

CONTEXTOS VULNERABLES: LAS APORTACIONES DE LA EVALUACIÓN

Vulnerable contexts: the contribution of assessment

FELIPE MARTÍNEZ RIZO

Universidad Autónoma de Aguascalientes

En diversas modalidades, durante las últimas décadas la evaluación educativa se ha desarrollado con fuerza en muchos sistemas educativos y ha adquirido un papel destacado en la política educativa de buen número de países. Si se diseña e implementa bien, y si los resultados se analizan y usan en forma apropiada, la evaluación puede sustentar decisiones que lleven a mejorar la calidad educativa, pero si esas condiciones no se cumplen puede tener resultados contraproducentes que se concreten en un deterioro de dicha calidad. Tanto las posibilidades positivas como las negativas de la evaluación tienen matices particulares tratándose de alumnos y escuelas que se sitúen en contextos socialmente vulnerables. Este artículo presenta consideraciones sobre el particular, especialmente a partir de la experiencia de México. Después de precisar la noción de contextos vulnerables (como expresión de la escasez de elementos fundamentales para el bienestar familiar y el desarrollo escolar) en relación con las desigualdades que caracterizan a la sociedad mexicana y su sistema educativo, se presenta la noción de calidad aplicada a un sistema educativo, se describen algunos tipos de evaluación y se analizan las implicaciones de las pruebas de rendimiento en gran escala y de la difusión de sus resultados en la forma de ordenamientos.

Palabras clave: *Contexto desfavorecido, Desigualdad social, Diversidad lingüística, Igualdad de oportunidades, Calidad de la educación, Evaluación, Resultado de aprendizaje, Prueba estandarizada, Evaluación del sistema educativo, Rankings.*

Qué son los contextos vulnerables

Es bien sabido que el desempeño escolar de los alumnos depende en parte de la calidad de la enseñanza que reciban en la escuela, pero en parte también de la presencia en el hogar de ciertas condiciones favorables, desde lo necesario para que estén bien nutridos y sanos, hasta el apoyo que reciban de sus padres, pasando por la existencia de libros, un lugar de trabajo, juguetes educativos, entre otros elementos.

Referida al contexto familiar, la noción de vulnerabilidad denota justamente las situaciones en que esos elementos son escasos o incluso están ausentes en ciertos hogares. En el caso del contexto escolar, se refiere a la escasez o ausencia en la escuela de los elementos que se consideran necesarios para que haya una buena enseñanza, comenzando por un docente cualificado, además de libros, material didáctico, etc.

Los sistemas educativos y las sociedades difieren considerablemente en cuanto a la forma en que se

distribuyen los elementos citados en hogares y escuelas. Los países escandinavos destacan porque tienen sociedades con poca desigualdad y también escuelas notablemente homogéneas. Las naciones de América Latina se sitúan en el caso opuesto, con desigualdad muy pronunciada tanto en la sociedad como en la escuela (INEE, 2007).

La calidad educativa

En el campo educativo la noción de evaluación puede referirse a la que llevan a cabo los maestros en relación con el progreso de sus alumnos, pero también a la evaluación de maestros, de escuelas, del currículo u otros objetos. La evaluación de la calidad de los sistemas educativos como tales es un caso particular que ha cobrado gran importancia en las últimas décadas.

Aplicada a un sistema educativo la noción de calidad no puede reducirse al aprendizaje que alcanzan los alumnos, sino que comprende otros aspectos: no se podría considerar bueno un sistema que atendiera solo a una parte de los niños en edad escolar, aunque los alumnos atendidos alcanzaran buenos niveles de aprendizaje. Por ello es razonable postular que, para ser bueno, un sistema educativo deberá al menos:

- Tener un currículo que responda a las necesidades de la sociedad y tenga en cuenta las condiciones de los alumnos.
- Lograr que todo niño en edad escolar acceda a la escuela y recorra el trayecto en los tiempos establecidos, terminándolo a la edad estipulada habiendo alcanzado niveles aceptables de competencia.
- Conseguir que los aprendizajes alcanzados se manifiesten en la vida adulta en la forma de conductas que lleven a una calidad de vida superior.
- Tener recursos suficientes para cumplir sus funciones y usarlos en forma eficiente.
- Atender diferencialmente a alumnos y escuelas, ofreciendo más apoyo a los que más lo necesiten.

En términos abstractos lo anterior se puede expresar diciendo que la calidad de un sistema educativo es una noción multidimensional que debe integrar los aspectos de relevancia, pertinencia, eficacia, impacto, eficiencia y equidad (Martínez Rizo *et al.*, 2007).

Las dimensiones de la calidad y los tipos de evaluación educativa

La noción de evaluación se relaciona estrechamente con la de calidad. Cuando se habla de evaluar cualquier cosa se entiende siempre que se trata de evaluar su calidad. Y si la noción de calidad es multidimensional, la de evaluación también deberá ser polifacética.

En otras palabras, y en referencia concretamente a la calidad de los sistemas educativos, la evaluación de cada dimensión de la calidad requiere de cierto tipo de acercamiento: las pruebas de rendimiento son adecuadas para evaluar el nivel de aprendizaje de los alumnos, pero otras dimensiones implican otros acercamientos, como indicadores estadísticos para valorar la cobertura y eficiencia del sistema, o estudios especiales para estimar la relevancia y la pertinencia del currículo o la equidad del sistema.

Desigualdad social y escolar en México

Numerosos estudios han mostrado que hay una estrecha relación entre las condiciones del hogar y el entorno social en que viven los alumnos y su desempeño escolar. Por ello se considera esperable que el rendimiento de los alumnos de familias con padres de menor escolaridad y, en general, inferior nivel socioeconómico y cultural (NSEC), sea también más bajo que el de los chicos con NSEC medio o alto. Por otra parte, se suele también esperar de la escuela —y, en particular, de la escuela pública gratuita— que compense de alguna manera las desigualdades sociales, permitiendo a los hijos de

las clases trabajadoras acceder a posiciones sociales mejores, gracias a los estudios.

Lo anterior supone, al menos, que todos los alumnos sean atendidos en escuelas que les brinden una educación de la misma calidad; de hecho debería esperarse incluso que las escuelas a las que asisten los niños de medios más vulnerables les brindaran un apoyo mayor, dado que las desfavorables condiciones del hogar los ponen en desventaja.

En muchos casos, sin embargo, lo que ocurre en realidad es que las escuelas que atienden a sectores particularmente desprotegidos cuentan también con recursos inferiores a los de las escuelas que operan en mejores condiciones de manera que, en lugar de contrarrestar las desigualdades sociales, la escuela de alguna manera contribuye a aumentarlas.

En el caso de México, estudios del Instituto Nacional para la Evaluación de la Educación (INEE) sustentan la aseveración anterior, con datos basados en una muestra representativa de las 98.575 escuelas primarias que tenía el país en el ciclo 2008-2009, con 14.815.735 alumnos en seis grados, y que se clasifican en cinco grupos, como sigue:

- *Cursos comunitarios*. Primarias multigrado en que un instructor, que no es maestro profesional, sino que tiene solo secundaria básica o superior, atiende a alumnos del primero al sexto grado; 11.187 escuelas (11,3% del total), en pueblos con menos de 100 habitantes, que atienden al 0,7% de la matrícula (109.141 niños).
- *Primarias indígenas*. Un conjunto de 9.918 escuelas primarias (10,1% del total) que atienden a 841.151 niños (5,7% del alumnado), con docentes profesionales por lo general bilingües, en zonas rurales con población cuya lengua materna no es el español.
- *Primarias públicas rurales*. Se trata de 37.886 escuelas (38,4% del total) ubicadas

en pueblos de hasta 2.500 habitantes, que atienden a 3.045.484 niños (20,6%).

- *Primarias públicas urbanas*. 31.742 escuelas (32,2%), localizadas en poblaciones de 2.500 habitantes o más, que atienden a 9.591.107 alumnos (64,7% de la matrícula), en general de clase baja y media baja.
- *Primarias privadas*. Son 7.842 escuelas (8,0% de los planteles), casi todas en medio urbano, que atienden a 1.228.852 alumnos (8,3% de la matrícula), en su gran mayoría de clase media alta y alta (INEE, 2009).

Puede hipotetizarse que el entorno familiar y social de los alumnos atendidos por las escuelas de los primeros de estos cinco tipos se caracterizará por una mayor vulnerabilidad, y que esta será menor en los tipos siguientes, lo que confirman claramente los datos empíricos recabados en las evaluaciones del INEE.

La tabla 1 presenta la distribución de los alumnos del sistema educativo mexicano en los cinco tipos anteriores de escuelas primarias, según su nivel socioeconómico y cultural, tomando como indicador la escolaridad de la madre.

Las diferencias del nivel socioeconómico y cultural del alumnado que asiste a cada tipo de escuela son evidentes. En las escuelas primarias privadas las madres de más de las dos terceras partes de los alumnos cuentan con estudios universitarios, menos de cuatro de cada cien tienen solo primaria, y menos de una de cada cien no fue a la escuela. En los cursos comunitarios y las primarias indígenas, por el contrario, más de dos terceras partes de las madres no fueron a la escuela o solo alcanzaron la primaria, mientras que poquísimas llegaron al bachillerato o a la universidad. El contraste entre las escuelas privadas, en un extremo, y estos dos tipos en el otro, es clarísimo, con las primarias públicas rurales y, sobre todo, las urbanas, en una posición intermedia.

TABLA 1. Porcentaje de alumnos según la escolaridad de su madre, por tipo de escuela

Escolaridad de las madres	Tipos de primaria				
	Cursos comunitarios	Indígenas	Públicas rurales	Públicas urbanas	Privadas
No fue a escuela	23,8	22,2	8,4	4,2	0,8
Primaria	53,5	42,3	43,3	18,7	3,9
Secundaria	17,1	19,9	27,6	27,6	12,0
Bachillerato	2,8	7,0	8,0	14,3	14,7
Universidad	2,8	8,7	12,7	35,2	68,7
Total	100,0	100,0	100,0	100,0	100,0

Fuente: INEE 2007: Tabla 1.1, p. 18.

Pero los datos derivados de los estudios del INEE muestran que las diferencias entre los tipos de primaria no se limitan a la extracción de sus alumnos, sino que se presentan en una forma muy similar en lo que se refiere a los recursos con que cuentan. La tabla 2 presenta datos sobre la calidad de la infraestructura de las escuelas, que se clasificaron en cuatro grupos: las escuelas del grupo I tienen condiciones de infraestructura muy deficientes; las del grupo II deficientes; las del III suficientes; y las del IV muy adecuadas.

Casi el 90% de las primarias privadas comprendidas en el estudio tuvieron condiciones de infraestructura muy adecuadas, y solo tres

de cada cien fueron ubicadas en las categorías de infraestructura deficiente o muy deficiente. En los cursos comunitarios la situación es la opuesta: más del 90% tienen infraestructura muy deficiente y deficiente, y menos de 10% suficiente o muy adecuada.

Es evidente el paralelismo de la distribución del nivel socioeconómico y cultural de los alumnos y la de la infraestructura de las escuelas a las que asisten. A diferencia de lo que debería ocurrir, las escuelas primarias mexicanas no compensan las desventajas del contexto de los hogares más pobres, sino que más bien las refuerzan, ya que los alumnos de circunstancias más vulnerables en el hogar son atendidos en escuelas de mayores carencias.

TABLA 2. Porcentaje de alumnos en escuelas con diferente calidad de infraestructura, por tipo

Calidad de la infraestructura escolar	Tipo de primarias				
	Cursos comunitarios	Indígenas	Rurales	Urbanas públicas	Privadas
Muy deficiente (I)	75,2	59,7	32,3	22,0	1,6
Deficiente (II)	15,7	23,1	30,9	24,3	1,5
Suficiente (III)	5,0	13,9	25,6	30,6	8,6
Muy adecuada (IV)	4,1	3,3	11,3	23,1	88,3
Total	100,0	100,0	100,0	100,0	100,0

Fuente: INEE 2007: Tabla 1,11, p. 32.

El potencial positivo y negativo de la evaluación

La reciente difusión de un tipo particular de evaluaciones educativas, como son las pruebas de rendimiento, suele ir acompañada por una visión que las considera acríticamente como un elemento indispensable para la mejora educativa, idea que expresan frases como la que sostiene que *lo que no se evalúa no se puede mejorar*.

Es necesario matizar estas afirmaciones: la evaluación es una herramienta que puede tener consecuencias muy positivas, pero también puede tenerlas contraproducentes. Por ello es necesario precisar las condiciones para potenciar las implicaciones favorables y minimizar el riesgo de que se presenten las negativas.

Para ello conviene tener en cuenta el carácter multidimensional de la calidad, que lleva a la consideración ya apuntada de que no se debe identificar la evaluación con una de sus formas, las pruebas de rendimiento. Con los debidos cuidados, estas pueden servir para valorar el nivel de aprendizaje alcanzado por un grupo de alumnos en algunos temas, pero no atienden otras dimensiones igualmente importantes como la cobertura o la eficiencia.

Deben tenerse en cuenta también las características técnicas de las evaluaciones aplicables a cada dimensión de la calidad, como la confiabilidad y la validez. Incluso los indicadores tradicionales de dimensiones como la cobertura pueden tener deficiencias que impiden la comparación en el tiempo, pero tratándose de instrumentos de construcción más compleja, como las pruebas, los cuidados que hay que tener para asegurar precisión y comparabilidad en el tiempo muchas veces rebasan la capacidad técnica de los organismos responsables, en especial si se busca medir dominios que no se reduzcan a la comprensión lectora y la competencia en matemáticas y ciencias, y se quiere contar con mediciones absolutas y no solo relativas, referidas a estándares o criterios precisos y no solo a la norma estadística.

Un tercer elemento que tener en cuenta para que las evaluaciones no tengan consecuencias negativas, tiene que ver con sus implicaciones para escuelas y alumnos vulnerables. Un caso paradigmático de usos inapropiados de las evaluaciones, precisamente por no atender los tres puntos anteriores, es el que se presenta en ocasiones, cuando se da un peso central en las políticas educativas a los ordenamientos de escuelas o de sistemas educativos (*rankings*) basados en los resultados de los alumnos en algunas pruebas de rendimiento.

En los casos de abusos más claros, el peso que se da a tales *rankings* implica que se dejan a un lado las dimensiones de la calidad no medidas por las pruebas para centrar la atención exclusivamente en lo que ellas miden, que ni siquiera es el aprendizaje en el sentido más amplio, sino únicamente el que puede ser medido por tales instrumentos, ignorando lo que queda fuera de su alcance, que suele incluir los aspectos de mayor demanda cognitiva de los dominios considerados, además de áreas como la expresión oral y escrita, la historia y otras ciencias sociales, la educación artística y todos los aspectos actitudinales y relativos a valores.

En los casos mencionados se ignoran también las limitaciones técnicas de los instrumentos, cuyos resultados tienen siempre márgenes de error que no se toman en cuenta y suelen ser inestables en el tiempo, por lo que dan lugar a tendencias engañosas. Además se suelen perder de vista también las implicaciones del contexto más o menos vulnerable en que se sitúan escuelas y alumnos, con lo que se incurre fácilmente en valoraciones injustas (Martínez Rizo, 2010).

¿Es justo medir a todos los niños con la misma vara?

Conviene ampliar el punto anterior, pues la postura que se sostiene no coincide con la de algunas personas que consideran injusto en sí

mismo utilizar las mismas pruebas para evaluar a niños de contexto vulnerable y a otros de medio más favorable.

Es atractiva la idea de que sería injusto utilizar cierto instrumento para medir a los hijos de padres analfabetos atendidos en un plantel de cursos comunitarios o en una primaria indígena, y aplicar *la misma vara* a los chicos de familias acomodadas que van a escuelas privadas, pero no es una idea correcta, porque confunde medición y evaluación.

Medir implica una acción técnica que permite llegar a afirmaciones en términos de *más qué o menos qué*, o de *cuántas unidades caben en*, pero no en términos de *bueno o malo*. Para llegar a juicios de este último tipo, o sea para evaluar, es necesario además contrastar el resultado de la medición con un referente no empírico, sino normativo, que establece lo que es deseable, en comparación con lo cual podemos llegar a la conclusión de que una realidad es satisfactoria o no, aceptable o inaceptable.

A partir de la distinción anterior se entenderá la postura que aquí se sostiene, en el sentido de que una evaluación del nivel de aprendizaje que alcanzan los alumnos de un sistema educativo debe necesariamente utilizar los mismos instrumentos de medición —*la misma vara*— para medir a todos los alumnos, pero que para llegar al juicio de valor sobre la situación de determinados grupos de alumnos podrá emplear referentes distintos y que, además, deberán siempre explorarse las razones de las diferencias de los rendimientos de unos alumnos y otros, para evitar sacar conclusiones que, esas sí, podrían ser injustas.

Para aclarar la idea es útil otro ejemplo: la diferente presencia de desnutrición y otras circunstancias en los hogares de niños de diferente nivel socioeconómico y cultural, además de incidir en su rendimiento escolar, se refleja también en diferencias de su estatura promedio. No

por ello deberán emplearse escalas de medición diferentes para registrar la estatura de unos niños y otros, en un estudio sobre esa variable. Para tener mediciones comparables es fundamental, desde luego, utilizar instrumentos iguales.

Para llegar a juicios de valor sobre la situación de ciertos niños puede ser razonable, en cambio, utilizar referentes distintos. El juicio de valor depende en parte del propósito de la evaluación; tener cierta estatura —digamos 160 centímetros— puede dar lugar al juicio de que se trata de un sujeto normal, de uno excepcionalmente alto, o uno de estatura insuficiente, si en el primer caso se trata de una valoración general del estado de salud de chicos de 3° de secundaria; en el segundo de uno de niñas de 6° de primaria; y en el tercero de seleccionar a los integrantes del equipo de baloncesto de un bachillerato.

De manera análoga, un valor de 500 puntos —igual a la media nacional de una prueba de competencia lectora o matemática que se haya estandarizado en esa forma— podrá ser interpretado como normal, sobresaliente o demasiado bajo, dependiendo de que se trate de un alumno de condiciones familiares y escolares cercanas también al promedio nacional, de uno de condiciones particularmente desfavorables, o de uno de medio claramente privilegiado. Lo que hace cambiar el juicio es el punto de referencia, pero la vara de medir tiene que ser la misma, pues de lo contrario toda comparación pierde sentido.

Lo injusto no es medir la estatura de todos los niños, o su rendimiento en matemáticas o lectura con la misma vara —suponiendo que esté bien construida, lo que en el segundo caso no es sencillo— sino utilizar referentes inapropiados y, sobre todo, llegar a la conclusión de que la menor estatura o el menor desempeño promedio de los niños de contexto vulnerable es algo reprochable por ser imputable a su menor esfuerzo, o algo similar.

La difusión de resultados de pruebas mediante ordenamientos

La evaluación es una herramienta poderosa que puede ser muy positiva para la toma de decisiones, pero que también puede tener un impacto negativo sobre ellas, si no se hace bien, o si sus resultados no se interpretan y utilizan correctamente.

Los *rankings* son una forma atractiva de presentar resultados de evaluaciones, en especial de pruebas estandarizadas, pero la idea de que ofrecen un orden inequívoco de la calidad de las escuelas o sistemas educativos considerados carece de sustento sólido.

Como se ha apuntado, las pruebas estandarizadas, en especial si se aplican masivamente, tienen niveles de precisión gruesos y solo consideran el rendimiento en unas áreas y temas del currículo que se pueden medir con preguntas de respuesta estructurada. Ordenamientos basados solo en sus resultados son imprecisos y dan lugar a cambios inexplicables de lugar de un año a otro de las escuelas, especialmente las que tienen pocos alumnos. Aunque su propósito de corregir las desigualdades de la situación inicial de los alumnos es correcto, por diversas razones técnicas y prácticas los recientes acercamientos que utilizan modelos de valor añadido (*value-added models*) tampoco son satisfactorios, al menos por ahora (hay buenas discusiones sobre los modelos de *valor añadido* en Rubin, Stuart y Zanutto, 2004; Braun, 2005; Glazerman *et al.*, 2010; Leckie y Goldstein, 2009; McCaffrey *et al.*, 2003; y Rothstein, 2011).

Aunque no lo pretendan, los *rankings* vuelven de alto impacto las pruebas en que se basan, lo que lleva a prácticas que afectan la calidad de la enseñanza, como reducirla al contenido de las pruebas o formas más o menos sutiles de hacer trampa. Los *rankings* llaman la atención por el desconocimiento de sus límites por parte de diversos actores: medios de comunicación, que muchas veces buscan formas fáciles de dar resultados sensacionalistas; empresarios que suelen

tener una visión simplista de la educación; políticos que, por falta de tiempo, preparación o interés, no hacen el esfuerzo necesario para aprovechar la riqueza de la información que pueden proporcionar las pruebas y otros estudios y se dejan llevar por visiones superficiales de las formas de evaluar y de mejorar la calidad.

Las estrategias de mejora basadas en la asignación de estímulos económicos y el establecimiento de competencia entre las escuelas con base en los resultados de pruebas parten de una transferencia no sustentada de la economía y no tienen en cuenta las peculiaridades de la oferta y la demanda educativas, que no siguen la lógica del mercado.

Muchas personas no tienen conciencia de lo difícil que es tener buenos resultados con alumnos pobres. Es frecuente que dirigentes empresariales vean con simpatía estrategias simplistas de mejora, pensando que las fallas de la escuela pública se podrían corregir fácilmente con escuelas privadas como las que atienden a sus hijos, ignorando que solo una minoría, generalmente privilegiada, asiste a ellas. Por eso abundan las opiniones de que bastará con aplicar pruebas masivamente y tomar medidas correctivas simples para que la calidad mejore sustancialmente. En otras palabras, las estrategias simplistas de mejora parten de un supuesto falso: que hacer buena educación en cualquier contexto es fácil: «Los sistemas de rendición de cuentas basados en pruebas se basan en la creencia de que la educación pública puede mejorar gracias a una estrategia sencilla: haga que todos los alumnos presenten pruebas estandarizadas de rendimiento, y asocie consecuencias fuertes a las pruebas, en la forma de premios cuando los resultados suben y sanciones cuando no ocurra así» (Hamilton, Stecher y Klein, 2002).

Conclusión

La proliferación de *rankings* y su excesivo peso en las políticas públicas sobre educación están trayendo ya consecuencias negativas serias:

- Banalización del debate público sobre la calidad educativa, reducido a superficiales debates de tan discutibles ordenamientos, olvidando la complejidad del tema.
- Mercadotecnia engañosa de las escuelas, sobre todo de sostenimiento privado, que buscan atraer alumnos con base en esos ordenamientos.
- Empobrecimiento del currículo, por la tendencia a enseñar para las pruebas, descuidando aspectos que no serán evaluados, aunque sean importantes.
- Cansancio y desaliento en aquellas escuelas que, pese a sus esfuerzos, no consiguen resultados comparables con los de planteles cuyos alumnos tienen condiciones más favorables, y actitud negativa de los alumnos frente a una educación centrada en prepararlos para las pruebas.
- Finalmente, y como síntesis: empobrecimiento de las políticas públicas, que tienden a buscar soluciones fáciles a problemas complejos, descuidando otros aspectos fundamentales, en particular la equidad (Martínez Rizo, 2009 y 2010).

Las evaluaciones de mala calidad pueden tener impactos negativos, pero estos pueden ser especialmente fuertes sobre escuelas y alumnos que se ubican en contextos vulnerables, y esto puede ocurrir incluso cuando los instrumentos utilizados son de buena calidad técnica, si no se tiene cuidado en la selección de los referentes a utilizar para contrastar los resultados y llegar al juicio y, sobre todo, si la manera de difundir los resultados no considera responsablemente las condiciones del contexto de los evaluados, como ocurre fácilmente cuando la difusión se hace en la forma de los multicitados ordenamientos.

La amplia difusión y la gran atención que suscitan estas formas de difusión se entienden en el contexto de una preocupación sobre la calidad de la educación que comparten muchas personas y tiene sustento real, ya que hay razones objetivas para considerar que deberían mejorar muchas cosas.

La preocupación, además, cobra sentido en el marco de algunas tendencias más amplias que distinguen los tiempos actuales, como el interés por la rendición de cuentas, la búsqueda de transparencia en el manejo de los asuntos públicos y, con frecuencia, también cierta desconfianza con respecto a la educación pública y, en general, a la gestión pública de cualquier servicio.

A todo ello debe añadirse la escasa cultura en cuanto a evaluación educativa que suele haber no solo entre el público general, sino también entre maestros y autoridades educativas, e incluso entre investigadores y especialistas, que lleva a esperar resultados casi milagrosos en las escuelas gracias a la aplicación de pruebas.

Para estar a la altura de los retos del mundo actual, en muchos países es necesario sin duda que se den mejoras significativas en la educación. Esto es verdad, en particular, en relación con la calidad que ofrecen las escuelas del ámbito iberoamericano. La evaluación puede contribuir a ese propósito, pero no cualquier tipo de evaluación, en particular no la que consiste en ordenamientos de escaso sustento, basados únicamente en resultados de pruebas de aplicación masiva, sin reconocer sus alcances y límites.

Lo que se necesita es un uso parsimonioso de pruebas de buena calidad técnica, junto con otros tipos de evaluación: indicadores de otras dimensiones de la calidad del sistema educativo, como cobertura, eficiencia, equidad y otras; y, sobre todo, buenas evaluaciones de enfoque formativo hechas día a día por cada maestro, del avance de cada uno de sus alumnos, con acercamientos finos, distintos de los que implican las pruebas en gran escala.

Si las pruebas en gran escala dejan de considerarse la panacea para la mejora educativa, para verse como un complemento del trabajo de los maestros, podrán contribuir realmente a la mejora de la educación, que se dará en cada escuela y cada aula, o no se dará.

La evaluación puede ayudar o estorbar, pero lo único irremplazable para alcanzar la calidad educativa es un trabajo intenso y bien orientado en cada escuela y cada aula. Para ello hacen falta muchas cosas, incluyendo buenas evaluaciones; lo que no hace falta es

un uso de los resultados que desorienta a tomadores de decisiones y sociedad, lleva a unos a jactarse de logros dudosos y a otros, entre los que se cuentan personas serias y comprometidas, a perder incluso la poca esperanza que aún tenían.

Referencias bibliográficas

- BRAUN, H. I. (2005). *Using Student Progress To Evaluate Teachers: A Primer on Value-Added Models*. Princeton: ETS. Policy Information Perspective.
- GLAZERMAN, S., LOEB, S., GOLDBERGER, D., STAIGER, D., RAUDENBUSH, S. y WHITEHURST, G. (2010). *Evaluating Teachers: The Important Role of Value-Added*. Brown Center on Education Policy. Washington, DC.: Brookings Institution.
- HAMILTON, L. S., STECHER, B. M. y KLEIN S. P. (eds.) (2002). *Making sense of test-based accountability in education*. Santa Monica, CA: Rand Corporation.
- INEE (2009). *Panorama Educativo de México. Indicadores del Sistema Educativo Nacional. 2009 Educación Básica*. México: Instituto Nacional para la Evaluación de la Educación.
- INEE (2007). *La educación para poblaciones en contextos vulnerables. Informe Anual 2007*. México: INEE.
- LECKIE, G. y GOLDSTEIN, H. (2009). *The limitations of using school league tables to inform school choice*. Reino Unido, Bristol: University of Bristol.
- MARTÍNEZ RIZO, F. (2010). Usos y abusos de la evaluación. *Este País*. 232 (agosto), 24-27.
- MARTÍNEZ RIZO, F. (2007). Propuesta metodológica para desarrollar un sistema nacional de indicadores educativos para evaluar la calidad de la educación en México. *Seminario Internacional de Indicadores Educativos: Memoria*. México: INEE. 37-56.
- MARTÍNEZ RIZO, F. (2009). La evaluación de la calidad de los sistemas educativos. Propuesta de un modelo: En MARTÍN, E. y MARTÍNEZ RIZO, F. (coords.) *Avances y desafíos en la evaluación educativa*, Madrid: OEI.
- MCCAFFREY, D. F., LOCKWOOD, J. R., KORETZ, D. M. y HAMILTON, L. S. (2003). *Evaluating Value-Added Models for Teacher Accountability*. Santa Monica: Rand Co.
- ROTHSTEIN, J. (2011). *Review of «Learning About Teaching: Initial Findings from the MET Project»*. USA: Boulder, National Education Policy Center.
- RUBIN, D. B., STUART, E.A. y ZANUTTO, E. L. (2004). A Potential Outcomes View of Value-Added Assessments in Education. *Journal of Educational and Behavioral Statistics*, 29 (1), 103-116.

Abstract

Vulnerable contexts: the contribution of assessment

Strongly developed in different ways, educational evaluation has now a relevant role in educational policies in many countries. If it is well designed and implemented, and if its results are correctly analyzed and used, evaluation can support decisions that improve educational quality, but when these conditions are not fulfilled it can bring negative results in the form of quality deterioration. And both possibilities, either positive or negative, have specific features when related with schools and pupils of disadvantaged social backgrounds. This article reflects on this issue, mainly from Mexico's point of view. It starts with a definition of disadvantaged backgrounds (as an expression of the shortage of essential elements for family welfare and school development) —in relation to social inequality, so

deeply embedded in the Mexican society and school system— and of the quality of an education system; after that, the consequences for educational evaluation and assessment are explored, in the case of large-scale standardized tests and the diffusion of their results in the form of rankings.

Key words: *Disadvantaged background, Social inequality, Linguistic diversity, Equal opportunity, Quality of education, Evaluation, Learning outcome, Standardized test, Evaluation of the education system, Rankings.*

Résumé

Contextes vulnérables: les apports de l'évaluation

Sous des formes diverses, au cours des dernières décennies l'évaluation éducative s'est fortement développée dans de nombreux systèmes éducatifs et elle a acquis un rôle de premier plan dans la politique éducative de bon nombre de pays. Si elle est bien conçue et bien implémentée, et si les résultats sont analysés et utilisés d'une façon appropriée, l'évaluation peut appuyer des décisions conduisant à l'amélioration de la qualité éducative, mais si ces conditions ne sont pas respectées, l'évaluation peut avoir des résultats contre-productifs conduisant à une détérioration de la qualité du système. Aussi bien les possibilités positives que les négatives de l'évaluation, présentent des nuances particulières dans les cas des élèves et des écoles qui se trouvent en situation de vulnérabilité sociale. Cet article présente quelques considérations à ce sujet, spécialement du côté de l'expérience du Mexique. Après avoir précisé la notion de contexte vulnérable (comme l'expression du manque d'éléments essentiels pour le bien-être familiale et le développement scolaire) en ce qui concerne les inégalités qui caractérisent la société mexicaine et son système éducatif, nous présentons le concept de qualité appliqué au système éducatif, nous décrivons quelques types d'évaluations et nous analysons les implications des tests de performance à grande échelle et la diffusion de ses résultats sous la forme d'ordonnances.

Mots clés: *Contexte défavorisé, Inégalité sociale, Diversité linguistique, Égalité de chances, Qualité de l'éducation, Évaluation, Résultats d'apprentissages, Test standardisé, Évaluation du système éducatif, Classements.*

Perfil profesional del autor

Felipe Martínez Rizo

En 2002 fundó el Instituto Nacional para la Evaluación de la Educación, que dirigió hasta 2008. De 2007 a 2011 publicó cinco libros y coeditó otros cuatro, además de 13 capítulos y 14 artículos en revistas científicas. Presentó 34 ponencias o conferencias en México y 21 en otros 11 países. En 2009 recibió el premio del Consejo Mexicano de Investigación Educativa y en 2011 el Doctorado Honoris Causa por la Universitat de València.

Correo electrónico de contacto: fmartin@correo.uaa.mx