

El debate sobre las pruebas estandarizadas de evaluación y el papel de los orientadores

Daniel Santacruz García¹

Resumen

Las pruebas estandarizadas se han convertido en los últimos años en objeto de debate en la comunidad educativa y en la sociedad. Estas pruebas presentan diversas ventajas e inconvenientes que se exponen a continuación. En el presente trabajo pretende ofrecerse un resumen de estos pros y contras y una reflexión sobre los mismos; y por último, una serie de propuestas que la orientación educativa tiene en su mano en este debate. Las pruebas estandarizadas tienen una utilidad, pero no pueden usarse como única medida de la calidad y la eficacia de un sistema educativo.

Palabras clave: orientación educativa, pruebas estandarizadas, incidencia política

1. Introducción

Las pruebas estandarizadas son evaluaciones creadas *a priori* para medir una o varias características en una población. Al administrarse y calificarse a una muestra representativa siguiendo un procedimiento determinado se obtiene información sobre dicha muestra y sobre la posición de un individuo o un grupo concreto en ella.

El debate existente sobre el uso masivo de estas pruebas es uno de los más proliferos para la comunidad educativa de nuestro país y sus argumentos se presentarán a continuación. Después, se plantearán qué actuaciones pueden llevarse a cabo desde la orientación educativa para contribuir a aclarar esta polémica.

¹ Psicólogo General Sanitario. Sexólogo. Trabaja como psicólogo en la clínica privada (AVANCE, Terapia y Más). Email: santacruz.dani@gmail.com

2. ¿Cuál es el debate sobre la evaluación estandarizada?

Las evaluaciones estandarizadas comenzaron a utilizarse en el contexto educativo en EEUU. En España, estas mediciones aterrizaron en los años ochenta con la Reforma Experimental de las Enseñanzas Medias (Álvaro et al., 1988), que supondría la semilla de la ESO. Durante los años noventa, las competencias en Educación se transfirieron a las Comunidades Autónomas y se creó un instituto específico –el INCE– para que ejerciera oficialmente las primeras labores de evaluación estandarizada.

La LOGSE (1990) fue la primera ley en contemplar la importancia de las evaluaciones globales para la mejora de la calidad. Sus evaluaciones se centraban más en la acción en el aula que en aspectos estructurales, pero ya incluían como objeto de evaluación tanto a los estudiantes como a los docentes, centros educativos, los procesos educativos y la administración (Puente, 2000).

En la actualidad, el sistema educativo español realiza dos estrategias de evaluación estandarizadas. Por un lado, el *informe PISA* que, desde el año 2000, se realiza a nivel mundial y mide el rendimiento académico de los estudiantes de quince años en matemáticas, ciencia y lectura (Maestro, 2006); tiene lugar cada tres años y pretende ofrecer información para que los países puedan mejorar sus políticas de educación. Por otro lado, las *pruebas de nivel* implementadas por la LOMCE (2013), en 3º y 6º de Primaria y en 4º de ESO, con una finalidad tanto diagnóstica, es decir, conocer puntos fuertes y débiles del sistema educativo; como muestral, presentar aspectos generales de la educación en nuestro país.

2.1. Ventajas del uso de pruebas estandarizadas

Tras la II Guerra Mundial comenzaron a buscarse y desarrollarse procedimientos de evaluación colectiva de la inteligencia para la selección de soldados. El auge de la psicometría amplió su uso a otros contextos como la selección de personal o la evaluación educativa, donde se medía el grado de adquisición de competencias definidas previamente como básicas (lectura, escritura, cálculo).

Este tipo de pruebas resultan más económicas que las individualizadas puesto que pueden aplicarse y corregirse rápidamente, permiten procesos de evaluación en masa, requieren de poca formación por parte del evaluador y facilitan la información a todos los

agentes implicados en el sistema educativo como familias, docentes o políticos (Calero y Padilla, 2004; Froemel, 2009).

El objetivo de cualquier evaluación, estandarizada o no, es obtener información; y ahí radica una de sus principales ventajas: las evaluaciones educativas permiten disponer de datos objetivos sobre el sistema educativo.

Estos datos permiten desmontar mitos e ideas preconcebidas, es decir, proporcionar información empírica que permitirá, a su vez, desarrollar actuaciones basadas en la evidencia (Spillane, 2012; Coburn, Hill y Spillane, 2016).

La información obtenida a través de pruebas estandarizadas ayuda a conocer mejor la realidad educativa de elementos individuales del sistema (estudiantes, docentes o directivos), elementos intermedios (centros educativos o programas) y elementos globales (el sistema educativo en sí, a nivel estatal o autonómico). Sus resultados, pues, permiten el diagnóstico y seguimiento de distintos procesos de enseñanza-aprendizaje (Skedsmo, 2011). Además, propicia el análisis de las políticas educativas, el seguimiento sobre la evolución del sistema y la comparación con otros sistemas (Puente, 2000; Maestro, 2006); lo que puede utilizarse para mejorar.

Para otros muchos autores (Eurydice, 2009; Darling-Hammond y Falk, 2013; Hopfenback, Flórez y Tolo, 2015) estas evaluaciones mejoran la enseñanza-aprendizaje, a largo plazo, mediante el *feedback* que proporcionan, pues sirve para promover el aprendizaje, contribuir a que docentes y estudiantes participen en la aplicación del currículum, motivar a estudiantes y docentes, etc.

En nuestro país (Monarca y Fernández-Agüero, 2018), los docentes señalan que la función diagnóstica de las pruebas es la fortaleza más valorada: ayudan a detectar áreas en las que trabajar más, permiten detectar necesidades específicas y ayudan a mejorar los procesos de enseñanza-aprendizaje.

Parece pues que las evaluaciones estandarizadas persiguen el ideal de mejora de la calidad del sistema educativo. Pero, como sucede en cualquier otro ámbito, este ideal utópico puede quedar reducido o eclipsado por usos más suculentos de los datos o por mediciones sesgadas. Especialmente a raíz del uso de esta información como estrategia de rendición de cuentas a la sociedad.

2.2. *Inconvenientes del uso de pruebas estandarizadas*

Un principio básico en metodología es que no existen pruebas infalibles. Cualquier evaluación ofrece una medición sujeta a error, esto es, si mide lo que está diseñada para medir (eficiencia o validez) y si lo mide adecuadamente (eficacia o fiabilidad).

Las pruebas estandarizadas en educación están limitadas por diversos factores, la proporción total que realmente pueden medir suponen sólo la punta del iceberg del sistema educativo (Froemel, 2009). Y sin embargo, en los años recientes se experimenta cierta obsesión con los datos aportados por una prueba específica –el informe PISA–, por pruebas de etapa –especialmente en 4º ESO–; o por una nota final.

Tomar la parte por el todo y entender que una prueba, con sus limitaciones inherentes, es *el* indicador de calidad del sistema educativo o del aprendizaje provoca graves distorsiones y perversiones.

Estas pruebas priorizan los resultados sobre los procesos y contextos, es decir, importa más que sepan y no cómo ni para qué lo saben (Fernández Navas, Alcaraz y Sola, 2017); lo que mantiene al sistema educativo español atrapado en el academicismo, la memorización y la reproducción de contenidos abstractos, descontextualizados y desconectados de los problemas reales de la vida cotidiana (Monarca 2012; Pérez Gómez, 2014; Conley, 2015).

Al centrarse en datos estáticos, tampoco permiten la evaluación de la evolución, es decir, las pruebas sólo informan de cómo están los estudiantes en comparación a ciertos estándares; pero no cómo han evolucionado –medidas *pre-post*–, si lo que saben es producto del sistema educativo u otros contextos, etcétera (Froemel, 2009; Monarca, 2015).

También se ha argumentado cómo empobrecen el currículum educativo (Klenowski y Wyatt-Smith, 2012; Monarca, 2012; Feniger, Israeli y Yehuda, 2015) mediante la estandarización del trabajo docente, que tiende a centrar la enseñanza en la superación de pruebas, el llamado *teaching to the test*, un ejemplo de cómo la evaluación influye en el objeto evaluado.

Estos fenómenos han supuesto las condiciones perfectas para el progresivo desarrollo de la peor consecuencia posible para nuestro sistema educativo: la jerarquización de las escuelas.

Al hacer depender únicamente de las pruebas estandarizadas aspectos como la continuidad en el sistema educativo, el reconocimiento del estudiante, la reputación del centro e incluso la financiación, se genera una obsesión finalista con estas pruebas.

Se dan así un sinnúmero de fenómenos como los rankings de colegios, legalmente prohibidos pero disponibles en la red e incluso en la prensa (Monarca, 2012; Koch y DeLuca, 2012; Bourke y Mentis, 2013; Simmonds y Webb, 2013, Pérez Gómez, 2014; Polesel, Rice y Dulfel, 2014; Bacon, 2015; Feniger et al., 2015) y la consiguiente concentración de recursos en estas escuelas; el sacrificio de asignaturas que no puntúan, como las artísticas o las filosóficas, necesarias para la creatividad y el razonamiento crítico (Barrenechea, 2010; Bayona, 2013; Pérez Gómez, 2014); las estratagemas para engañar a las pruebas, como alterar sus resultados, invitar a algunos estudiantes no acudir el día de la prueba (Haladyna, Bobbit-Nolen y Haas, 1991); la presión y ansiedad en estudiantes y docentes (Smith, 2014; Winter, 2017), etcétera.

El resultado: una progresiva mercantilización del sistema educativo, propia de sistemas neoliberales (Feniger et al., 2015; Pettersson, Popkewitz y Lindblad, 2016; Komatsu y Rappleye, 2017) y un aumento de las desigualdades (Falabella, 2014; Smith, 2014).

Además, se ha señalado en la literatura que los efectos de estas pruebas pueden tener otros efectos negativos: perjudicar a aquellos estudiantes que no se ajustan al funcionamiento de la prueba (Barrenechea, 2010); favorecer las motivaciones extrínsecas para aprender, esto es, por notas y puntuaciones (Armstrong, 1999); obviar la evaluación a profesores, centros educativos o contextos familiares (Fernández Navas et al., 2017); ignorar metodologías menos mecanicistas para enseñar-aprender; no considerar el estilo de aprendizaje que caracteriza a los jóvenes actuales (Pérez Gómez, 2014); y por último, no proporcionar conocimientos profundos para la práctica docente ni ofrecer recomendaciones u orientaciones para la mejora (Conley, 2015; Klenowski y Wyatt-Smith, 2012).

En la encuesta global realizada en España por Monarca y Fernández-Agüero (2018), la debilidad más señalada de las pruebas se centraba en la absoluta descontextualización de las mismas con respecto a las individualidades de cada estudiante, centro o familia; y, especialmente en la Comunidad de Madrid, el efecto ranking provocado por su uso.

Si los efectos del peso de una o varias pruebas no podían ser más perniciosos, se les suman las consecuencias sobre la opinión pública que tienen los resultados. A nivel individual, las puntuaciones contribuyen a etiquetar a los y las estudiantes y a que se articule el famoso efecto Pigmalión. Por otro lado, la población general, lega en cuanto al funcionamiento de la educación, entra en esta lógica obsesiva por las notas y acaba pensando que las notas son indicadores fiables del rendimiento. Por último, los datos y resultados acaban utilizándose como arma política entre los principales partidos (Barrenechea, 2010; Skedsmo, 2011; Koch y DeLuca, 2012).

De hecho, parece que son las únicas consideraciones tenidas a cuenta para las reformas educativas cuando el sistema educativo cojea de muchas de sus patas, y no todas están siendo evaluadas por pruebas estandarizadas; por ejemplo, el abandono escolar prematuro o las altas tasas de depresión y ansiedad entre el cuerpo docente.

3. ¿Cómo garantizamos la evaluación de la calidad del sistema educativo?

Parecería pues que las pruebas estandarizadas son una suerte de mal endémico del sistema educativo. Nada más lejos de la realidad. Siguiendo la lógica de “matar al mensajero” se tiende a culpar al uso de estas pruebas cuando realmente los principales problemas están en *el uso del uso* de estas pruebas.

La corriente política que promueve la mercantilización del sistema educativo requiere necesariamente de herramientas que justifiquen su postura. Las pruebas estandarizadas utilizadas para ordenar y separar –que no informar o diagnosticar– cumplen este cometido a la perfección. Por tanto, el cambio legislativo es necesario. Desgraciadamente, la instauración de los cambios propuestos por la LOMCE no augura un futuro clarificador (Viñao, 2015).

Las pruebas estandarizadas son útiles para evaluar y deberían seguir utilizándose pero no como únicos métodos de recogida de información. Desde los años setenta se maneja el término de “evaluación formativa” para referirse al proceso de evaluación constante de los avances en relación a objetivos de aprendizaje entre docentes y estudiantes. Es decir, un tipo de evaluación que permita establecer objetivos de aprendizaje, analizar las capacidades y errores de cada estudiante en base a los objetivos y desarrollar y usar estrategias para promover la mejora de la enseñanza (Morales Vallejo, 2009; Agencia de Calidad de la Evaluación, 2016).

Para ello, se propone el uso de otras estrategias de evaluación como la observación del cuerpo docente, la auto-evaluación del estudiante, el trabajo por proyectos o en pequeños grupos, el uso frecuente de pruebas cortas, la creación de portafolios, etcétera. Es decir, un sistema de evaluación mucho más parecido al universitario.

Esto requeriría obviamente una mayor profesionalización del cuerpo docente, algo que el pensamiento conservador impide sibilamente como argumento contra el sistema educativo. También implicaría el cambio de aspectos organizativos en las escuelas como una mayor autonomía a los centros, la implementación de estilos de liderazgo alternativos en las comunidades educativas (Bolívar, López y Murillo, 2013), la disminución de la ratio estudiantes-docentes por aula (Pérez Gómez, 2014), la creación de trayectorias formativas flexibles (Froemel, 2009) y un largo etcétera que puede resumirse en un estilo de educación más personalizada.

El sistema educativo sienta las bases de cualquier sociedad, por tanto, como sociedad tenemos la obligación de evaluar el sistema educativo vigente, analizar qué mejoras se pueden realizar y pactar cuáles son las metas finales.

4. ¿Qué puede hacer la orientación educativa?

Se pueden distinguir dos principales y posibles objetivos. En primer lugar, se trataría de *informar* a la comunidad educativa. Por lo general, muchas familias e incluso muchos docentes pueden no conocer los pros y contras de las evaluaciones estandarizadas ni los oscuros vericuetos que pueden ofrecer sus implicaciones con el sistema político. Dentro de las funciones de la orientación educativa se encuentra la coordinación con la comunidad educativa en aras de la participación (Bisquerra, 1998).

Un paso más consistiría en realizar tareas de asesoramiento comunitario (Grañeras y Parra, 2009), es decir, tratar de mejorar la capacidad de los sujetos para tomar decisiones (Rodríguez Romero, 1996; Vélez de Medrano, 2003; Vélez de Medrano, 2005; Rodríguez Romero, 2006). Este asesoramiento se focaliza en grupos organizados como redes de colaboración y uno de sus objetivos (Vélez de Medrano, 2003) consiste en “capacitar a sus miembros para la participación democrática en la vida de la comunidad a la que pertenecen”. En este contexto, el rol del orientador es el de asesorar a dichos grupos.

Desde esta perspectiva, cuando existe información y concienciación sobre los problemas asociados, otra posibilidad de la orientación educativa consistiría en coordinar

la participación del entorno en una de sus máximas expresiones, que apenas suele citarse en los manuales generales: la *incidencia política*. Efectivamente, los cambios legales del sistema educativo corresponden a los más altos niveles curriculares. En los últimos años, la organización de movimientos sociales (como las mareas verdes) se han manifestado y han luchado por la incorporación de cambios al sistema educativo español, entre ellos, la realización de pruebas estandarizadas.

Sería función del orientador educativo conducir los intereses participativos de alumnos, profesores y familias hacia la colaboración con representantes políticos. Por ejemplo, llevando a cabo mapeos y análisis políticos, contactando con tomadores de decisiones y figuras públicas que puedan abanderar la lucha, buscar posibles aliados, etc. En definitiva, coordinar la búsqueda del cambio en su nivel más macro: el político.

5. Conclusiones

Las pruebas estandarizadas ofrecen una simple radiografía de la realidad de la enseñanza. Utilizarlas para informar de la calidad del sistema educativo sería como utilizar únicamente un mapa hidrográfico para visitar una ciudad.

Es obvio que una prueba colectiva homogeneiza a los participantes, pero por eso se utilizan otras estrategias de evaluación adicionales. Cambiar esta realidad tiene que partir de una sociedad concienciada y participativa, y el catalizador de estos cambios puede ser la orientación educativa.

Referencias bibliográficas

- Agencia de Calidad de la Educación (2016). *Guía de Evaluación formativa* Santiago de Chile: Agencia de calidad de la Educación.
- Álvaro, M., et al. (1988). *Evaluación externa de la Reforma Experimental de las Enseñanzas Medias*. Ministerio de Educación y Ciencia, Madrid.
- Armstrong, T. (1999). *Las inteligencias múltiples en el aula*. Buenos Aires: Manantial.
- Bacon, J. (2015). The impact of standards-based reform on special education and the creation of the 'dividual'. *Critical Studies in Education*, 56 (3), 366-383.
- Barrenechea, I. (2010). Evaluaciones estandarizadas: Seis reflexiones críticas. *Archivos Analíticos de Políticas Educativas*, 18 (8), 3-27.

- Bayona, B. (2013). Los ejes de la LOMCE. *Fórum Aragón*, 7, 13-15.
- Bisquerra, R. (1998). *Modelos de orientación e intervención psicopedagógica*. Barcelona: Praxis.
- Bolívar, A., López, J. y Murillo, F. J. (2013). Liderazgo en las instituciones educativas. Una revisión de líneas de investigación. *Revista Fuentes*, 14, 15-60.
- Bourke, R. y Mentis, M. (2013). Self-assessment as a process for inclusion. *International Journal of Inclusive Education*, 17 (8), 854-867.
- Calero, M. D. y Padilla, J. L. (2004). Técnicas psicométricas: los tests. En R. Fernández-Ballesteros (Dir.), *Evaluación psicológica. Conceptos, métodos y estudio de casos*, Pirámide, Madrid.
- Coburn, C. E., Hill, H. C. y Spillane, J. P. (2016). Alignment and accountability in policy design and implementation: the common core state standards and implementation research. *Educational Researcher*, 45 (4), 243-251.
- Conley, D. (2015). A new era for educational assessment. *Education Policy Analysis Archives*, 23 (8). <http://dx.doi.org/10.14507/epaa.v23.1983>
- Darling-Hammond, L. y Falk, B. (2013). *Teacher learning through assessment. How student-performance assessments can support teacher learning*. Washington: Center for American Progress.
- Eurydice (2009). *National testing of pupils in Europe: objectives, organisation and use of results*. Brussels: Education, Audiovisual and Culture Executive Agency.
- Falabella, A. (2014). The performing school: the effects of market and accountability policies. *Education Policy Analysis Archives*, 22 (70). <http://dx.doi.org/10.14507/epaa.v22n70.2014>
- Feniger, Y., Israeli, M. y Yehuda, S. (2015). The power of numbers: the adoption and consequences of national low-stakes standardised tests in Israel. *Globalisation, Societies and Education*, 14 (2), 183-202.
- Fernández Navas, M., Alcaraz, N. y Sola, M. (2017). Evaluación y pruebas estandarizadas: una reflexión sobre el sentido, utilidad y efectos de estas pruebas en el campo educativo. *Revista Iberoamericana de Evaluación Educativa*, 10 (1), 51-67.
- Froemel, J. E. (2009). La efectividad y la eficiencia de las mediciones estandarizadas y de las evaluaciones en educación. *Revista Iberoamericana de Evaluación Educativa*, 2 (1), 10-28.

- Grañeras, M. y Parras, A. (2009). *Orientación educativa: fundamentos teóricos, modelos institucionales y nuevas perspectivas*. Centro de Investigación y Documentación Educativa, Ministerio de Educación.
- Haladyna, T., Bobbit-Nolen, S. y Haas, N. (1991). Rising standardized achievement test scores and the origins of test score pollution. *Educational Researcher*, 20 (5), 2-7.
- Hopfenback, T., Flórez, M. T. y Tolo, A. (2015). Balancing tensions in educational policy reforms: large-scale implementation of Assessment for Learning in Norway. *Assessment in Education: Principles, Policy and Practice*, 22 (1), 44-60.
- Klenowski, V. y Wyatt-Smith, C. (2012). The impact of high stakes testing: the Australian story. *Assessment in Education: Principles, Policy and Practice*, 19 (1), 65-79.
- Koch, M. J. y DeLuca, C. (2012). Rethinking validation in complex high-stakes assessment contexts. *Assessment in Education: Principles, Policy and Practice*, 19 (1), 99-116.
- Komatsu, H. y Rappleye, J. (2017). A new global policy regime founed on invalid statistics? Hanushek, Woessmann, PISA and economic growth. *Comparative Education*, 53 (2), 166-191. 10.1080/03050068.2017.1300008
- Ley 1/1990 de 3 de octubre, *Orgánica de Ordenación General del Sistema Educativo*, Ministerio de Educación.
- Ley 8/2013, de 9 de diciembre, *Orgánica para la Mejora de la Calidad Educativa*, Boletín Oficial del Estado.
- Maestro, C. (2006). La evaluación del sistema educativo. *Revista de Educación, extraordinario 2006*: 315-336.
- Monarca, H. (2012). La influencia de los sistemas nacionales de evaluación en el desarrollo del currículum. *Perfiles educativos*, 34 (135), 164-176.
- Monarca, H. (2015). Sistemas de evaluación: disputas simbólicas y configuración de prácticas educativas en los procesos de globalización y reforma del Estado. En H. Monarca (Coord.), *Evaluaciones externas. Mecanismos para la configuración de representaciones y prácticas en educación* (pp. 17-42). Buenos Aires, Argentina: Miño y Davilá.
- Monarca, H. y Fernández-Agüero, M. (2018). Opinión del profesorado sobre las fortalezas y debilidades de las evaluaciones externas de aprendizaje en España. *Educación XXI*, 21 (2), 249-273.

- Morales Vallejo, P. (2009). La evaluación formativa. En P. Morales Vallejo, *Ser profesor: una mirada al alumno*. Universidad Rafael Landívar, Guatemala.
- Pérez Gómez, A. I. (2014). Evaluación externa en la LOMCE. Reválidas, exclusión y competitividad. *Revista Interuniversitaria de Formación del Profesorado*, 81 (28.3), 59-71.
- Pettersson, D., Popkewitz, T. y Lindblad, S. (2016). On the use of educational numbers: comparative constructions of hierarchies by means of large-scale assessments. *Espacio, Tiempo y Educación*, 3 (1), 177-202.
- Polesel, J., Rice, S. y Dulfer, N. (2014). The impact of high-stakes testing on curriculum and pedagogy: a teacher perspective from Australia. *Journal of Education Policy*, 29 (5), 640-657.
- Puente, J. (2000). La evaluación del sistema educativo español. *Revista de Educación*, 321: 81-96.
- Rodríguez Romero, M. M. (1996). *El asesoramiento en educación*. Granada: Aljibe.
- Rodríguez Romero, M. M. (2006). El asesoramiento comunitario y la reinención del profesorado. *Revista de Educación*, 339, 59-75.
- Simmonds, M. y Webb, P. T. (2013). Accountability synopticism: how a think tank and the media developed a quasimarket for school choice in British Columbia. *The International Education Journal: Comparative Perspectives*, 12 (2), 21-41.
- Skedsmo, G. (2011). Formulation and realisation of evaluation policy: inconcistencies and problematic issues. *Educational Assessment, Evaluation and Accountability*, 23, 5-20.
- Smith, W. C. (2014). The global transformation toward testing for accountability. *Education Policy Analysis Archives*, 22 (116). <http://dx.doi.org/10.14507/epaa.v22.1571>
- Spillane, J. (2012). Data in practice: conceptualizing the data-based decisión-making phenomena. *American Journal of Education*, 118, 113-141.
- Vélez de Medrano, C. (2005). El asesoramiento en educación no formal: una mirada desde los orígenes de este modelo de intervención. En C. Monereo y J. I. Pozo (Coords). *La práctica del asesoramiento educativo a examen*. Barcelona: Graó.
- Vélez de Medrano, C. (Coord.) (2003). *Orientación comunitaria. El asesoramiento educativo para la resolución de problemas de los menores vulnerables o en conflicto social*. Madrid: UNED.

- Viñao, A. (2015). La Ley Orgánica de Mejora de la Calidad Educativa (LOMCE) de 2013: ¿una reforma más? *Historia y memoria de la Educación*, 3, 137-170.
- Winter, C. (2017). Curriculum policy reform in an era of technical accountability: “fixing” curriculum, teachers and students in English schools. *Journal of Curriculum Studies*, 49 (1), 55-74.

Fecha de recepción: junio de 2018

Fecha de aceptación: septiembre de 2018