

LA EVALUACIÓN EXTERNA DE LOS APRENDIZAJES COMO GARANTÍA DE CALIDAD EDUCATIVA

/

EXTERNAL EVALUATION OF LEARNING AS A GUARANTEE OF QUALITY EDUCATION

José M^a Arribas Estebaranz

Universidad de Valladolid.

Facultad de Educación (Campus de Segovia).

Departamento de Pedagogía. ¹

Chema2@pdg.uva.es

Resumen

La evaluación externa del rendimiento académico de los estudiantes y/o de los centros es un asunto ciertamente polémico -en estos momentos de plena actualidad- que suscita una fuerte oposición en muchos sectores sociales y educativos.

Exponemos de la manera más aséptica posible y con toda honradez científica los argumentos y recelos esgrimidos habitualmente, muchos de los cuales, aluden más a los efectos perversos -muchas veces ciertos- derivados de una mala praxis que a su misma existencia, cuya bondad o maldad, como en tantos otros órdenes de la vida, dependerá, en gran medida, de su puesta en práctica: por qué y para qué se evalúa, cómo se evalúa, qué utilización se hace de los resultados, qué consecuencias se derivan de ellos...

Casi todos los países y todos los informes al uso contemplan, en mayor o menor medida, la evaluación externa –complementaria siempre a la evaluación realizada

¹ Campus Universitario de Segovia "María Zambrano" Plaza Alto de los Leones, 1. 40005. Segovia

por los profesores- como garantía imprescindible de calidad educativa en sus distintos niveles de concreción curricular.

Palabras Clave: Evaluación; calidad educativa; control del rendimiento; eficacia del centro escolar; evaluación externa

Abstract

The external evaluation of the academic performance of students and / or schools is an admittedly controversial right now-now-fully arouses strong opposition in many social and educational sectors.

We present as aseptically as possible and with all scientific honesty arguments usually put forward and misgivings, many of which allude more perverse-effects often certain derivatives of malpractice that its very existence, the goodness or badness as in so many other spheres of life, depend to a large extent on its implementation: Why and what is assessed, how it is evaluated, what use is made of the results, what consequences follow from them ...

Almost all countries and all reports use contemplated, in varying degrees, the always-complementary external evaluation by the teachers-as an essential guarantee of educational quality at different levels of detail curriculum evaluation.

Key Words: Evaluation; quality of education; achievement control; school effectiveness; External evaluation

Introducción

La evaluación externa del rendimiento académico de los estudiantes o de los centros es un asunto ciertamente polémico; cualquier intento de instaurar este tipo de evaluación sobre cualquier etapa, programa o proceso educativo suscita, inmediatamente, una fuerte oposición por parte de determinadas personas o grupos sociales, generalmente los afectadas, de uno u otro modo, por dicha evaluación, alegando, fundamentalmente, objeciones de carácter ideológico, de orden práctico, a veces intereses espurios: ocultar la realidad o la incompetencia; o a la posibilidad, en ocasiones cierta, de una utilización perversa e interesada de los resultados.

Frecuentemente se asimila el concepto de evaluación objetiva, externa y pública con la concepción más peyorativa de los conceptos de clasificación, jerarquización, eficacia, efectividad, competencia, competitividad, excelencia, liberalismo... Se contraponen, incomprensiblemente, como si fueran incompatibles entres sí, calificándolas de *buena y mala*, respectivamente, la evaluación continua, formativa... y la evaluación final, sumativa...

Muchas de estas objeciones, aparentemente irrefutables, aluden más a los efectos perversos -muchas veces ciertos- derivados de una mala praxis que al principio teórico. Creemos que es ahí donde radica el principal problema de la mala imagen de la evaluación certificadora y externa, cuya bondad o maldad, como en tantos otros órdenes de la vida, dependerá, en gran medida, de su puesta en práctica: por qué y para qué se evalúa, cómo se evalúa, qué utilización se hace de los resultados, qué consecuencias se derivan de ellos...

La dificultad en la constatación de los objetivos, las dificultades de elaboración, o de aplicación, la mala praxis o la utilización espuria o malintencionada de los resultados, por ejemplo, no han de cuestionar la validez y necesidad del principio, cuya calidad dependerá, fundamentalmente, de la validez de los indicadores, de la fiabilidad de los instrumentos y de las condiciones de aplicación.

Exposición inicial

Partimos de un hecho paradigmático de lo que queremos exponer y en absoluto exclusivo: en el manual de evaluación de la actividad docente del profesorado (Proyecto Docentia-ANECA), una de las subdimensiones consideradas para la valoración de la actividad docente del profesorado es el rendimiento académico de los alumnos -Tasa de Éxito, Tasa de Rendimiento y Tasa de Rendimiento Académico-

¿Qué fiabilidad puede tener la evaluación de la calidad de la actividad docente de un determinado profesor tomando como referencia el número de alumnos presentados al examen y el número de alumnos que de ellos superan dicho examen cuando es ese mismo profesor juez y parte en su propio proceso de evaluación?

A este respecto, García Landa, en su Blog (20 de abril de 2007), manifestaba una opinión que muchos compartimos:

Empiezan a saltar a la prensa noticias sobre rectorados presionando directamente a los profesores para que aprueben a más alumnos; incluso hay alguna universidad que planea introducir entre los índices de rendimiento y eficacia del profesor el número de alumnos aprobados. No aprobados por otros: ¡aprobados por él mismo! Vamos, esto es de traca, si no fuera porque entra tan de lleno en la lógica de los tiempos. O sea, que si quiero subir espectacularmente la calidad de mi docencia, sólo tengo que dar aprobado general a final de curso: todos aprueban, ergo misión cumplida, y profesor excelente.

¿Qué garantías de calidad en la formación de los estudiantes tiene el hecho de haber superado una determinada asignatura o incluso una titulación completa si no existe un sistema de evaluación objetivo y externo que certifique el grado de consecución de los objetivos y competencias previamente establecido y verificable a través de unos criterios de evaluación válidos y fiables?

¿Qué interpretación cabe hacer del hecho de que en una misma asignatura, un grupo -con un determinado profesor- obtenga unos resultados significativamente dispares a los del grupo contiguo -con otro profesor-?

¿Cabe, en esta situación, iniciar programa alguno de mejora? ¿Qué aspectos hemos de potenciar? ¿Cuáles hemos de subsanar?

Sin embargo, en nuestro sistema educativo –desde la educación infantil hasta la obtención del grado- no existe ni una sola prueba objetiva y externa –a excepción de la Prueba de Acceso a la Universidad, con todas sus objeciones- que certifique la calidad y el grado de consecución de los objetivos y competencias previstos en cada etapa.

En el ámbito universitario -si exceptuamos, tal vez, la Educación a Distancia donde la propia naturaleza de la formación no presencial, exige una uniformidad de contenidos, de objetivos y de criterios de evaluación y calificación para cada una de las asignaturas y titulaciones- no existe, sobre todo en determinadas titulaciones, ningún control externo –en muy pocos casos interno, de la propia universidad- sobre los objetivos propuestos, los contenidos impartidos, o los criterios de evaluación y calificación exigidos por cada profesor en su asignatura.

Se ha consagrado hasta tal punto el concepto de libertad de cátedra que cada profesor tiene la posibilidad de convertir su aula en un coto privado en el que goza de casi total impunidad; la calidad docente queda al arbitrio de la profesionalidad, ética y dedicación que cada profesor, libremente, quiera ejercitar.

Aceptamos como única garantía de calidad y de consecución de los objetivos propuestos, la calificación que cada profesor expide en su asignatura al albur de su criterio personal -o cada centro, según el nivel de exigencia que libremente quiera imprimir a sus titulaciones-. En este sentido se expresa Quintanilla (1998, p. 90):

No existe un criterio de calidad externo e independiente del que la propia institución (o el conjunto del sistema universitario o de sus profesores) define. Esto seguramente explica, aunque no justifica, algunos hechos bastante comunes en muchos cursos universitarios: no se usan libros de texto, a veces ni siquiera se especifica de antemano el programa de las asignaturas, en todo caso es frecuente que no se cumpla todo el programa y que cuando varios profesores imparten la misma asignatura, rara vez coincidan en sus contenidos y no digamos ya en sus enfoques, lo que hace difícilmente comparable el nivel de «calidad» de cada una de las diferentes prestaciones de lo que en teoría debería ser un mismo servicio.

No estamos con ello instigando, ni mucho menos, a una actitud de recelo generalizado –la mayoría de los profesores como la mayoría de los taxistas, ONGs, alcaldes o incluso concejales de urbanismo (permítaseme la broma) son excelentes profesionales- sin embargo, todos estaremos de acuerdo en que deben existir mecanismos de control externos que prevengan y actúen frente a los inevitables casos de mala praxis profesional, entendiendo por tal, no solo los casos aislados y tipificados, para los que el sistema ya tiene previsto un protocolo de actuación sino, fundamentalmente, la ineficacia docente, personal, institucional o del sistema – menos escandalosa pero mucho más grave-.

No es nuestro propósito, en este momento, valorar qué consecuencias han de derivarse de la evaluación por objetivos, competencias, resultados... –tema ciertamente polémico-, ni tan siquiera la validez de los resultados como indicador de la calidad docente e institucional, sino tan solo poner de manifiesto la necesidad de un sistema de evaluación, complementario en la proporción que se quiera, al de los profesores y centros, de modo que se constituya como un elemento más de garantía de los aprendizajes adquiridos.

Revisión de la literatura

Existe una corriente de pensamiento que tiende a contraponer las funciones formativa y certificadora de la evaluación ponderando la primera en detrimento de la segunda (Alonso, Asencio, López, Carriedo y Rycheck, 2007; Carless, 2003; Gibbs y Simpson, 2003; Santos, 1999, 2003; McDonald, Boud, Francis y Gonzi, 2000; García, 2004; Villardón, citado en Gairín, 2009, 69; Álvarez, 2001). Sintetizaremos los principales argumentos aducidos:

- Existe una influencia predominante de la cultura psicométrica, de la evaluación de tipo cuantitativo, sobrevalorando los criterios de fiabilidad y validez.
- Se mide mucho, se evalúa poco.
- Este tipo de evaluación no tiene una finalidad formativa sino fundamentalmente selectiva, clasificatoria. En términos coloquiales: “pesar al

pollo no significa darle de comer”. Coll (Citado en Pérez de Pablos, 2002, p. 32): “Lo que hacen las pruebas, en el mejor de los casos, es identificar el nivel de calidad, pero no mejorarlo”. Rodríguez Zapatero (Citado en Bedoya, 2002, p. 18): “No se consigue calidad con más exámenes. Con ellos se mide la calidad de unos y otros, y cuanto antes se mida, antes puede existir tendencia a la separación, pero no se mide la calidad. No por saltar mas vallas se será mejor atleta”.

- Se persigue la constatación de resultados, el “rendimiento” del estudiante con carácter meramente acreditativo: “resultadismo” -“taxonomización de los objetivos” lo denominó Santos Guerra (1998)- menospreciando el proceso, las causas del fracaso, los efectos secundarios o los aprendizajes no deseados.
- La evaluación centrada en productos, en contenidos “evaluables”, mayoritariamente conceptuales, con un estilo de aprendizaje memorístico, deja de lado aprendizajes importantes y favorece las metas orientadas al rendimiento por encima de las metas orientadas al aprendizaje, instigando la implantación de métodos de enseñanza orientados únicamente a la satisfacción de la prueba. Así lo expresa Angulo, (2009, p. 19): “[Las pruebas externas] determinan qué contenidos curriculares –por ser evaluados- son los que hay que valorar, y qué otros –por no ser evaluados- marginar; orienta la enseñanza y, a *Fortiori*, el aprendizaje hacia el formato técnico de las pruebas”. “No aprendemos para la vida sino para las pruebas PISA.”
- Sólo se evalúa a los alumnos, ni al profesor, ni al sistema.
- Los resultados obtenidos en un momento determinado pueden no ser significativos. El hecho de superar una prueba no implica necesariamente un mayor grado de competencia.
- Generalmente es una evaluación heterónoma, realizada casi exclusivamente por el profesorado, sin la participación de los principales agentes implicados, los alumnos, a través de la co-evaluación o la autoevaluación.
- Los resultados de las pruebas son definitivos, sin posibilidades de auto-mejora, centrados en determinar más las debilidades que las fortalezas del

aprendiz. Los errores se penalizan, no constituyen un elemento de aprendizaje.

- Por último, y tal vez la objeción más grave de todas, es la que denuncia que la evaluación objetiva de los resultados de aprendizaje y la toma de decisiones en función de dichos resultados perjudica a las personas y grupos sociales más desfavorecidos y favorece, por el contrario, a las clases privilegiadas. Este tipo de argumentos se sitúa en una corriente de pensamiento que queda perfectamente dibujada en esta cita de Miguel Ángel Santos Guerra:

Poner objetivos cuantificables, evaluar su consecución mediante pruebas estandarizadas, hacer clasificaciones elementales, realizar procesos atributivos interesados, distribuir los recursos mediante criterios coherentes con los resultados... He ahí una forma de hacer triunfar una rigurosa racionalidad. La ciencia es neutra, los números cantan. La calidad total, el control de calidad, los círculos de calidad... son, todas, expresiones que dan vueltas en la órbita de la sociedad neoliberal y que se convierten en trampas mortales para los «desheredados de la tierra» (1999, p. 8).

Discusión

Reflexionamos acerca de algunas de estas objeciones:

- Ciertamente, “pesar al pollo no significa darle de comer”; la mera medición, aislada, descontextualizada, sin consecuencias, es un ejercicio estéril que, en el mejor de los casos, solo produce pérdida de recursos y de tiempo pero, también es cierto, que no existe valoración de calidad y por consiguiente toma de decisiones en función de esa valoración si no se parte de un conocimiento cierto y profundo de aquello que se quiere valorar. Por lo tanto, la *medición*, entendiendo por tal la constatación de una circunstancia o estado en un momento determinado, al principio o al final de un proceso es imprescindible aunque no suficiente.

La evaluación es esencial para mejorar el sistema educativo. Si no se cuenta con información fiable sobre la situación de rendimiento de los alumnos y los aspectos que pueden determinarlo, resultará muy complicado iniciar procesos de mejora. Es necesario llevar a cabo análisis en profundidad de aspectos concretos y una mayor explotación de estos datos que

permitan extraer claves que orienten esos procesos de mejora. (MEC-INEE, 2011, p. 134)

- Una evaluación objetiva y externa no excluye la evaluación interna; ambas presentan ventajas e inconvenientes que es necesario conjugar (García Llamas, 2003, p. 279-284). Por ello, ambas son necesarias, complementarias y mutuamente enriquecedoras tal como recogen Kempfert y Rolf, (1999) y Bos y Postlethwaite (citado en Bos y Schwippert, 2009, p. 11). ¿Estamos ante dos formas distintas y ¿contrapuestas? –pruebas objetivas y calificaciones académicas- de valorar el rendimiento académico? A este respecto González (1988, p. 25) establece una distinción entre ambas que nos parece esclarecedora:

Las pruebas objetivas pretenden una medida controlada, carente de subjetivismo y analizable estadísticamente, de un conjunto de saberes o conocimientos. El caso de las calificaciones del profesorado es ligeramente distinto. Si bien en líneas generales se evalúa, de forma fundamental, la adquisición de conocimientos, en la evaluación académica entra también en juego un conjunto de factores que conllevan un elevado riesgo de subjetivismo: subjetivismo del profesor, relación profesor-alumno, *efecto Pigmalión*, política del centro a la hora de evaluar, consenso grupal en las sesiones de evaluación, etc.

[...] El complejo constructo del rendimiento escolar no puede reducirse a la adquisición de conocimientos reflejados en unas pruebas objetivas, aunque estas son un importante componente de aquel, pero, por otro lado tampoco podemos aceptar las calificaciones del profesorado como óptima medida del rendimiento, carentes de la posibilidad de un estudio metodológico objetivo y desconociendo la mecánica del complejo proceso de asignación de estas evaluaciones.

En el estudio “Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados”, Eurydice (2009, p. 65) se abunda en la necesidad de complementariedad de la evaluación externa e interna del rendimiento académico:

En la mayoría de los países europeos las pruebas nacionales tienen una gran importancia para los alumnos, ya que sus resultados determinan las decisiones sobre su trayectoria educativa posterior. En la mayoría de los casos, los resultados se valoran de forma conjunta con los de otras evaluaciones, especialmente la evaluación continua de los profesores y los exámenes internos. Este enfoque permite a los profesores tener voz en las decisiones que afectan a sus alumnos. Además, combina los puntos fuertes de los distintos instrumentos de evaluación, lo que permite paliar algunas de las limitaciones de las pruebas nacionales, especialmente la de que reflejan únicamente el rendimiento de los alumnos en un momento concreto y únicamente en algunas áreas.

Estas pruebas “se convierten en una garantía de que los certificados escolares son comparables independientemente del centro donde se obtengan”.

Dichas pruebas son simultáneas en todos los casos con las tradicionales evaluaciones continuas y sumativas en las que los profesores valoran aquellos otros aspectos de carácter menos cognitivo y objetivable.

Los países que tienen una larga tradición en la realización de pruebas nacionales se encuentran actualmente en un proceso de elaboración de políticas y estrategias específicas para lograr un equilibrio entre la evaluación realizada por el profesor o el centro y las pruebas y exámenes nacionales.

- Sin evaluación externa no existe la función certificadora de la evaluación. Queda invalidada *per se*. ¿Qué fiabilidad puede tener la evaluación de la calidad de la actividad docente de un determinado profesor o centro cuando es ese mismo profesor o centro juez y parte en su propio proceso de evaluación? Incluso los productos de mercado incluyen en su etiquetado, como garantía de calidad, el hecho estar certificados según determinada norma europea por un organismo independiente ¿Qué credibilidad aportaría que un productor elogiase la bondad de su producto sin un referente externo que lo avale?

Por ejemplo, el número de alumnos de un centro que superan las pruebas de acceso a la universidad, respecto del número de alumnos presentados por ese centro o los resultados de la evaluación de diagnóstico o de la evaluación PISA proporcionan, sin lugar a dudas, y al margen de cuantas objeciones quieran hacerse –muchas de ellas seguramente justificadas- una cierta garantía de calidad educativa o al menos una información significativa en absoluto despreciable e imprescindible como punto de referencia para la toma de decisiones. O aunque tan solo sea como mecanismo de control.

Por ello es imprescindible verificar la diferencia existente entre lo que la escuela y la universidad certifican que sus alumnos saben y lo que estos realmente saben; (Dickinson, 2000; Sullivan, 2002; Chisholm y Davis, 2007) sobre todo, tal como señala Álvarez (2008), en el caso de las universidades, si se pretende avanzar en los procedimientos de acreditación y homologación de los estudios a nivel internacional.

...Algún tipo de supervisión pública, externa a los centros, parece necesaria, habida cuenta la fuerte presencia de un sector privado y la evidencia de que puede haber centros que sobrecalifiquen a sus alumnos para atraer *clientes*, en detrimento de los públicos o privados más rigurosos. Es práctica común en todos los países educativamente desarrollados”. (Editorial, 2002).

Pero es que, además, una medición y valoración objetiva y fiable es el mejor aliado de la evaluación formativa ¿cómo mejorar sin saber con certeza de dónde partimos y adónde hemos llegado?

Los resultados del estudio de Serván (2011, p. 221) muestran que a pesar de las limitaciones, “las comunidades educativas valoran la información proporcionada por las pruebas de evaluación de diagnóstico y los planes de mejora asociados a ellas como un estímulo para la reflexión sobre los procesos de enseñanza y aprendizaje”. Igualmente Monereo y Morcillo (2011, p. 183) afirman que “las evaluaciones de carácter externo pueden actuar como un andamiaje educativo, suscitando cambios efectivos en las prácticas docentes que habitualmente se desarrollan en clase siempre y cuando se cumplan determinadas condiciones”.

En esta misma idea abunda el informe final del estudio realizado a petición de la Dirección General de Educación y Cultura de la Comisión Europea acerca de las causas del rendimiento escolar en función de los resultados de los estudios internacionales PISA, TIMSS y PIRLS (Henrik, 2005, 34):

Los análisis de datos multivariados de PISA 2000 y TIMSS 1995 proporcionan algunas pruebas de que los exámenes a nivel nacional tienen un impacto positivo en el rendimiento escolar en matemáticas y ciencias (Bishop, 1997, Wössmann, 2003, Fuchs y Wössmann 2004a) y los análisis de PISA 2000 (Fuchs y Wössmann 2004, p.16) en concreto indican que los exámenes finales externos podrían aumentar el rendimiento de los colegios autónomos, sirviendo como herramienta para controlar tal rendimiento.

- Por otro lado, creemos que la simple existencia de una valoración externa de los resultados supone ya, en sí misma, un estímulo, un acicate, una *presión*, ciertamente, que nosotros consideramos, en principio, positiva. “Las pruebas puntuales de reválida influyen en el esfuerzo de alumnos y profesores durante los cursos cubiertos por la misma” (Editorial, 2002); así lo reconocen y verbalizan muchos profesores de segundo de Bachillerato, único curso en el que existe una prueba externa de rendimiento académico. Es condición inherente a la naturaleza humana el rendimiento en función de las demandas.

Así se reconoce literalmente en el informe Mckinsey (2012, p. 9), Educación en España-motivos para la esperanza y en (Mckinsey, 2007, p. 64).

En términos generales, la publicación de los resultados de las evaluaciones externas genera mayor responsabilidad y conciencia pública, lo que a su vez impulsa mejoras adicionales. “Hacemos pública toda la información; esto crea tensión en el sistema – transparencia acerca de los problemas- e impulsa las mejoras”.

Ahora bien, la experiencia de muchos países nos dice también que:

“Si una escuela no sabe cómo mejorar, o si carece de capacidad de mejora, no mejorará su instrucción independientemente del grado de presión ejercido sobre ella [...] una mayor presión sólo lleva a conductas

regresivas [por ejemplo, preparar para el examen, poner excesivo énfasis en las preguntas de la prueba, impedir que los peores alumnos rindan el examen y conductas potencialmente fraudulentas]”. (Mckinsey, 2007, p. 65).

Es evidente. La recogida de información es solo una parte de la evaluación - imprescindible pero insuficiente-, después hay que interpretar esa información y arbitrar las medidas de mejora necesarias; todo ello y no alguna de sus partes, es lo que constituye la verdadera evaluación.

- Por *resultados de aprendizaje* entendemos el grado de consecución de los objetivos y el grado de desarrollo de las competencias previstos de antemano en el proceso de enseñanza-aprendizaje, sean estos cuales fueren, no solo, ni principalmente, a los de carácter cognitivo.

Ciertamente, resulta inevitable la tentación de orientar los esfuerzos, tanto por parte del profesorado como del alumnado, más a la satisfacción de la prueba que a la consecución de los objetivos y competencias previstos (Angulo, 2009); por ello, o bien concluimos que existen competencias que no pueden ser valoradas objetivamente y por lo tanto renunciamos a hacerlo, confiando a ciegas en el buen hacer de los docentes y en su apreciación subjetiva o bien, previendo este efecto colateral negativo de la evaluación externa, se aplican criterios e instrumentos de valoración cuya superación implique, necesariamente, haber desarrollado los objetivos y competencias deseados. Por ello, “algunos países han incorporado a sus pruebas un enfoque centrado en las competencias, y varios países evalúan determinadas competencias transversales” (Eurydice, 2009, p. 65).

No obstante, y a pesar de todo, creemos que la existencia de una valoración externa de los aprendizajes adquiridos –complementaria a la evaluación del profesor- acentúa los efectos positivos: promueve la utilización de estrategias de aprendizaje que favorecen la comprensión y la memoria a largo plazo, estimulan la motivación y alientan la voluntad del alumno y del profesor, liberando a este último de la responsabilidad absoluta de la calificación, recuperando así, la función genuina de la docencia y de la evaluación: la función formativa.

En todo caso, ¿Por qué vamos a dejar de valorar la consecución de aquellas competencias más fáciles de medir por la dificultad de hacerlo con aquellas otras

más difíciles de valorar? En el peor de los casos, mejor eso que nada. Si ni siquiera valoramos aquello que se ve, ¿cómo vamos a valorar aquello que no se ve? Si los resultados obtenidos en las competencias valoradas no fueran los esperables, ¿porque se ha de suponer que sí se están desarrollando el resto de competencias?

En casi todos los países donde se realizan pruebas estandarizadas externas se ha producido un intenso debate al respecto –véase (Eurydice, 2009, p. 60-62)- optándose en cada uno de ellos por diversas soluciones pero en ningún caso eliminando dichas pruebas, cuyas ventajas parecen estar fuera de toda duda.

- Otra objeción que con frecuencia se le imputa a la evaluación externa del rendimiento académico es que prima, la competitividad, la eficacia, la excelencia, los resultados... sobre el proceso o las circunstancias, atribuyéndosele un espíritu más empresarial, mercantilista, que pedagógico. No quisiéramos entrar en esta cuestión que por su marcado trasfondo ideológico excede nuestro espacio y pretensiones; tan solo quisiéramos apuntar al respecto, que el concepto de eficacia es definido por el DRAE (22º edición) como “la capacidad de lograr el efecto que se desea o se espera” y la excelencia como “la superior calidad o bondad que hace digno de singular aprecio y estimación algo”. Así las cosas, ¿quién no quiere la excelencia para sí y para los suyos?, la disensión estriba, como casi siempre, en la polisemia de los conceptos, en los objetivos pretendidos, en el modo de alcanzarlos o en el uso que de los resultados obtenidos se haga.

Evidentemente no es lícito alcanzar el éxito a cualquier precio, la forma, el proceso, sin duda, es fundamental; no obstante, al valorar los resultados de aprendizaje se está valorando, en gran medida, todo el proceso de enseñanza-aprendizaje: la obtención de unos resultados positivos nos permite inferir con cierta fiabilidad –no totalmente- la bondad del proceso de enseñanza-aprendizaje; por el contrario unos resultados negativos, nos obliga a revisar qué aspectos del proceso han sido deficitarios.

Cabe aducir que los resultados no siempre correlacionan con dicho proceso; No es fácil que eso ocurra; lo cual, no obstante, sería más bien atribuible a deficiencias del proceso evaluador que no ha sido capaz de desenmascarar esas incongruencias, bien por una mala definición de los objetivos, bien por la falta de validez de los

indicadores utilizados -no han medido lo que se quería medir-, por la escasa fiabilidad de los instrumentos o por deficiencias en el proceso de aplicación.

La evaluación ha de implicar a todos los agentes intervinientes, según su nivel y responsabilidad. Los éxitos y los fracasos no suelen responder a causas unilaterales o unifactoriales, sino multipersonales y multifactoriales; por ello ha de concebirse la evaluación de una manera global, que fomente el espíritu de cooperación, el sentimiento de corresponsabilidad; de lo contrario, suscita suspicacias y enfrentamientos entre los que se sienten examinados y los examinadores. En el ámbito educativo, la evaluación, utilizada tradicionalmente como medida concebida únicamente para medir el nivel de adquisición de conocimientos por parte de los alumnos, considerados como únicos responsables de su éxito o fracaso, ha de hacerse extensiva al resto de factores –humanos, materiales y organizativos- que interactúan con aquellos.

En consecuencia, ha de evitarse esa concepción negativa de la misma que supone el considerarla como el final de un camino tras el cual sólo cabe esperar una sentencia inapelable; por el contrario, la valoración de los resultados obtenidos ha de tener, fundamentalmente, un carácter marcadamente propedéutico. Así pues, no ha de percibirse como una amenaza externa y malintencionada, sino que si nuestra intención es realmente noble y contiene un deseo de mejora, la evaluación interna, por supuesto, pero también la externa constituye un reto, un baremo de nuestra posición relativa respecto al medio en que nos movemos; sin esta referencia externa perdemos objetividad y corremos el peligro de caer en el chovinismo y la autocomplacencia o en la apatía al no existir referentes a los que aspirar o ante los que responder.

- La objeción, a veces expresada, de que el hecho de superar una prueba externa no significa estar más capacitado que quien no lo hace ¿no pone en duda más la validez y fiabilidad de la capacidad discriminatoria de la prueba que su misma existencia?

- Los padres, los alumnos tienen el derecho de conocer el proyecto educativo, el ideario, los objetivos, el funcionamiento, los resultados... del centro educativo; los contribuyentes, el derecho a conocer en qué se invierten sus impuestos y con cuanta

eficacia se administran y ello solo es posible si existe una evaluación externa. No debería extrañar, por tanto, Tal como manifiesta Pérez Juste (2007, p. 25):

...que la sociedad asuma funciones de *control* sobre el sistema educativo; asegurarse de que se respetan las normas, que el personal cumple sus obligaciones, que los recursos invertidos se utilizan con eficiencia y eficacia es, por tanto, otra importante función de la evaluación; o mejor, dos importantes funciones, las conocidas desde Nevo como *administrativa* y *sociopolítica*.

- Por último, creemos que la evaluación externa de los resultados académicos, lejos de ser un elemento al servicio de los más privilegiados puede y debe erigirse en un elemento realmente democrático y promocional. Un sistema de valoración objetivo y externo, que exija unos mínimos estándares de calidad –los establecidas por Ley- a todos los alumnos y centros escolares, independientemente de sus condicionamientos personales, sociales, ubicación o titularidad, favorece, precisamente, a los más débiles del sistema que no disponen de más salida que su propia promoción personal para escapar del ambiente poco estimulante que les ha tocado en suerte. En esta misma idea abunda Fernando Savater (2002, p. 32):

Lo que no cabe es calificar de “elitismo” cualquier forma de valoración o selección entre esfuerzos desiguales. El verdadero elitismo perverso es el falso igualitarismo que hermana en una generosa mediocridad la pereza siempre respaldada por los papás de los niños mimados con el esfuerzo de quienes se empeñan en mejorar la modestia de sus orígenes. Si ningún derecho jerárquico se recibe en la cuna, tampoco el de ser bachiller superior o el de entrar como fuere en cualquier universidad [...] pero en cualquier caso, me rebelo contra que la única forma de distinguir la excelencia sean las votaciones populares de *Operación Triunfo*, con toda su estomagante parafernalia lacrimosa y sus ventajismos mediáticos. ¿Va a ser ésa la única “reválida” que cuente con el consenso de jóvenes y adultos?

Resulta difícil sustraerse a la aparente injusticia que supone utilizar los mismos criterios de evaluación con quienes son desiguales entre sí. No se puede valorar, dicen, con el mismo baremo a un centro privilegiado por su ubicación, por el origen de su alumnado,

por la abundancia de recursos, etc. que a otro con unas condiciones adversas. Se instiga, incluso desde los organismos oficiales a que se tenga en cuenta de dónde parte cada uno a la hora de valorar sus resultados, dándose por satisfechos si los segundos han mejorado a pesar de que se encuentren aún lejos de los primeros.

Y, sin embargo, nosotros creemos que no debe ser así. La ausencia de una evaluación objetiva y externa de unos mínimos resultados de aprendizaje o el establecimiento de distintos niveles de exigencia en función de las circunstancias oculta las diferencias artificial y temporalmente pero no las elimina, se convierte en una trampa sutil precisamente para aquellos a los que se pretende beneficiar; la realidad acaba imponiéndose. Una actitud paternalista que fomente la discriminación positiva de los más desfavorecidos, en cuanto a la valoración de sus aprendizajes pero sin facilitarles cuantas posibilidades requieran para acceder a ellos, supone una infravaloración de sus capacidades y una rendición encubierta de la función compensatoria del sistema educativo, que vendría a reconocer subrepticamente dos categorías, al estilo de los *olímpicos y para-olímpicos*, perpetuando y dando carta de naturaleza a esa fractura social.

Por el contrario, para lo que hay que tener en cuenta las circunstancias sociales o personales no es para ser *condescendientes* con los menos favorecidos y mucho menos para reducir los recursos de los centros que precisamente más lo necesitan, sino, muy al contrario, para compensar las deficiencias, multiplicando los recursos, eliminando los guetos, evitando la selección encubierta de los alumnos por parte de algunos centros sostenidos con dinero público so pretexto de libertad de elección de los padres... con el fin de lograr, en la medida de lo posible, la igualdad de oportunidades que les permitan alcanzar las metas de las que cada uno de ellos sea capaz. Así lo recoge el artículo 71 de la LOE en referencia a los *alumnos con necesidad específica de apoyo educativo*:

”...Corresponde a las Administraciones educativas asegurar los recursos necesarios para que los alumnos y alumnas que requieran una atención diferente a la ordinaria [...] pueden alcanzar el máximo desarrollo posible de sus capacidades personales, y en todo caso, los objetivos establecidos con carácter general para todo el alumnado”.

Y añade en el artículo 80.1 y 2:

“Con el fin de hacer efectivo el principio de igualdad en el ejercicio del derecho a la educación los Poderes públicos desarrollarán las acciones de carácter compensatorio en relación las por personas, grupos y ámbitos territoriales que se encuentren en situaciones desfavorables y proveerán los recursos económicos para ello”.

“Las políticas de educación compensatoria reforzarán la acción del sistema educativo de forma que se eviten las desigualdades derivadas de factores sociales, económicos, culturales, geográficos, étnicos o de otra índole”.

En definitiva, en términos generales, la solución no está en *bajar el listón*, sino en *aupar al niño*. Ese es el reto: compatibilizar la extensión de la oferta educativa con la calidad de la misma, más aún en el caso de la enseñanza post-obligatoria, especialmente, en la etapa universitaria, en la cual, aun siendo necesario considerar, por razones pedagógicas -en aras a los principios del aprendizaje significativo- el punto de partida de los estudiantes, ha de priorizarse la consecución de las competencias académicas y profesionales previstas para cada titulación en cada universidad, al margen de cualesquiera circunstancias personales.

La evaluación externa en el ámbito internacional

La consultora estadounidense Mckinsey (2007) publicó uno de los más importantes estudios a nivel mundial de educación comparada: *¿Cómo hicieron los mejores sistemas educativos del mundo para alcanzar sus resultados?* en esa misma línea se publica en 2010 un segundo informe: *¿Cómo los sistemas educativos que más mejoran continúan mejorando?*, en ellos se analizan las características comunes que comparten los mejores sistemas educativos del mundo y qué hicieron aquellos que más han mejorado. De entre todas las medidas puestas en práctica nosotros nos vamos a centrar en la que en este trabajo nos ocupa: las evaluaciones estandarizadas de los estudiantes y de los centros educativos. Traemos a colación algunas de las consideraciones que dicho informe manifiesta al respecto.

- Todos los sistemas con alto rendimiento reconocen que es imposible mejorar algo que no se mide. Constatar los resultados les permite identificar y difundir mejores prácticas, indicar con precisión las áreas más débiles y hacer que las escuelas respondan por sus resultados.

- Por lo general, la intensidad de esta constatación de los resultados es inversamente proporcional al buen funcionamiento del centro, es decir, a medida que el sistema educativo mejora disminuye la intensidad de la evaluación externa.

- La evaluación de los resultados garantiza que el sistema tenga la información que necesita para poder intervenir cuando las escuelas comienzan a mostrar dificultades.

- Los sistemas con mejores resultados utilizan dos mecanismos para evaluar la calidad de la enseñanza y el aprendizaje: los exámenes y las inspecciones oficiales. Los exámenes evalúan los conocimientos, la comprensión y las habilidades de los alumnos y ofrecen una medida objetiva de los resultados reales con alto grado de detalle.

- Los exámenes tienen además un poderoso efecto para impulsar el rendimiento de cualquier sistema educativo. “Lo que se evalúa es lo que se aprende, y cómo se evalúa determina cómo se lo aprende”.

- Las inspecciones escolares a diferencia de los exámenes, miden tanto los resultados como los procesos que los impulsaron, y como consecuencia de ello pueden ayudar a las escuelas y a los sistemas a identificar áreas específicas que requieren mejoras. Las inspecciones escolares también permiten a los sistemas medir algunos resultados deseados más sutiles y complejos, que resultan difíciles o imposibles de medir por medio de exámenes.

- En muchos de los sistemas con mejor rendimiento, los responsables de la evaluación de los resultados son agentes externos al proceso de enseñanza-aprendizaje: “La gente responsable por la mejora de la educación no puede ser la misma que determine si las mejoras necesarias han sido efectivamente implementadas o no”.

En definitiva parece indiscutible que la evaluación externa del rendimiento académico de los estudiantes y los centros escolares es un elemento fundamental de calidad educativa; ahora bien debe estar encaminado fundamentalmente a la detección de dificultades con el fin de subsanarlas multiplicando los recursos precisos en dichos centros

Hanushek y Woessmann en “The Economics of International Differences in Educational Achievement” (2012, 25) afirman que “Students in countries that have external exit-exam systems very consistently perform significantly and substantially better on the international student achievement tests than students in countries without external exit-exam systems”.

La OCDE (MEC, 2012, 47), basándose en los resultados del análisis de regresión bivariante que relacionan la puntuación en Lectura obtenida por el alumnado de 15 años y la existencia de pruebas externas estandarizadas manifiesta:

Los resultados del estudio PISA 2009 ponen de manifiesto que en los países en los que se aplican las pruebas externas estandarizadas, el rendimiento del alumnado es mayor, incluso descontando el efecto de otros factores que puedan condicionar dicho rendimiento. Concretamente, en los países cuyos sistemas de evaluación incluyen exámenes externos estandarizados, la puntuación en lectura es 16 puntos superior de media que en los que no se aplican pruebas de estas características.

De los treinta y cinco países de la red Eurydice sólo cinco no realizan ninguna prueba estandarizada nacional. El resto de países utilizan dichas pruebas con tres objetivos fundamentales: balance de los alumnos a nivel individual “evaluación del aprendizaje”; evaluar los centros y/o el sistema educativo y la calidad del profesorado e identificar necesidades individuales de aprendizaje “evaluación para el aprendizaje”.

Se considera que las pruebas nacionales son necesarias para obtener una medición equiparable y estandarizada del rendimiento escolar. El debate se centra, por lo general, sobre el contenido, el formato y la organización de las pruebas y sobre la utilización que se hace de sus resultados. Una cuestión clave es la necesidad de garantizar la validez y la pertinencia de las pruebas

nacionales, así como su exactitud técnica, su objetividad y su rentabilidad. (Eurydice, 2009, p. 63)

Por último un asunto especialmente polémico por sus implicaciones éticas y prácticas –que en esta ocasión no abordaremos- es el mayor o menor grado de publicidad de los resultados obtenidos en las evaluaciones. Cada país mantiene una política diferente al respecto en función de sus ventajas e inconvenientes, de sus circunstancias particulares y de los objetivos que persigan.

Conclusiones

Casi todos los países y todos los informes al uso -los informes PISA de la OCDE, los informes sobre el panorama de la Educación de la OCDE, los informe del seguimiento de la educación para todos en el mundo de la UNESCO, los estudios internacionales TIMSS y PIRLS del IEA, el informe Eurydice sobre las pruebas nacionales de evaluación del alumnado en Europa, los informes McKinsey, el informe de Hanushek y Woessmann: “The Economics of International Differences in Educational Achievement” o los informes sobre el sistema educativo español de la fundación BBVA-IVIE- contemplan en mayor o menor medida la evaluación externa como garantía imprescindible e indiscutible de calidad educativa en sus distintos niveles de concreción curricular: a nivel individual –los alumnos-, a nivel intermedio – los centros- y a nivel nacional –los sistemas educativos-.

Ello no es óbice para que esté exenta de múltiples problemas como hemos puesto de manifiesto a lo largo del trabajo, tanto a nivel teórico como en su aplicación práctica: peligros o inconveniencias del contenido, del formato, del riesgo de la utilización espuria de los resultados, de la pertinencia de su publicidad... a pesar de lo cual consideramos que las ventajas exceden a los inconvenientes. En todo caso, en todos los países –así lo contemplamos también nosotros- la evaluación externa es complementaria, en mayor o menor grado, de la evaluación realizada por los profesores; dicha complementariedad minimiza los riesgos potenciales o reales que presentan cada una de ellas por separado.

Financiación

Sin financiación expresa

Conflicto de Intereses

Ninguno”.

Referencias bibliográficas

- Alonso, J., Asencio, F., López, I., Carriedo, N., y Rycheck, F. (2007). Evaluación para aprender: Las ideas de los profesores de ciencias sociales sobre la evaluación del entendimiento causal. *Electronic Journal of Research in educational Psychology*, 5 (3), 593-616.
- Álvarez, J.M. (2001). *Evaluar para conocer, examinar para excluir*. Madrid: Morata.
- Álvarez, I. (2008). Evaluación del aprendizaje en la Universidad: una mirada retrospectiva y prospectiva desde la divulgación científica. *Revista Electrónica de Investigación Psicoeducativa*, 14, Vol. 6 (1), 235-272.
- Angulo, F. (2009). Evaluaciones estandarizadas, una vuelta al pasado. *Docencia*, 38.
- Bedoya, J. G. (2002, 6 de febrero). Zapatero llama a movilizarse contra la regresión educativa. *El País*, p. 18.
- Bos, W. y Schwippert, K. (2009). TIMSS, PISA, IGLU y demás: Razón y sinrazón de los estudios internacionales de rendimiento escolar. *Profesorado. Revista de Currículum y Formación de Profesorado*, 13 (2), 1-15.
- Carless, D. (2003, November). *Learning-oriented assessment*. Paper presented at the evaluation and assessment conference. Adelaide, University of south Australia.
- Chisholm, C., y Davis, M. (2007). Analysis and evaluation of factors relating to accrediting 100% of prior experiential learning in UK work-based awards. *Assessment and Evaluation in Higher Education*, 32 (1) 45-59.

Dickinson, M. (2000). An analysis of the accreditation of transferable skills in extra curricular activities within higher education. *Assessment and Evaluation in Higher Education*, 25 (1) 61-70.

Editorial (2002, 10 de febrero). *El País*.

EURYDICE (2009). Pruebas nacionales de evaluación del alumnado en Europa: objetivos, organización y utilización de los resultados. Brussels: Education, Audiovisual and Culture Executive Agency. Disponible en: http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/109ES.pdf [consulta 2013, 15 de octubre]

Gairín, J., Carbonell, J.L., Paredes, J. y Santos Guerra, M. Á. (2009). *Glosario*. En Paredes, J. (Coord.), De La Herrán A. (Coord.), Santos Guerra, M.Á., Carbonell, J.L., y Gairín, J. (373-378) *La práctica de la innovación educativa*. Madrid: Síntesis.

García Llamas, J. L. (2003). *Métodos en investigación en educación*. (V. II) Madrid: UNED.

García Landa, (2007, 20 de abril). La evaluación externa. Mensaje abierto.

Gibbs, G., y Simpson, C. (2003). Does your assessment support your students' learning? Disponible en: <http://arts.unitec.ac.nz/documents/GrahamGibbAssessmentLearning.pdf>. [Consulta [2011, 12 de diciembre].

Hanushek, E y Woessmann, L. (2012). "The Economics of International Differences in Educational Achievement", (National Bureau of Economic Research, USA, 2010). *IZA Discussion Paper*, 4925. Disponible en: <http://ftp.iza.org/dp4925.pdf> [consulta 2013, 30 de octubre]

Henrik, J. (2005). *Explicación del rendimiento escolar. Resultados de los estudios internacionales PISA, TIMSS y PIRLS*.(Informe final). Danish Technological Institute.

McKinsey and Company. (Barber, M. y Mourshed, M.) (2007). Cómo hicieron los sistemas educativos con mejor desempeño del mundo para alcanzar sus objetivos. *PREAL*, 41. Disponible en:

http://mckinseysociety.com/downloads/reports/Education/Como_hicieron_lo_s_sistemas_educativos.pdf [consulta 2013, 30 de septiembre]

McKinsey and Company (2012). Educación en España. Motivos para la esperanza. Disponible en: http://dl.dropboxusercontent.com/u/58452177/Educacion_en_Espana_Motivos_para_la_esperanza.pdf [consulta 2013, 30 de septiembre]

MEC (2012) PIRLS-TIMSS 2011. Estudio Internacional de Progreso en Comprensión lectora, matemática y Ciencia IEA Volumen I: Informe español. Disponible en: <http://www.mecd.gob.es/dctm/inee/internacional/pirlstimss2011vol1.pdf?documentId=0901e72b8146f0ca> [consulta 2013, 15 de julio]

MEC (2012). Panorama de la Educación. Indicadores de la OCDE 2012. Informe español. Disponible en: <http://www.todofp.es/dctm/todofp/panorama-ocde2012.pdf?documentId=0901e72b81416fd3> [consulta 2013, 20 de julio]

Pérez de Pablos, S. (2002, 6 de febrero). Los expertos y la oposición ponen en duda que la reválida mejore la calidad de la enseñanza. *El País*, p. 34.

Pérez Juste, R. (2007). La evaluación externa y sus implicaciones. Aspectos técnicos, prácticos y éticos. *Avances en Investigación Educativa*, 6. Disponible en: http://adide.org/revista/index.php?option=com_content&task=view&id=197&Itemid=47. Consulta: [2011, 15 de julio].

Santos, M. A. (1998). *Evaluar es Comprender*, Buenos Aires: Magisterio del Río de la Plata.

Santos, M.A. (1999). 20 Paradojas en la evaluación del alumnado en la Universidad Española. *Revista electrónica Interuniversitaria de Formación del Profesorado*. Disponible en: <http://www3.uva.es/aufop/publica/revelfop/v2n1bl2>. Consulta [2011, 25 de mayo].

Santos, M. A. (2003). *Una flecha en la Diana: La evaluación como aprendizaje*. Madrid: Narcea.

Savater, F. (2002, marzo 8). Motivación e indiferencia. *El País*, p. 22.

- Serván, M^a J. (2011). Percepciones e interpretaciones de las comunidades educativas en torno a la evaluación externa de los aprendizajes escolares. Estudios de caso sobre la evaluación de diagnóstico en Andalucía. *Cultura y Educación*, 23, 2.
- Sullivan, K. (2002). Credit and grade transfer within the European's Union SOCRATES Programme: unity in diversity or head in the sand? *Assessment and Evaluation in Higher Education*, 27 (1), 65-74.