

EVALUACIÓN EN GRAN ESCALA Y EVALUACIÓN EN AULA

Felipe Martínez Rizo

En su forma más conocida, la evaluación educativa no es algo reciente. La tarea del maestro, en su interacción cotidiana con los alumnos, ha incluido siempre, como una dimensión fundamental, el evaluar los avances de cada uno.

En las formas tradicionales de enseñanza que prevalecieron hasta bien entrado el siglo XIX, cuando surgieron los sistemas educativos de concepción moderna con los que estamos familiarizados —con cobertura que tendía a ser universal y, por ello, con muchos alumnos, organizados en grados de edad y avance similar— la tarea de los maestros era más de evaluación que de docencia. En las escuelitas en que un *dómine* atendía a una docena de chicos de distintas edades y niveles, para enseñarles a leer, escribir, contar y rezar, la lección magistral estaba ausente; el trabajo del maestro se limitaba fundamentalmente a *tomar la lección* a cada alumno, indicándole luego la siguiente tarea, en función de su avance.

La que sí es bastante reciente es *la evaluación en gran escala*, la que consiste en la aplicación estandarizada de pruebas a grandes números de alumnos, para apreciar el nivel de aprendizaje que se alcanza en el sistema educativo de todo un país, región o distrito, ante la imposibilidad de agregar las evaluaciones que hacen los maestros, siempre ligadas al contexto en que trabaja cada uno.

Con antecedentes que se remontan al final del XIX, las evaluaciones en gran escala se desarrollaron en los Estados Unidos, durante la primera mitad del siglo XX, adquirieron importancia a lo largo de su segunda mitad y se extendieron a la mayor parte de los países del mundo en las dos o tres últimas décadas. Además de evaluaciones nacionales, se desarrollaron proyectos internacionales que hoy atraen poderosamente la atención cada vez que se difunden sus resultados. Las más conocidas son las pruebas del Proyecto para la Evaluación Internacional de Estudiantes (PISA, por sus siglas en inglés), de la Organización para la Cooperación

y el Desarrollo Económico (OCDE), cuyo desarrollo comenzó a planearse en 1995, y se aplican cada tres años desde 2000.

Más de tres décadas antes, en 1958, comenzaba a gestarse la Asociación Internacional de Evaluación del Rendimiento Académico (IEA), con la planeación del Primer Estudio Internacional sobre Matemáticas que se llevó a cabo en la década de 1960, del que se deriva el Estudio de Tendencias en Matemáticas y Ciencias, TIMSS, una de las evaluaciones vigentes más importantes.

En México, el desarrollo de pruebas en gran escala para educación básica dio inicio desde la década de 1970, y se desarrolló sobre todo a partir de la de 1990, con las pruebas para evaluar el Factor Aprovechamiento Escolar del Programa de Carrera Magisterial. La tendencia se acentuó en la última década, con las pruebas del Instituto Nacional para la Evaluación de la Educación (INEE), a partir de 2003, y las pruebas censales de la Secretaría de Educación Pública, desde 2006.

Las pruebas en gran escala pueden ser una herramienta valiosa para apoyar los esfuerzos de mejora de la calidad educativa, si se las ve como un complemento de las evaluaciones a cargo de los maestros, insustituibles para evaluar de manera detallada todos los aspectos del currículo y para hacerlo de manera que puedan ofrecer a cada niño y cada niña la retroalimentación precisa sobre sus puntos fuertes y débiles, esencial para que mejorar el aprendizaje de cada uno.

Es preciso, sin embargo, advertir sobre un riesgo que no se puede ignorar: el peligro de que las pruebas en gran escala produzcan consecuencias negativas para la calidad educativa, si se les comprende y utiliza mal.

En efecto: para poder dar resultados confiables de los niveles de aprendizaje que alcanzan muchos miles de alumnos, una prueba estandarizada necesariamente tiene que reducirse a la medición de un número relativamente pequeño de temas, y debe hacerlo

mediante preguntas que no pueden atender los aspectos más complejos de las competencias que pretende desarrollar la escuela. Por ello hay que reiterar la idea central de que las pruebas en gran escala no pueden sustituir el trabajo de evaluación de los maestros, el único que puede atender con precisión los aspectos más complejos de la enseñanza y el aprendizaje, y hacerlo de modo que se pueda brindar retroalimentación detallada y oportuna a cada estudiante.

Si no se entienden bien los alcances y límites de los resultados de las pruebas en gran escala, es fácil que se usen en forma inapropiada. El peligro más claro en este sentido es la tendencia a tomar como referente para la tarea docente el contenido de las pruebas y no el de los programas de estudio, enseñar para las pruebas, por la visibilidad de los resultados de éstas. Con lo anterior la tarea de la escuela se empobrece, al descuidar aspectos esenciales que no evalúan las pruebas en gran escala, como la expresión escrita y oral, la formación de actitudes y valores, la educación artística, e incluso los niveles cognitivos más complejos de las áreas tradicionalmente cubiertas de lectura, matemáticas y ciencias.

Otros ejemplos de manejo inapropiado de los resultados de las evaluaciones en gran escala, son la asignación de estímulos a los docentes o la elaboración de ordenamientos simples de escuelas, en ambos casos supuestamente en función de la calidad de los maestros o las escuelas, sin tener en cuenta los numerosos factores que inciden en los resultados de los alumnos en las pruebas ni tener en cuenta las limitaciones de éstas.

Este tipo de errores, además, produce un explicable rechazo de toda evaluación en gran escala por parte de muchos maestros, que se dan cuenta de sus graves consecuencias para una educación realmente buena.

Por lo anterior, es importante reiterar que la evaluación en gran escala puede ser muy valiosa para la mejora de la calidad, a condición de entenderla y usarla de otra manera, en particular viéndola como complemento del trabajo del maestro, y no como sustituto del mismo. Ahora bien: para que esto ocurra, es necesario que los maestros tengan una comprensión correcta tanto de

los alcances y límites de la evaluación en gran escala, como de la que ellos deben llevar cabo en el aula.

En México la formación inicial que reciben muchos maestros no los prepara bien ni para una cosa ni para la otra. Un indicio de lo anterior es una solicitud que el INEE suele recibir, de maestros, supervisores y directivos de escuelas normales, para impartir talleres de elaboración de reactivos de opción múltiple, gracias a lo cual se espera que mejore la calidad de las evaluaciones que deben hacer los maestros.

En el INEE hemos mantenido la posición anterior sobre los alcances y límites de la evaluación en gran escala, y sobre la necesidad de verla como complementaria de la evaluación a cargo de los maestros, y nos enfrentamos permanentemente con una dificultad considerable para conseguir que esta postura sea comprendida no sólo por la sociedad en general, sino también por las autoridades educativas y por los maestros, como muestran las demandas a las que alude el párrafo anterior.

Tratando de enfrentar esa dificultad, el INEE desarrolló la herramienta informática llamada *Explorador de Excale*, dirigida en especial a los maestros de educación básica, buscando propiciar un mejor conocimiento de las pruebas desarrolladas por el Instituto (*Excale*) y, a la vez, su utilización en el contexto del aula, para fortalecer la calidad de las evaluaciones hechas por los maestros mismos. Con el mismo propósito el INEE, junto con la Subsecretaría de Educación Básica, la Universidad Pedagógica Nacional y la Fundación para la Cultura del Maestro, lanzó una convocatoria para invitar a los maestros del país a utilizar el *Explorador*.

De los trabajos recibidos como respuesta a esa convocatoria, algunos de los que han sido considerados merecedores de reconocimiento por su calidad, tienen que ver con la relación entre evaluación en gran escala y evaluación en aula, y son los que se agrupan en este bloque.

Las propuestas que se presentan a continuación son interesantes pero, al mismo tiempo, muestran la tendencia mencionada a considerar la evaluación en gran escala como el paradigma al que debería

subordinarse la evaluación que llevan a cabo los maestros en el salón de clases. Eso se puede apreciar en el trabajo que abre el bloque, “Evaluamos mejor para que los alumnos aprendan más” de Gloria del Carmen Alfaro Cuevas, cuya propuesta para una zona escolar se refiere al desarrollo de la capacidad de los maestros para elaborar reactivos. Una idea similar se aprecia en el segundo trabajo “Apropiación de los reactivos de Excale” de Abril Mayanín Vázquez Buenfil.

Los dos textos siguientes proponen el uso de portafolios, en el caso de “Planeación del trabajo escolar a nivel de aula dentro de la escuela primaria” de Reyna Lizeth Caballero Botello y Perla Yadira Lara Mendoza, su uso es para mejorar la planeación del trabajo escolar, y en el caso de “Excale un referente para evaluar la enseñanza a través del portafolio docente” de Sandra Conzuelo Serrato, se propone el portafolios para la evaluación de la docencia misma. El quinto trabajo “Identificación y atención de necesidades de actualización docente, a nivel escuela o zona” de Madain Vázquez García y Fernando Méndez Hernández, señala también la posibilidad de identificar las necesidades de actualización de los maestros usando el Explorador, pero con el propósito de desarrollar la cultura de la evaluación.

Sin desconocer el valor de los aspectos atendidos en las cinco propuestas que se recogen en esta publicación, es importante ver que es necesario ir más allá, buscando mejoras más profundas en la comprensión y el uso de la evaluación.

Si se comprenden las características de las evaluaciones en gran escala y en aula, se entenderá también que las preguntas de opción múltiple son esenciales en las primeras, pero que las segundas, las evaluaciones que los maestros llevan a cabo en el aula, pueden utilizar acercamientos diferentes y mejores para evaluar aspectos finos y complejos, que difícilmente se pueden atender en gran escala, pero que en el ámbito del aula es posible emplear. Las preguntas de opción múltiple pueden ser usadas también en el aula, y son adecuadas para evaluar algunos aspectos del aprendizaje, pero otros deben valorarse de formas distintas, como mediante la producción de textos amplios, la realización de ejercicios en vivo, la observación del trabajo individual y grupal de los alumnos, entre otros.

La expresión *evaluación en aula* (*classroom assessment*) se refiere a este tipo de acercamientos alternativos. Es importante que escuelas normales y programas de actualización de maestros en servicio presten la atención que merecen a estos enfoques, relativamente recientes y poco conocidos en nuestro medio.

Además de la distinción referente al ámbito en que tiene lugar una evaluación y a sus dimensiones —*en gran escala o en aula, macro o micro*— para clasificar los distintos acercamientos a la misma es importante tener en cuenta también sus propósitos. En este sentido es importante la tradicional distinción entre evaluación *formativa* y *sumativa* o, en los términos preferidos por autores de las nuevas corrientes: evaluación *del* aprendizaje vs. evaluación *para* el aprendizaje.

Las nuevas tendencias de *evaluación en aula* son congruentes con los enfoques pedagógicos llamados *constructivistas*, puesto que unas y otros se derivan de la *revolución cognitiva*. La idea clave común a las dos corrientes establece que el aprendizaje no consiste simplemente en la acumulación de información, sino en la construcción de nuevos conocimientos, integrando datos nuevos en estructuras cognitivas previamente existentes que, en algunos casos, deben redefinirse en una dialéctica de deconstrucción-construcción.

A partir del punto central anterior, se llega a la idea de que el aprendizaje de los alumnos no se dará sencillamente por el hecho de que el maestro exponga un tema, sino que es necesario un trabajo mucho más complejo para que el proceso tenga lugar, en especial cuando se trata de competencias cognitivas complejas.

En ese caso es fundamental que se distingan con el mayor detalle posible los componentes de esas competencias complejas, y que se trabaje de tal suerte que los alumnos identifiquen con precisión los aspectos que manejan bien y aquellos en los que no lo consiguen, como condición necesaria para orientar sus esfuerzos para avanzar. En otras palabras, para que los alumnos aprendan es necesario un tipo de evaluación que permita detectar los puntos débiles de cada uno, seguida por una retroalimentación inmediata para que sea posible orientar con precisión los esfuerzos siguientes. Eso es lo que ha dado en llamarse

evaluación para el aprendizaje, y sólo puede hacerla adecuadamente un buen maestro, en pequeña escala: *evaluación en aula*.

Para contribuir al desarrollo de estas innovadoras ideas, el INEE ha publicado un texto importante de una de sus principales defensorasⁱ de cuya presentación se toma buena parte de las ideas anteriores. El texto contiene una bibliografía que podrán aprovechar quienes deseen ampliar sus conocimientos al respecto. Se recomienda en especial la obra de Stiggins y colaboradoresⁱⁱ.

Ojalá que la evaluación educativa en México se desarrolle combinando el avance técnico de las pruebas en gran escala, un uso de sus resultados que no ignore sus límites y un avance substancial de la evaluación en aula, a cargo de los maestros. Así, y sólo así, la evaluación podrá contribuir realmente a la mejora educativa.

Felipe Martínez Rizo
Julio de 2008.

BIBLIOGRAFÍA

ⁱ Shepard, Lorrie A. (2008). *Evaluación en el aula*. México. INEE. Versión original en Brennan, Robert L. Ed. (2006). *Educational Measurement*. 4a ed. Westport. ACE/Praeger. Cap. 17. Págs. 623-646.

ⁱⁱ Stiggins, Rick, J. Arter, J. Chappuis y S. Chappuis (2007). *Classroom Assessment for Student Learning. Doing It Right – Using It Well*. Merrill/ETS College Textbook Series. Upper Saddle River, NJ. Pearson Education.

También se recomienda consultar:

http://www.assessmentinst.com/staff_rick.php