, se considera que la evaluación del sistema no tiene por qué tener una periodicidad anual y que la evaluación de diagnóstico para los centros puede organizarse de manera flexible en los centros educativos, dada la autonomía que otorgan los recursos tecnológicos.

ii) *Claves para el uso de la evaluación externa*

El punto anterior presenta un marco que facilitaría un proceso de construcción de modelo de evaluación externa a través de la definición de fines y objetivos, tipo de prueba, objeto, población, responsabilidad en la evaluación y tiempos. Estas ideas generales deberían constituir un modelo consistente, alineado con el currículum, que dote de suficiente información a sociedad (formativa), centros educativos (diagnóstico) y que faciliten las oportunidades para todos los alumnos al final del periodo obligatorio (sumativo).

Sin embargo, el uso de la información que la evaluación externa puede seguir distintos planteamientos, finalidades y estrategias. Sin ánimo de pretender ser exhaustiva, se plantea una lista de objetivos y principios basados en experiencia de las últimas dos décadas que podría contribuir a guiar dicho debate. El uso de las pruebas de evaluación externa debe estar:

- **Orientada a la mejora del alumno, docente, centro educativo, profesorado y el sistema:** la falta de propósito (o la existencia de un propósito negativo y sancionador) de la evaluación externa debilita su propia existencia y compromiso por parte de la comunidad educativa. Por ello, la evaluación externa debe facilitar información de calidad a la administración (en torno a la evaluación de políticas mediante cooperación con la comunidad investigadora), inspección educativa (en torno a su capacidad orientadora de los centros educativos), centro educativo (en torno a la consecución de objetivos y planes), profesorado (en torno a su estrategias de evaluación interna y su relación con la evaluación externa), y alumnado (en torno a su progreso educativo y sus logros).
- **Basada en el conocimiento científico:** la literatura al respecto muestra algunos resultados importantes que deben ser tenidos en cuentas en este proceso. Por poner dos ejemplos: (i) la información devuelta a los centros educativos, profesorado y familias a partir de las evaluaciones diagnóstico y en conjunción con la evaluación interna puede generar mejoras y como un proceso de empoderamiento de centros; (ii) la publicación de rankings de resultados puros, sin contexto ni propósito, puede tener efectos nocivos sobre la igualdad de oportunidades y la segregación escolar; (iii) el tipo de respuesta, compromiso e interés del alumnado varía enormemente cuando se trata de pruebas low-stakes (diagnóstico, pruebas internacionales) respecto a las pruebas high-stakes (sumativas, con consecuencias para el alumnado), resultando por tanto diferentes las decisiones que deben tomarse a partir de las mismas.
- **Construida en torno a un marco evaluativo alineado con los objetivos del currículum.** En el presente documento se pone un esfuerzo en concretar un modelo de evaluación y se incide enormemente vincular la evaluación con el currículum. En este sentido, el uso de la evaluación externa debe ser el de contribuir a la profundización del modelo

curricular y el enfoque por competencias, tanto para el profesorado como para las familias y la sociedad en su conjunto.

- Que facilite la **construcción consensuada de un sistema público transparente de indicadores** orientado a la mejora, adaptado a la realidad de cada CCAA, que refleje, entre otros, el proyecto educativo y sus indicadores de logro mediante la evaluación interna e indicadores de valor añadido (a partir de la evaluación externa), incluyendo tendencias y evolución, descartando cualquier posibilidad de rankings o datos de composición del centro que pueda generar estigmatización o simplificación del proceso educativo. Que no se realicen rankings educativos no significa que no se realicen comparativas entre centros que compartan las mismas condiciones sociales y económicas. La comparación puede ser una ayuda para interpretar los datos e identificar qué se está haciendo bien en cada centro.

iii) Claves para un nuevo sistema nacional de indicadores orientado al ODS4

El actual Sistema Estatal de Indicadores de la Educación se ha consolidado como herramienta de referencia de información del sistema educativo español. Desde su primera edición en 2000, el sistema ha evolucionado y ha ido incorporando indicadores y métricas acorde con la realidad educativa, el contexto y los objetivos que ha ido incorporando el sistema. En este sentido, se plantea un nuevo salto que integre como clave fundamental el objetivo ODS4 como eje vertebrador del Sistema Estatal de Indicadores, conjugándose con el modelo actual y con nuevos indicadores que reflejen los nuevos retos del sistema educativo español en el momento actual.

El ODS4 está formulado de la siguiente manera “Garantizar una educación inclusiva, equitativa y de calidad y promover oportunidades de aprendizaje durante toda la vida para todos”¹³. Por otro lado, durante los últimos años ha existido un amplio debate sobre la equidad y la inclusión del sistema educativo español, entendida esta como igualdad de oportunidades del alumnado según su origen socioeconómico. La falta de consenso ha sido como “Puzzle de la equidad educativa”¹⁴. Es por tanto relevante incorporar una dimensión de indicadores de equidad, definida como igualdad de oportunidades según el origen del alumno, en los resultados de evaluación externa (pruebas externas nacionales e internacionales) y evaluación interna (repetición de curso y escolarización post-obligatoria), o acceso a centros con entornos socioeconómicos diversos u homogéneos (inclusión o segregación social en los centros educativos). Por su parte, sería igualmente relevante considerar indicadores de la educación especial y su relación con la equidad y la inclusión.

V. La transición a un nuevo modelo de evaluación

La transición hacia una “nueva evaluación educativa” no es un proceso lineal, inmediato o sencillo. Al contrario, en el modelo arriba descrito, se superponen decisiones de tipo administrativo y legal o iniciativas de construcción de nuevas instituciones y capacidades, implementación de nuevos marcos en los centros educativos y en la competencia profesional docente. Por su naturaleza y su historia, este modelo conlleva una transformación cultural en torno a la idea de aprendizaje y la idea de evaluación auténtica (un tipo de evaluación por competencias) en contraposición con la calificación. Se trata por tanto de un horizonte que va

¹³ El ODS4 está regido por los siguientes objetivos: (i) graduación de primaria y secundaria; (ii) servicios de atención en primera infancia y educación prescolar; (iii) equidad en el acceso a la formación técnica y superior; (iv) mejorar los resultados en las competencias necesarias, especialmente técnicas y profesionales; (v) eliminar disparidades de género; (vi) mejorar los resultados en competencias en lectura, escritura y aritmética; (vii) mejorar las competencias para promover el desarrollo sostenible. Además, se consideran indicadores los siguientes recursos: (a) instalaciones escolares de calidad; (b) mejora sustancial del número de becas; (c) mejorar la calificación y el número de maestros.

¹⁴ Martínez García (2017).

desde medidas fácilmente alcanzables, consensos que deben ser (re)construidos, pasando por formación a la comunidad educativa en la cuanto al marco de evaluación planteado. En todo, caso este proceso debe lograr:

- La capacitación integral del sistema educativo en torno a estrategias para la mejora de la evaluación interna, ya sea a nivel centro educativo, profesorado o familias hacia una apuesta decidida por la evaluación criterial y auténtica. Esta capacitación requerirá de recursos y nuevos marcos de responsabilidades. Un programa de cooperación territorial podría ejercer el liderazgo en dicha capacitación.
- En torno a esta capacitación, el establecimiento de objetivos concretos en torno a la repetición de curso.
- La introducción de un marco de evaluación interna de centro, que guíe la construcción de proyectos de centro.
- La creación de un sistema de indicadores de centros educativos como bien público orientado a la mejora.
- El fortalecimiento de la institucionalidad de la evaluación externa, tanto a nivel nacional (INEE) como a nivel autonómico (inspección, orientación y agencias autonómicas de evaluación). Además de en la aplicación de pruebas, diseño de ítems, este fortalecimiento debería incorporar la responsabilidad de generar conocimiento científico que provea al sistema de información sobre las políticas educativas.

Propuestas a corto plazo (2020)	Propuestas a medio plazo (2023)	Propuestas a largo plazo (2026)
<ul style="list-style-type: none"> • Plan de formación y actualización (mediante PCT y planes autonómicos) de las capacidad y estrategias de evaluación interna de los centros educativos y el profesorado. 	<ul style="list-style-type: none"> • Implementación de las primeras pruebas de evaluación externa a través del nuevo marco (CCAA). 	<ul style="list-style-type: none"> • Creación de un sistema de indicadores de centros, en base a los marcos de evaluación interna y externa (CCAA, con apoyo del MEFP)
<ul style="list-style-type: none"> • Construcción del marco de evaluación externa que guíe el sistema educativo en el periodo 2020-2030 (MEFP y CCAA) 	<ul style="list-style-type: none"> • Reforma de la Prueba de acceso a la Universidad (EBAU), planteando un modelo alternativo basado en el marco de competencias (MEFP y CCAA). 	<ul style="list-style-type: none"> • Reducción de la tasa nacional de repetición de curso anual en ESO del 10% al 5% (y del 30% acumulado al 15%).
<ul style="list-style-type: none"> • Actualización del sistema estatal de indicadores de la educación con el marco ODS4 e indicadores sobre equidad (MEFP). 	<ul style="list-style-type: none"> • Construcción de marcos de evaluación interna de centro nacional y autonómicos (MEFP y CCAA). 	<ul style="list-style-type: none"> • Universalizar la evaluación por competencias o auténtica en todo el periodo obligatorio.
<ul style="list-style-type: none"> • Construcción del marco de evaluación del docente y sus características (MEFP y CCAA). 	<ul style="list-style-type: none"> • Desarrollar programas piloto de evaluación externa del profesorado mediante instrumentos de observación de aula (mediante PCTs) 	<ul style="list-style-type: none"> • Creación de un sistema de certificación de los aprendizajes y las competencias al final del periodo obligatorio que sustituya al título de ESO (MEFP y CCAA).

VI. Glosario de evaluación de los aprendizajes

Calificación

Es el proceso que usa números para significar el logro de un estudiante. (Nota tr.: ‘marking’ en nuestro contexto se puede asimilar a *corregir, poner una nota a un examen o prueba*). También es el proceso de realización de un juicio global sobre la calidad del trabajo (‘grading’), a veces mediante la suma de calificaciones numéricas. (Isaacs, Zara y Herbert, 2013).

Proceso de asignación sistemática de valores (generalmente numéricos, pero que también incluyen letras y comentarios verbales) a los resultados de las pruebas, cuestionarios, etc. (ERIC Thesaurus: *Scoring*)

Proceso de calificación del desempeño de un individuo o grupo, logro, o de forma menos frecuente, el comportamiento, utilizando escalas de valores específicamente establecidas (ERIC Thesaurus: *Grading*)

Criterios de evaluación

Cuestiones objetivas, especificaciones, requerimientos o estándares que actúan como referencia para valorar o juzgar individuos, organizaciones o cosas. (ERIC Thesaurus)

Estándares académicos

Reglas, principios o criterios establecidos por una institución educativa mediante los que se determinan niveles o grados de logro de los estudiantes, generalmente en términos de la medición de la adecuación, aceptabilidad, cantidad, calidad o valor. (ERIC Thesaurus)

Evaluación

Es el proceso de identificación, recogida y análisis de información relevante –que podrá ser cuantitativa o cualitativa-, de manera sistemática, rigurosa, planificada, dirigida, objetiva, creíble, fiable y válida, para emitir juicios de valor basados en criterios y referencias preestablecidos para determinar el valor y el mérito del objeto educativo en cuestión a fin de tomar decisiones que ayuden a optimizar el mencionado objeto. (Lukas y Santiago, 2009, pp. 91-92).

Evaluación alternativa

Estimación del conocimiento o las habilidades por otros medios distintos a los empleados tradicionalmente, pruebas objetivas, especialmente las pruebas de elección múltiple estandarizadas. (ERIC Thesaurus)

Evaluación auténtica

Se trata de un tipo de evaluación por competencias. Evaluación basada en la presentación de problemas o cuestiones importantes en las que el alumnado debe utilizar el conocimiento para mostrar desempeños de forma efectiva y creativa. Las tareas (auténticas) pueden ser también réplicas o análogas al tipo de problemas a los que se enfrentan los ciudadanos o profesionales en un campo determinado. (Wiggins, 1993).

Evaluación basada en el desempeño

Prueba sobre la capacidad de aplicar conocimiento, destrezas y comprensión, habitualmente en un contexto real. (Isaacs, Zara y Herbert, 2013)

Evaluación con referencia al criterio

Evaluación donde se especifican los criterios para el desempeño de un estudiante en diferentes niveles. Los estudiantes que demuestran competencia en cada uno de los criterios logran el éxito en el dominio evaluado, independientemente del desempeño de otros evaluados de la misma forma. (Isaacs, Zara y Herbert, 2013)

Evaluación continua

Estrategia de aula implementada por profesorado para determinar el conocimiento, comprensión y destrezas logradas por los estudiantes. Los profesionales administran evaluaciones de diversos modos a lo largo del tiempo para permitirse observar múltiples tareas y recoger evidencias sobre qué es lo que los estudiantes saben, comprenden y pueden hacer. [...] (Isaacs, Zara y Herbert, 2013)

Evaluación del aprendizaje (sumativa)

Conocida como evaluación sumativa, proceso formal que puede incluir juicio del profesorado y aplicación de pruebas en el que el sentido sumativo es una medida de los resultados al final de una unidad, programa, curso, cualificación o experiencia educativa. (Isaacs, Zara y Herbert, 2013)

Evaluación diagnóstica

Identifica las características, fortalezas, debilidades, conocimiento, destrezas, habilidades y cualquier dificultad o problema en el aprendizaje en un estudiante nuevo con el objetivo de proporcionar consejo y guías iniciales tanto para el estudiante como para el profesor sobre un curso académico. (Isaacs, Zara y Herbert, 2013)

Evaluación para el aprendizaje (formativa)

Proceso de búsqueda e interpretación de evidencias para uso de los estudiantes y profesorado con la finalidad de decidir si los éstos aprenden, dónde necesitan llegar y cuál es la mejor forma de hacerlo. (Assessment Reform Group). (Isaacs, Zara y Herbert, 2013)

Exámenes

Un examen o prueba es un intento de medir el conocimiento, comprensión o habilidad del estudiante en una materia o dominio en un tiempo limitado. (Isaacs, Zara y Herbert, 2013)

Feedback

Proceso a través del cual los estudiantes conocen qué logro están alcanzando y qué necesitan para mejorar su desempeño. [...] (Isaacs, Zara y Herbert, 2013)

Logro

La aptitud es la capacidad y potencial de una persona para desarrollar habilidades o adquirir conocimiento. La evaluación de la aptitud se refiere a medir las destrezas cognitivas, capacidad y velocidad de aprendizaje de una persona. El logro se refiere al cumplimiento – qué es lo que una persona sabe y es capaz de hacer. No hay una línea que divida la aptitud de las pruebas de logro y las evaluaciones. (Isaacs, Zara y Herbert, 2013).

Nivel de desempeño en relación con una medida estándar de desempeño, o de éxito en lograr un fin deseado. (ERIC Thesaurus: *Achievement*).

Medida

Medir algo es cerciorarse del grado o cantidad de una cosa mediante su comparación con una unidad fija o con un objeto de tamaño conocido. La medida en evaluación se define frecuentemente como el acto de asignar letras, números o símbolos de acuerdo a un conjunto de reglas o estándares conocidos. (Isaacs, Zara y Herbert, 2013).

El proceso de obtención de una descripción numérica del grado en el que personas, organizaciones o cosas poseen unas características específicas. (ERIC Thesaurus: *Measurement*).

Objetivos (educativos)

Propósitos o fines hacia los que se dirige un esfuerzo, relacionados con los resultados o la organización del proceso educativo – se incluyen los establecidos por autoridades educativas en todos los niveles. (ERIC Thesaurus)

Resultados de aprendizaje

Declaraciones que describen aprendizajes significativos y esenciales que los estudiantes han alcanzado y pueden demostrar al finalizar un curso o programa. (ERIC Thesaurus)

VII. Bibliografía

Angelo. Gender gaps in different grading systems: Evidence from Portugal. FEUNL Working Papers, 2016.

Bolivar, A. B. (2016). ¿ Cómo puede la evaluación institucional contribuir para mejorar la escuela?. Estudos em Avaliação Educacional, 27(65), 284-313.

Chatterji, M. (2003). Designing and Using Tools for Educational Assessment. Pearson Education.

Calsamiglia, C., & Loviglio, A. (2018). Grading on a curve: When having good peers is not good. Revise and Resubmit at Economics of Education Review.

Comisión Europea/EACEA/Eurydice (2015). La garantía de la calidad en la educación: Políticas y enfoques para la evaluación de los centros educativos en Europa. Informe de Eurydice. Luxemburgo: Oficina de Publicaciones de la Unión Europea.

Cosnefroy, O. y Rocher, T. (2005). Le redoublement au cours de la scolarité obligatoire: nouvelles analyses, mêmes constants. Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, France.

Di Liberto, A., & Casula, L. (2016). Teacher Assessments Versus Standardized Tests: Is Acting 'Girly' an Advantage?.

Eurydice (2004). Evaluación de los centros de enseñanza obligatoria en Europa. Madrid: Subdirección General de Información y Publicaciones.

ERIC Thesaurus. <https://eric.ed.gov/>

Fullan, M., Rincón-Gallardo, S., & Hargreaves, A. (2015). Professional capital as accountability. *education policy analysis archives*, 23, 15.

Gauna Gorostiza, J. G. R., Balsera, P. D., Murgiondo, J. E., & Fernández, J. S. (2013). Pruebas de selectividad en Matemáticas en la UPV-EHU. Resultados y opiniones de los profesores. *Revista de educación*, (362), 217-246.

Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

Isaacs, T., Zara C. Y Herbert, G. (2013) *Key Concepts in Educational Assessment*. London: SAGE.

K. Santiago, J.F. Lukas (2013). *Evaluación educativa*. Madrid: Alianza Editorial.

Marcenaro-Gutierrez, O., & Vignoles, A. (2015). A comparison of teacher and test-based assessment for Spanish primary and secondary students. *Educational Research*, 57(1), 1-21.

Martinez Garcia: 2017. *EL PUZZLE DE LA EQUIDAD EDUCATIVA*. Agenda Pública

Moya Otero, J. M. (2018). Capacidad profesional docente: una nueva visión de la Escuela. *Cuadernos de pedagogía*, (487), 82-87.

Principado de Asturias (2007). *Guía de autoevaluación y mejora de los centros*.

Junta de Andalucía (2015). *Guía orientativa del proceso de autoevaluación y mejora de los centros educativos*.

OCDE. (2009) Santiago, P. y Benavides, S. (2009). *Teacher Evaluation: a conceptual framework and examples of country practices*. OECD.

Wiggins G. P. (1993). *Assessing student performance*. San Francisco: Jossey-Bass Publishers.)