

Liliana Miranda Molina:

Diseñar e implementar evaluaciones es un desafío complejo

La jefa de la Oficina de Medición de la Calidad de los Aprendizajes (UMC) del Ministerio de Educación reflexiona sobre la trayectoria, aporte y perspectivas de la Evaluación Censal de Estudiantes (ECE) en nuestro país. Al mismo tiempo, brinda propuestas con miras a contar con un sistema de evaluación nacional que aporte a la mejora permanente de la calidad educativa.

Liliana Miranda Molina: Designing and implementing standardised evaluations is a complex challenge

The Chief of the Learning Quality Measurement Office (UMC) at the Education Ministry reflects on the performance, contribution and perspectives of the Student Census Evaluation (ECE) in Peru. Ms. Miranda also provides proposals for a national evaluation system for constantly improving education quality.

LILIANA MIRANDA MOLINA

lmiranda@minedu.gob.pe

Jefa de la Oficina de Medición de la Calidad de los Aprendizajes (UMC) del Ministerio de Educación, es licenciada en Sociología, estudios concluidos en Educación, estudios concluidos de maestría en Antropología, especialización en Estadística Aplicada por la Pontificia Universidad Católica del Perú.

PALABRAS CLAVE:

Aprendizajes
Calidad educativa
Estándares curriculares
Evaluación censal
Evaluación longitudinal

nes estandarizadas

TAREA: A diez años de aplicación de la ECE, ¿cuál es el balance desde la UMC, los mayores avances, los desafíos que aún quedan por delante y la agenda pendiente?

Liliana Miranda (LM): Con este año serían nueve. El uso de los resultados de la ECE se ha ido incrementando con el tiempo, y eso es positivo. En los últimos años ha ido ganando legitimidad entre los directores, docentes y funcionarios regionales como una herramienta que, al ofrecer información relevante, aporta a la reflexión y a la mejora en sus prácticas de gestión directiva y pedagógica, así como al diseño de intervenciones educativas, decisiones de política, entre otros. Sin embargo, todavía es necesario realizar esfuerzos adicionales con el propósito de preparar adecuadamente a estos actores para que puedan aprovechar mejor los resultados de las pruebas y utilizarlos en la mejora de su práctica.

Considero que un acierto desde que se empezó a implementar la ECE fue la decisión que tomó la UMC de elaborar informes específicos para los docentes que buscaban dar a los resultados, a las cifras, un contenido pedagógico. Asimismo, ha aportado de alguna manera a concretar, a hacer explícitas, aun con las limitaciones que conllevan los instrumentos estandarizados, las expectativas curriculares. Me atrevería a señalar que ahora hay una mayor claridad y acuerdo entre los docentes y directores sobre lo que esperamos que todos los estudiantes aprendan en Matemática al finalizar segundo grado o el nivel de lectura que deben haber desarrollado los de cuarto grado que asisten a una escuela de Educación Bilingüe Intercultural.

Hay un mejor posicionamiento de la idea de que todos los estudiantes deben alcanzar los aprendizajes esperados y tienen la capacidad para hacerlo, y de que esos aprendizajes deben ser de calidad, no mecánicos ni ru-

tinarios. Esto, obviamente, no solo se debe a la ECE, pues existe en este proceso otro conjunto de factores que han coadyuvado a ello; por ejemplo, los recursos pedagógicos que el ministerio ha producido estos años, como las *Rutas del Aprendizaje*, o los programas de acompañamiento, pero creo que la ECE ha cumplido un rol en este proceso. En términos más sociales y políticos, me parece que la ECE ha puesto su grano de arena en posicionar en la agenda pública, no solo nacional sino también local, la centralidad de los aprendizajes en la educación.

Desde el punto de vista de la función de la UMC, la ECE ha constituido un proceso de aprendizaje importante, pues se han podido ir mejorando los procedimientos de elaboración de instrumentos, implementación y controles de calidad del operativo, estrategias de análisis de datos y elaboración de los informes de resultados. Sin embargo, quedan aún muchos aspectos por mejorar y trabajar en las distintas dimensiones que supone, de los que podemos citar la ampliación de la evaluación a otros grados y competencias. Este año se realizará la primera evaluación de carácter censal en Secundaria, se evaluará a los estudiantes de segundo grado. Además, por primera vez se incluirán preguntas de formato de respuesta abierta. Asimismo, estamos preparando la evaluación en Ciencias y Ciencias Sociales para los siguientes años. En lo que toca a la diversidad, estamos ampliando la evaluación de dos lenguas originarias adicionales y se han elaborado pruebas para algunas necesidades educativas especiales asociadas a discapacidad.

Otros desafíos importantes, y a veces complejos de manejar, van por el lado de comunicar y difundir los resultados de una manera apropiada y estableciendo los controles para neutralizar los efectos no deseados que las evaluaciones de carácter censal traen consigo.

En otra dimensión, pero que tampoco se puede soslayar, persiste el desafío de tener un mercado con una oferta suficiente y competitiva de instituciones, empresa, organizaciones para los distintos procesos logísticos que requiere una operación de la dimensión de la ECE.

TAREA: ¿Qué aprendizajes nos deja la evaluación longitudinal que el MINEDU acaba de realizar?

LM: Uno de los hallazgos del estudio longitudinal realizado en escuelas públicas de Lima Metropolitana es que las curvas de crecimiento de los aprendizajes en Lectura y Matemática presentan una tendencia a incrementarse conforme transcurren los años de Educación Primaria, lo cual es alentador. Es decir, el desempeño de los estudiantes evaluados mejora a medida que éstos transitan en su escolaridad, de manera tal que a mayor cantidad de años en la escuela se obtienen mejores logros de aprendizaje. Sin embargo, los aumentos entre años son cada vez más pequeños. Concretamente, el crecimiento es mayor en los tres primeros grados que en los tres grados superiores. Entre otros aspectos, esto demanda que la formación de nuestros docentes atienda, tanto en lo disciplinar como en lo didáctico, a la consideración de la mayor complejidad de las capacidades y de los contenidos desarrollados desde el cuarto al sexto grado de Primaria, así como las estrategias metodológicas más pertinentes para su enseñanza.

Además, se ha encontrado, de manera consistente con los resultados en otras evaluaciones, que existe una brecha entre los aprendizajes establecidos en los documentos curriculares vigentes y, por ejemplo, las capacidades de comprensión lectora desarrolladas por los estudiantes promedio evaluados. Los resultados obtenidos muestran, por ejemplo, que la reflexión sobre el contenido del texto, capacidad que se espera que todos los estudiantes desarrollen desde los dos primeros grados, presenta evidencias de logro apenas a partir de cuarto grado, y la capacidad de reflexionar sobre la forma del texto no es lograda ni por los estudiantes de sexto grado.

Otros hallazgos del estudio son que, en el caso de Lectura, las variables que muestran una relación estadísticamente significativa con las diferencias en el rendimiento al finalizar el primer grado de Primaria fueron: sexo del estudiante, comunicación de padres e hijos, participación de los padres en actividades de la institución educativa, capital cultural, infraestructura

del hogar, satisfacción del docente, trabajo colegiado de los docentes, autoeficacia docente e infraestructura de la escuela. Una situación similar ocurre en Matemática, pues el modelo de factores asociados halló significancia estadística, pero con un tamaño del efecto muy pequeño, para variables como: sexo del estudiante, capital cultural, satisfacción docente y trabajo colegiado.

Desde un punto de vista metodológico, el estudio longitudinal ha permitido ganar experiencia en temas como la construcción de instrumentos pensados con una lógica de análisis longitudinal, la equiparación vertical de medidas y la aplicación de modelos lineales mixtos a datos de carácter longitudinal.

TAREA: Si bien una prueba nacional es parte de una evaluación estandarizada, nos encontramos en un país diverso, multiétnico y pluricultural; sabemos de los esfuerzos por las pruebas a población indígena en su lengua y en castellano, no obstante lo cual persiste la brecha negativa. ¿Es suficiente esta adecuación para abordar la diversidad? ¿Cuáles son los desafíos de una evaluación nacional en contextos de alta diversidad como el peruano? ¿Qué nuevos esfuerzos se están preparando?

LM: Diseñar e implementar evaluaciones estandarizadas es un desafío complejo en cualquier sociedad que por definición es diversa; un aula de clase también lo es, así que es probable que los esfuerzos se queden siempre cortos para aproximarse a ofrecer información pertinente a esa diversidad; pero es esencial que los equipos a cargo de las evaluaciones seamos conscientes del carácter diverso de los estudiantes. Es importante entender que este trabajo constituye un proceso en el que se busca mejorar continuamente, aunque, por otro lado, no siempre es fácil hacer los cambios a la velocidad que uno quisiera, pues se debe considerar que uno de los objetivos de las evaluaciones de logro de aprendizaje es ofrecer información comparable en el tiempo. Si se cambian aspectos sustantivos del diseño de una evaluación, se pierde la posibilidad de hacer comparaciones, pues no habría bases comunes desde donde hacerlos.

En paralelo a las evaluaciones para la población que tiene una lengua materna originaria y recibe una educación bilingüe, a partir del año pasado se han incorporado principios del diseño universal a los ins-



trumentos de evaluación, y este año se han diseñado pruebas orientadas a estudiantes con necesidades educativas especiales asociadas a discapacidad sensorial (visual y auditiva). Se han elaborado pruebas con macrotipos y en Braille que se aplicarán en la ECE 2015 a los estudiantes de segundo de Secundaria en algunas regiones. Además, se han trabajado adecuaciones para los estudiantes con discapacidad auditiva. Éstos son los esfuerzos en los que actualmente estamos trabajando.

TAREA: Existe una aparente disociación entre los resultados de la ECE y los que obtienen los mismos docentes con los mismos estudiantes aplicando pruebas en las mismas aulas. ¿Cuál es la relación entre la evaluación ECE y la que desarrollan los docentes en las aulas? ¿Existe algún nivel de articulación entre la ECE con los estándares de aprendizajes y mapas de progreso?

LM: Las evaluaciones de logro de aprendizajes, como la ECE, y las evaluaciones de aula, tienen objetivos distintos y usan procedimientos diferentes. Es evidente que las posibilidades con las que cuenta un docente para evaluar a un estudiante son diversas —no tienen límites de tiempo ni de formatos— y parten, o por lo menos se espera que así sea, de un conocimiento sobre las características específicas de sus estudiantes, lo que le debería permitir un acercamiento más detallado de sus necesidades y potencialidades. La ECE se realiza en un momento determinado y, de alguna manera, tiene que hacer una abstracción de lo

que todo estudiante promedio del país al finalizar un grado o ciclo debe saber y saber hacer. No obstante, la ECE busca ofrecer al docente la posibilidad de ubicar el desempeño de sus estudiantes y de las propias evaluaciones que él realiza en un marco de referencia más amplio: aquellos aprendizajes que como sociedad hemos convenido que son valiosos que todos nuestros estudiantes desarrollen. En este sentido, como en cualquier otro trabajo, siempre es bueno tener algún referente externo con el cual podamos contrastar nuestra propia praxis.

Ahora bien: no deja de llamar la atención que en la actualidad, si bien los resultados de una y otra evaluación no tienen necesariamente que coincidir, difieran tanto. Esto es una llamada de atención sobre las expectativas o estándares curriculares de nuestros docentes, pero también sobre la claridad con que éstos están expresados. Sobre esto último, el ministerio viene trabajando en años recientes. Es importante asegurar que todos los estudiantes del país logren desarrollar ciertos aprendizajes comunes que como sociedad hemos considerado que son valiosos. Es un derecho de nuestros estudiantes.

Esto me da pie para responder la segunda pregunta. El referente principal para la construcción de las pruebas de la ECE son los documentos curriculares, además de otros recursos pedagógicos como las *Rutas del Aprendizaje* y los textos escolares que usan los docentes, y referentes internacionales. En este marco, para la construcción de las pruebas también se han utilizado los

mapas de progreso en cuya construcción el equipo de especialistas de las áreas curriculares de evaluación de la UMC ha participado.

TAREA: Algunas regiones del país vienen implementando procesos de evaluación desde una perspectiva de descentralización. ¿Son estos esfuerzos convenientes? ¿Cuáles son los límites y posibilidades de evaluaciones de aprendizajes a escala regional? ¿Cómo se articulan los esfuerzos nacionales y regionales?

LM: En efecto, sabemos que en los últimos años algunas regiones han implementado diversas evaluaciones de logro de aprendizajes. Considero que estos esfuerzos son importantes de valorar en la medida en que buscan obtener información más aproximada al aprendizaje de los estudiantes de sus regiones que aquella que puede hacerse desde una perspectiva nacional; aun cuando debemos ser conscientes de que cualquier evaluación o investigación que busque generalizar tiene que asumir cierto grado de abstracción, pues las localidades y regiones también son diversas. Sin embargo, y ésta es la orientación que compartimos con nuestros colegas de las regiones cuando nos piden asistencia para esta labor, deben tener claro el propósito de la evaluación que quieren hacer, porque de ello dependerán las decisiones metodológicas que se deben tomar (no es lo mismo, por ejemplo, diseñar pruebas de carácter muestral que censal), y cuidar la calidad del diseño e implementación de las evaluaciones. Las evaluaciones de logro de aprendizaje que son implementadas por las instancias de gestión educativa siempre llevan un mensaje sobre lo que es relevante para los docentes y directores. Ello demanda que aquellos que desarrollemos evaluaciones lo hagamos con responsabilidad. Entre la disyuntiva de hacer malas pruebas y no hacerlas, es mejor no hacerlas, sin ninguna duda.

Por otro lado, sería pertinente que las regiones que estén implementando evaluaciones complementen los esfuerzos nacionales y no dupliquen los grados y competencias que ya son evaluados. El objetivo de las evaluaciones de logro de aprendizajes no es entrenar a los estudiantes en la resolución de pruebas. Al estar a cargo de territorios más acotados, sería deseable que las regiones evalúen habilidades importantes, cosa que muchas veces, desde una lógica nacional, no es posible hacer, y sobre las cuales es importante conocer cuál es el nivel de competencia de los estudiantes (por ejemplo, escritura).

TAREA: El 2013, el CNE hizo llegar a la UMC algunas recomendaciones. ¿Cuáles se han implementado y con qué resultados?

LM: Gracias por la pregunta. En efecto, en 2013 el CNE nos alcanzó un conjunto de sugerencias, lo cual agradecemos, pues nos permiten mejorar nuestro trabajo y estar alertas sobre aspectos que debemos atender. Por el tiempo, solo me referiré a algunas de ellas. Entre las recomendaciones estaba el continuar con la evaluación de estudiantes indígenas y considerar la evaluación de estudiantes con discapacidad que asisten a escuelas regulares. Respecto a ello, la UMC continúa con el esfuerzo evaluador en lenguas originarias y en castellano como segunda lengua; de hecho, somos el único país de la región que actualmente desarrolla este tipo de evaluaciones. En la última ECE se incluyó la evaluación en quechua-chanka, y este año se está piloteando una en asháninka. A los esfuerzos para evaluar de manera más pertinente a los estudiantes con necesidades educativas especiales asociadas a discapacidad ya me referí.

Una segunda recomendación que hizo el CNE fue presentar los resultados de rendimiento asociados al nivel socioeconómico de las familias de los estudiantes evaluados, considerando la evidencia acumulada en los últimos años sobre la fuerte asociación existente entre esas dos variables. Estamos trabajando en ello, así que tenemos previsto que la ECE 2015 presente los resultados ajustados por tal indicador.

Otra de las recomendaciones del CNE fue publicar un informe técnico de las evaluaciones con el objetivo de dar cuenta y hacer explícitos los procedimientos seguidos en el diseño de los instrumentos y los cálculos de los resultados. Hemos procedido con la elaboración y publicación, a través de la web de la UMC, de los reportes técnicos de la ECE 2013 y 2014. Sus contenidos incluyen: construcción de las pruebas aplicadas en la ECE; definición de la población y de la muestra de control; organización y descripción del operativo de campo, propiedades psicométricas de las pruebas aplicadas, modelo de medición adoptado, criterios y resultados del procesamiento de la información recogida, entre otros.

Finalmente, se propuso desarrollar un plan de evaluación a largo plazo que incluyera evaluaciones longitudinales y muestrales y la participación en evaluaciones internacionales. Como ya mencioné, la UMC acaba de publicar los resultados del estudio longitudinal titula-



do “Aprendizajes de primero a sexto de primaria en lectura y matemática”. Se ha llevado a cabo la evaluación muestral de sexto grado de Primaria, y este año, aparte de las evaluaciones latinoamericanas y de PISA, participaremos en la evaluación internacional de Cívica y Ciudadanía.

TAREA: Existe una corriente crítica respecto a las pruebas estandarizadas: por un lado, se afirma que fomenta prácticas de “entrenamiento hacia los estudiantes”; por otro, se sostiene que estamos en una lucha en la que va ganando la razón jerárquica, la uniformidad, la inequidad, la reducción a una parte de los aprendizajes, entre otros. ¿Cuáles son las ventajas y cuáles los límites de una prueba nacional como la ECE? ¿Cómo se están neutralizando los riesgos de los efectos inesperados que puede generar?

LM: En primer lugar, los problemas aparecen cuando se quiere equiparar la calidad educativa con los resultados de las evaluaciones de logro educativo. La calidad de la educación no se puede reducir a los resultados de las evaluaciones de logro educativo. Estas evaluaciones se limitan a medir lo que los estudiantes saben y deben saber hacer en determinados grados según lo establecido en nuestros documentos curriculares, y que son posibles de ser medidos a través de pruebas de lápiz y papel. Es decir, buscan ofrecer información sobre una parte de aquello que como sociedad hemos acordado que es relevante que nuestros estudiantes aprendan. Como cualquier evaluación o investigación,

evidentemente ésta tiene límites. La UMC, por medio de los informes de resultados y de la difusión que hace, siempre intenta que estas ideas queden claras. Desde el inicio de la ECE se apostó por elaborar informes pedagógicos para cada docente de las áreas evaluadas, y no limitarnos a la mera presentación de cifras; se trataba de dar contenido pedagógico a esas cifras (por ejemplo, detallar lo que significa, en materia de conocimientos y habilidades, que un estudiante esté ubicado en el nivel satisfactorio o en el nivel en inicio; que, dicho sea de paso, no significa que no sepan nada); cosa que no todos los sistemas de evaluación hacen.

Por otro lado, estas limitaciones no deberían ser problemáticas si todos entendemos el propósito de las evaluaciones estandarizadas y ayudamos a hacer sentido común de esto. Las evaluaciones buscan ofrecer información sobre el nivel de aprendizaje de los estudiantes en determinados hitos de su escolaridad obligatoria para apoyar la reflexión y toma de decisiones sobre política educativa. Son importantes en la medida en que el Estado y el sistema educativo necesitan dar cuenta a la sociedad, a los ciudadanos, en qué medida está logrando asegurar el derecho de las personas a una educación de calidad, y para ello las evaluaciones aportan con algunos indicadores. La evaluación del sistema educativo requiere de un conjunto amplio de factores, no solo de los que tienen que ver con los aprendizajes, sino también, por ejemplo, del grado de eficiencia con que estamos asegurando la universalidad en el acceso, el tránsito



TAREA / CÉSAR SALHUANA

y la conclusión de la educación obligatoria, además de otros aspectos como las oportunidades de aprendizaje de los estudiantes.

Habiendo señalado esto, es cierto que las evaluaciones estandarizadas tienden a traer consigo efectos no deseados. Lo que mencionas es conocido en la literatura como *consecuencias indeseadas de los test*, que podrían afectar el derecho de los estudiantes y constituir una amenaza a la validez de una evaluación. Debemos estar alertas al estrechamiento curricular, al sobreenentrenamiento al que pueden estar sometiendo en ciertas escuelas a estudiantes de los grados evaluados, al descuido de las cohortes de estudiantes que no son evaluados o a la selección de estudiantes de bajo rendimiento. Todos debemos estar alertas a esto. Como ya mencioné, desde la UMC tratamos, en los procesos de difusión, de señalar explícitamente que estas prácticas son negativas, y prendemos las alertas a las direcciones y oficinas del Ministerio y a las instancias de gestión educativa descentralizada para que tengan también los cuidados del caso. Además, estamos empezando a recoger información y a hacer algunos análisis donde hay datos disponibles para ponderar los efectos de estas prácticas.

TAREA: ¿Por qué tendríamos que seguir apostando por la ECE en el Perú?

LM: Considero que debemos seguir apostando a tener un sistema de evaluación que en su conjunto, a través de lo que aportan las evaluaciones muestrales,

las censales y las internacionales, nos permita contar con información sobre el nivel de aprendizaje de nuestros estudiantes en determinadas competencias claves e hitos de su educación básica obligatoria. En ese marco, la ECE busca aportar en la dirección de ofrecer información pertinente a los distintos actores para que se responsabilicen por su desempeño en el ámbito que le compete. Su carácter censal hace posible que la información se desagregue y así se devuelva a cada actor educativo que tiene un rol que cumplir en la tarea del desarrollo de aprendizajes de nuestros estudiantes.

Indudablemente, estas evaluaciones tienen límites, y por ello el esfuerzo continuo que debemos hacer todos para que sus resultados se entiendan en el marco y propósito para el que fueron diseñados, de ponderar las interpretaciones al momento en que se difunden los resultados. Sin embargo, por otro lado, pensar que por sus límites se puede prescindir de las evaluaciones de logro de aprendizaje, no sería adecuado. En los últimos años las evaluaciones han permitido, con toda las limitaciones que pueden tener, conocer que los aprendizajes han mejorado, no en la medida en que todos esperamos, aunque sí lo han hecho —los resultados de las últimas evaluaciones nacionales e internacionales son consistentes con esta afirmación—; pero, por otro lado, también han permitido visibilizar las preocupantes brechas de equidad que actualmente existen entre los distintos grupos poblacionales, y a cuyo esfuerzo de reducción todos nos debemos avocar. **T**