

Repensando el entusiasmo evaluador y las pruebas

Rosa María Torres

1. La situación en general

Siguiendo recomendaciones internacionales y "siguiendo la corriente" al creciente peso asignado en los últimos años a la **evaluación** en el campo educativo, muchos países de América Latina y el Caribe han venido montando sistemas nacionales de evaluación. El auge de las **pruebas estandarizadas**, por su parte, es parte del actual Movimiento Global de Reforma Educativa (GERM por su nombre en inglés: Global Education Reform Movement). Todo el paquete ha sido importado acríticamente en muchos países latinoamericanos, justo en momentos en que dichas [pruebas estandarizadas son debatidas, cuestionadas y rechazadas](#) en muchos países del Norte, empezando con Estados Unidos, que es donde ganaron notoriedad y desde donde se han expandido al resto del mundo.

Además de las iniciativas y sistemas nacionales, existen dos pruebas internacionales de rendimiento escolar en las que han participado varios países de la región, una de alcance regional y otra de alcance mundial:

(a) **el Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE)** coordinado por la oficina regional de la UNESCO ([OREALC](#)). El LLECE ha realizado hasta la fecha dos estudios, en [1997](#) y en [2005-2006](#), los cuales han evaluado dos áreas del currículo: Lenguaje y Matemáticas, en tercero y cuarto grados, y en tercero y sexto grados, respectivamente. En el segundo estudio ([SERCE](#)) se agregó Ciencias Naturales para algunos países. Un tercer estudio - Tercer Estudio Regional Explicativo y Comparativo ([TERCE](#)) - aplicado en el 2013 y dará a conocer sus resultados en 2014.

En el primer estudio participaron 13 países. En el segundo estudio participaron 16 países - Argentina, Brasil, Colombia, Costa Rica, Cuba, Chile, Ecuador, El Salvador, Guatemala, México, Nicaragua, Panamá, Paraguay, Perú, República Dominicana, Uruguay - y el estado de Nuevo León en México. Cuba ocupó el primer lugar en los resultados de las pruebas y República Dominicana el último. Ecuador, El Salvador, Nicaragua, Guatemala, Panamá, Paraguay, Perú y Rep. Dominicana conformaron el grupo de países con peores resultados.

(b) **las pruebas internacionales PISA** (*Programme for International Student Assessment* - Programa para la Evaluación Internacional de Alumnos) diseñadas por la [OCDE](#) (los países "desarrollados") que vienen aplicándose desde el año 2000 y cada tres años a estudiantes de 15 años de edad. La prueba del 2000 se centró en competencias lectoras, la del 2003 en competencias matemáticas, la del 2006 en competencias científicas y la del 2009 nuevamente en competencias lectoras (incluyendo por primera vez lectura en medios digitales). Varios países latinoamericanos han participado hasta ahora en la prueba PISA: Argentina, Brasil, Colombia, Chile, México, Panamá, Perú, Uruguay. Estos han conformado – junto con España y Portugal, en Europa - el [Grupo Iberoamericano de PISA](#). México y Chile son los dos únicos países latinoamericanos incluidos dentro del grupo de países OCDE.

2. Evaluando los impactos de las evaluaciones

La evaluación, en educación, se vende como instrumento indispensable e infalible para "mejorar la calidad de la educación". No obstante, la fiebre evaluadora – de los alumnos de diversos niveles, de los docentes, de las instituciones escolares (y ahora, ¡incluso de los padres de familia!, según se propone en EE.UU...) – que viene creciendo desde la década de 1990, no se ha traducido en la anunciada "mejoría". Y es que **la evaluación, por sí misma, no mejora ni cambia nada**. Más bien se observan varios efectos negativos de las evaluaciones estandarizadas, así como debilidades técnicas y problemas de comparabilidad entre ellas, a nivel regional e internacional, lo que dificulta establecer con claridad conclusiones y tendencias. Por ejemplo, los resultados del

primer y segundo estudios del LLECE no son comparables entre sí, entre otros porque: en 1997 se evaluó a alumnos de 3º y 4º grados, y en 2005-2006 a alumnos de 3º y 6º grados; el segundo estudio evaluó no sólo saberes sino “habilidades para la vida”; y 8 de los países que participaron en el segundo estudio no participaron en el primero. A su vez, los métodos, indicadores y resultados del LLECE (regional, UNESCO) y de PISA (mundial, OCDE) no son plenamente comparables entre sí.

Todo esto obliga a revisar el camino recorrido y algunas de las premisas que sustentan el entusiasmo evaluador, entre otras:

- a. que las evaluaciones miden lo que saben o son capaces de hacer los alumnos y los docentes, y
- b. que las evaluaciones son indispensables para mejorar la calidad de la educación, en tanto ayudan a visibilizar los resultados de la enseñanza (y a explicar algunos de los factores que contribuyen a las diferencias de sus resultados) y, de ese modo, a informar mejor - y eventualmente rectificar - políticas y acciones educativas.

La experiencia muestra que, a pesar de los recursos invertidos en reforma y en evaluación educativas desde los 1990s, no hay mejoría consistente y/o significativa en los resultados escolares de los países en los campos o aspectos evaluados. Se observan pequeños incrementos en los puntajes de las pruebas, pero también bajones. En cualquier caso, los países latinoamericanos siguen ocupando los últimos lugares de la lista de países participantes en PISA.

La divulgación de los resultados en forma de *ránkings* que comparan entre sí a los planteles escolares y a los países, no ha tenido los impactos esperados sobre las políticas ni sobre las prácticas docentes. Más bien, ha contribuido a reforzar el malestar y la resistencia docentes, pues son los docentes los primeros a quienes se responsabiliza de los malos resultados, sin ver el conjunto de [“factores asociados” a los rendimientos escolares](#) y sin que se implementen las políticas que las propias evaluaciones nacionales e internacionales muestran indispensables. [Los gobiernos de los países participantes en las sucesivas pruebas PISA no están teniendo en cuenta las conclusiones y recomendaciones de dichas pruebas.](#)

Asimismo, las evaluaciones de las evaluaciones muestran que se avanza poco en términos de construir una **cultura de evaluación** en el campo educativo. Más bien, vienen desarrollándose varias **tendencias negativas** previsibles:

- ▶ **desmoralización** entre estudiantes, docentes y planteles que obtienen bajos resultados, lo que contribuye no a su superación sino más bien a reforzar sentimientos de minusvalía e impotencia.
- ▶ **“estudiar para la prueba”** (tanto en el caso de los alumnos como de los docentes), lo que no incide realmente sobre la comprensión y el aprendizaje sino más bien sobre la memorización y la retención de información en el corto plazo.
- ▶ **atención a las asignaturas o contenidos sujetos a evaluación**, y descuido de aquellos que se sabe no entrarán en las pruebas.
- ▶ **toda la sociedad concentrada en las pruebas y en sus resultados**, asumiendo estos como indicador de la calidad de la educación escolar a nivel nacional y de cada plantel, olvidando aspectos tanto o más importantes que no se ven a través de las pruebas tales como: la relevancia y pertinencia del currículo y de los contenidos que se enseñan en las aulas, la validez pedagógica de los métodos de enseñanza utilizados, la calidad profesional y humana de los educadores, las relaciones entre profesores y alumnos y entre alumnos, el clima de aula y el clima escolar en general, el interés y la satisfacción de los alumnos y de las familias.
- ▶ **presionar a los profesores a que sus alumnos obtengan buenas calificaciones en las pruebas**, pues de ellas depende la evaluación del “desempeño docente”, aumentos salariales,

incentivos, premios y castigos, y hasta despidos (En EEUU y en otros países muchos profesores dependen hoy de las notas de sus alumnos para sobrevivir y permanecer en el sistema).

► **activar la previsible "cultura del engaño"**, incluyendo filtración y compra-venta de pruebas, venta de cursos y materiales para prepararse y aprobarlas, evitar que se presenten a las pruebas los "malos alumnos", los "migrantes", los con "problemas de aprendizaje", etc. La selección de alumnos al momento de la prueba es un fenómeno que ha sido analizado y cuestionado desde hace mucho en [Chile](#). Casos de engaño denunciados y ventilados públicamente se han dado por ejemplo en [México](#) ([copia en la prueba nacional Enlace, venta del 'Examen Universal' - aplicado en julio 2012 - a través de Mercado Libre, en internet](#)), [Perú](#) (pruebas aplicadas a los docentes), Brasil ([ENEM](#): Exame Nacional do Ensino Médio), Estados Unidos ([Atlanta](#), pruebas estandarizadas en escuelas) y en la comunidad autónoma de [Cataluña](#) en España (pruebas PISA).

► **obtener buenos resultados en las pruebas como objetivo**, a fin de:

- quedar bien parado en el marco de parámetros competitivos según los cuales son evaluadas las instituciones escolares, e incluso juzgadas como aptas o con mérito para recibir diversos tipos de incentivos,
- "ubicarse lo mejor posible" en los rankings internacionales.

Mejorar los puntajes en la prueba PISA y el ranking del país en dichas pruebas ha pasado a ser la razón de ser y el objetivo principal de las reformas educativas en varios países (por ejemplo [México](#) de cara a PISA 2012), así como un logro esperable de metas de alcance regional (por ejemplo, las [Metas 2021](#) promovidas por la [OEI](#). Meta 5: Mejorar la calidad de la educación. Indicador 13: Porcentaje de alumnos con niveles satisfactorios de logro en competencias básicas en las pruebas nacionales e internacionales. Nivel de logro: Disminuye en al menos un 20% el número de alumnos situados entre los dos niveles bajos de rendimiento en las pruebas LLECE 6º grado, PISA, [TIMMS](#) o [PIRLS](#) en las que participan diferentes países. Aumentan en la misma proporción los alumnos en los dos niveles altos en dichas pruebas).

Hay asimismo puntualizaciones importantes que hacer y que a menudo siguen sin hacerse:

► *evaluar la **calidad de la educación** es mucho más amplio que evaluar los **rendimientos escolares**.*

► ***rendimiento escolar** no es lo mismo que **aprendizaje**; los aprendizajes que se dan en el sistema escolar van mucho más allá de la asimilación de determinadas asignaturas e incluso del currículo prescrito.*

► las pruebas pueden medir solo una pequeña porción de lo aprendido/sabido, **dejando afuera valores, aptitudes, actitudes y competencias claves que hacen al corazón mismo de la educación, la socialización y la convivencia en el medio escolar, y de las múltiples inteligencias** (por ejemplo: pensamiento crítico, colaboración, iniciativa, esfuerzo, curiosidad, respeto, amabilidad, honestidad, humor, etc.).

► **no cabe aplicar los mismos instrumentos de evaluación a países, zonas, grupos humanos, culturas, diferentes.** A objetivos, condiciones, contextos, contenidos diversos, corresponden obviamente diferentes para qué, qué y cómo en términos de evaluación.

3. América Latina y los países de la OCDE

Es fundamental preguntarse acerca de la pertinencia y validez de aplicar las mismas pruebas - y posteriormente comparar sus resultados - (a) a toda la población escolar, muy heterogénea dentro de un mismo país o región, y (b) a países muy diferentes entre sí, como son los países de la OCDE - países ricos con necesidades básicas satisfechas por parte de la mayoría de la población

(y que son los países que sistemáticamente obtienen los mejores resultados en las pruebas PISA) - y los países de América Latina, con una alta proporción de su población viviendo bajo la línea de pobreza. Como corroboran las mismas evaluaciones (LLECE, PISA), [el contexto socio-económico - no sólo de cada escuela sino del país en su conjunto - tiene gran peso sobre las condiciones de enseñanza-aprendizaje y sobre los resultados escolares](#) en los diversos niveles del sistema.

Las enormes diferencias – históricas, sociales, económicas, políticas culturales - entre los países del Norte y los países del Sur, dentro de la propia América Latina y en el interior de cada país, exigen enfoques, políticas y programas diferenciados. Esto se aplica también a la evaluación, la cual incide de manera directa sobre los currículos, las pedagogías, las expectativas, los perfiles y quehaceres docentes.

Antes que herramienta de rectificación de desigualdades, las pruebas y la evaluación en el medio escolar puede reforzar la discriminación hacia los sectores más pobres y vulnerables. En particular, los grupos y pueblos indígenas resultan los más perjudicados pues las pruebas se elaboran por lo general sin atención a sus especificidades lingüísticas y culturales. [Esto ha sido oficialmente denunciado, por ejemplo, en el caso de México y su prueba nacional Enlace.](#)

4. La experiencia cubana y la experiencia finlandesa

Dos notas de cierre que llaman a la reflexión:

► [Cuba es el país que ha obtenido los mejores resultados escolares en ambas pruebas LLECE](#) aplicadas hasta hoy (1997 y 2005-2006), en ambos casos con una gran diferencia respecto de los demás países que han participado en estas pruebas en América Latina y el Caribe (cabe recordar que las pruebas LLECE incluyen tanto a escuelas públicas como privadas). Cuba, no obstante, está insatisfecha con su sistema educativo y reconoce la necesidad de cambios mayores en sus políticas educativa, cultural y social. Una muestra de que los resultados de las pruebas escolares son insuficientes para dar cuenta de la calidad educativa y que los *rankings* formalmente establecidos e internacionalmente validados pueden tener escasa relevancia frente a las expectativas y a los parámetros que pueden fijarse internamente los propios países.

► [Finlandia](#), país que viene ocupando los primeros lugares en las pruebas PISA a nivel mundial, se distancia radicalmente de los preceptos que dominan la evaluación en la "reforma educativa global" en la actualidad;

- es crítica de las pruebas estandarizadas y de los *rankings*;
- no aplica evaluaciones estandarizadas a sus alumnos antes de los 16 años;
- no divulga públicamente los resultados de las pruebas y no promueve la competencia entre instituciones escolares en base a dichos resultados;
- no aplica a sus profesores esquemas de "pago por desempeño" (entendiendo y valorando dicho desempeño en relación a los rendimientos escolares, según lo medido por pruebas).

** Texto revisado y actualizado (2012), basado en la conferencia "Sociedad del conocimiento, políticas de evaluación y cambio educativo", conferencia de clausura del X Congreso de [ADIDE-Federación](#), Zaragoza, España, 5-7 noviembre 2008.*