



Julio De 2010

25 AÑOS DEL SIMCE

Por Iván Ortiz
Investigador del CIDE

El Sistema de Medición de la Calidad de la Educación (SIMCE) cumple 25 años de vida, según señala su sitio web institucional. Desde su creación, el lugar que ocupa la educación nacional en el debate público ha crecido enormemente, en parte gracias a él. Sin embargo, es una institución que despierta mucha controversia en torno a sus usos; los críticos del SIMCE son legión entre los profesores y profesoras, y entre los académicos de las facultades de educación. En las próximas páginas revisaremos esta controversia y qué podría estar en juego en ella.

El origen del actual SIMCE se remonta al año 1982 en que se inicia un programa de evaluaciones del desempeño escolar en 4º y 8º año de enseñanza básica, conducido por la Universidad Católica de Chile (Programa de Evaluación del Rendimiento, PER). Si bien este programa no fue la primera experiencia de evaluación estandarizada del desempeño escolar en Chile –lo fue la Prueba Nacional aplicada a los 8º básicos desde 1968 a 1971– puede ser considerado el antecesor del SIMCE actual por la unidad de *propósitos* de ambos y por la continuidad de sus creadores. En efecto, el PER pretendía producir y comunicar resultados educativos para ayudar a los actores de la educación a tomar decisiones informadas; la novedad radicaba en que entre estos actores, los padres y apoderados figuraban como destinatarios cruciales. Se esperaba de ellos que usaran esa información para elegir un buen establecimiento para sus hijos, generando una demanda calificada, capaz de movilizar una oferta educativa de calidad. Tal propósito era, por cierto, complementario a la privatización de la educación y al financiamiento de la misma mediante el sistema de vouchers, controvertidas reformas de la época.

Sin embargo, en los dos años que duró el PER el propósito anterior no se concretó. Eyzaguirre y Fontaine (1999)¹ relatan que el Ministerio de Educación de la época se resistió a la publicación de los resultados con el argumento de que se estaba juzgando la labor del profesor sin tomar en cuenta las realidades socio-económicas del alumnado. Vemos que la controversia en torno a este punto acompaña al SIMCE desde sus orígenes². Posteriormente, la LOCE de 1990 incluyó la obligación del Ministerio de Educación de publicar los resultados de los establecimientos en diarios de circulación nacional o regional, doblando así su resistencia. Pero hubo que esperar hasta 1995 para que se cumpliera la ley y los resultados escolares de los establecimientos fueran publicados en los diarios por primera vez.

El mismo equipo que creó el PER dio origen al SIMCE actual, en la Universidad Católica en 1988, fecha de la primera evaluación nacional del rendimiento escolar bajo tal sigla. En 1992 el SIMCE fue traspasado totalmente al Ministerio de Educación.

¹ Eyzaguirre, B y Fontaine, L (1999): “¿Qué mide realmente el SIMCE?”, *Estudios Públicos*, 75, pp. 107-161.

² Vemos también al Ministerio de Educación de la época más identificado con el campo educacional, probablemente en disputa con otros campos.

El carácter controversial del SIMCE se refiere a una diversidad de aspectos. El más emblemático es tal vez su uso mediático como ranking de establecimientos, en la medida en que hace visible y respalda una visión de la provisión del servicio educativo basada en el funcionamiento del mercado. No se trata solo de una crítica conceptual, o ideológica, de aquellos que prefieren que la educación *no* sea regulada por el mercado; se trata también de la estigmatización que sufren los profesores, directores y estudiantes de los establecimientos que obtienen bajos resultados, los que coinciden en gran medida con los establecimientos más pobres.

Pero hay otros cuestionamientos. Uno de ellos sería la reducción de la noción de calidad de la educación a los aprendizajes que mide el sistema. Es cierto que los aprendizajes que mide el SIMCE son solo una parte del currículum nacional, la parte que más preocupa a los establecimientos, pero es difícil asegurar que esto estaría empobreciendo las oportunidades de aprendizaje que los establecimientos ofrecen a sus alumnos en las demás asignaturas o en los Objetivos Fundamentales Transversales. Es posible que esta crítica esconda otra, relacionada con el estatus diferencial de las distintas disciplinas, lo que tiene raíces más hondas que el SIMCE.

Otro cuestionamiento, más grave, se refiere a los aprendizajes que la prueba evalúa, los que no serían todo lo relevantes que se quisiera, y modelarían negativamente la enseñanza. Esto no es tan cierto. Ha habido esfuerzos deliberados de la institución por mejorar las pruebas y progresos reales; si se comparara una prueba de mediados de los 90 con una actual, se encontrarían ítemes de mejor calidad y más relevantes en la prueba actual. Pero es una línea de trabajo en la que aún hay mucho por hacer, y donde la investigación especializada ha estado ausente³.

Si las críticas al SIMCE fueran ciertas, implicaría que este habría adquirido un enorme poder de conducción (o de extravío) de la educación, tanto de la demanda por el servicio educativo como de los procesos de enseñanza al interior de los establecimientos. En nuestra opinión, ni lo uno ni lo otro es evidente.

En el año 2003 el Ministro de Educación convocó a una comisión de expertos para hacer un diagnóstico del SIMCE, especialmente de la información que éste proporcionaba y hacer recomendaciones para optimizar su impacto sobre la calidad de la educación. Por esa fecha el descontento de muchos profesores, directores y académicos de la educación con el SIMCE era notorio. Lo que gatilló la decisión del Ministro fue la crítica según la cual el SIMCE no solo estigmatizaba a los establecimientos más pobres, sino que además la información que proporcionaba a los profesores y directores no era suficientemente útil como para fundamentar acciones pedagógicas específicas. Todo esto en el marco de una crítica generalizada de la calidad de la educación y del compromiso del Ministerio de mejorarla.

³ Ver Eyzaguirre, B y Fontaine, L (1999, op cit). Las autoras revisaron pruebas del Simce y levantaron una fuerte crítica respecto a la relevancia de los aprendizajes medidos. Por esa fecha, el Ministerio de Educación, que había formulado una reforma al currículo, impulsó modificaciones en las pruebas del SIMCE para alinearlas con las exigencias de la reforma curricular.

El diagnóstico que hizo la comisión de expertos reconoció aspectos positivos y negativos del SIMCE⁴. Entre los primeros, el hecho de que el sistema estaba instalado en la tradición del sistema educativo y gozaba de legitimidad y credibilidad; que había permitido centrar la atención de la opinión pública y de los docentes en los resultados de aprendizaje; que la calidad de sus instrumentos y metodologías de comparación de resultados había experimentado avances importantes; que era ampliamente utilizado como herramienta de apoyo a la formulación y evaluación de políticas educativas, y que pese a todo, era utilizado por los docentes para analizar los logros de sus establecimientos.

Entre los aspectos negativos, la Comisión citó el hecho que los resultados reportados a los establecimientos no tenían mayor significado para orientar las prácticas pedagógicas, pues dejan en la oscuridad qué saben o no saben los estudiantes y si los resultados alcanzados son o no suficientes. Asimismo, reconoció que el SIMCE incentivaba prácticas no deseadas de exclusión y selección de alumnos como estrategia para mejorar los resultados, y de concentración de esfuerzos en las generaciones de estudiantes evaluados por este instrumento, en desmedro de las demás. También se mencionó que el manejo comunicacional de los resultados por los medios ha distorsionado el sentido de la medición, cuando presenta ranking de establecimientos sin tomar en cuenta la influencia en ellos de los antecedentes familiares de los estudiantes. Otro aspecto negativo señalado fue el insuficiente conocimiento de los resultados SIMCE por parte de los padres y apoderados y el hecho de que no había datos concluyentes acerca de cómo y cuánto ellos usaban esta información.

Un aspecto interesante del diagnóstico de la Comisión se refiere al mal manejo comunicacional de los resultados: “significa una presión enorme para los gestores de política educacional, que en ocasiones los puede llevar a anunciar medidas impulsadas más por la fuerza de la opinión pública, que informados por un análisis profundo y sistemático” (op .cit. pag. 41). No hay más detalles en el texto que expliciten el sentido de tal enunciado, pero podría interpretarse como que las autoridades educativas no deberían suprimir la publicación de los resultados SIMCE solo porque el manejo comunicacional de los mismos por los medios genera descontento en los usuarios.

La Comisión hizo una serie de recomendaciones frente a este diagnóstico, partiendo con un respaldo total al SIMCE en cuanto sistema encargado de entregar públicamente información válida y confiable sobre los resultados de aprendizaje de los alumnos de todos los establecimientos. De este modo, reafirmó la importancia de todos los usos que había tenido el SIMCE: con fines pedagógicos, de rendición de cuentas, de insumo para formulación y evaluación de políticas y de investigación. En general, la literatura admite que diferentes usos requieren diferentes instrumentos. Además, la Comisión prescribió entregar *más y mejor* información sobre estos resultados: aumentar la periodicidad de las mediciones en 4° básico, incluir nuevas áreas curriculares, reportar resultados en relación a estándares de desempeño, y canalizar directamente la entrega de información a todos los docentes y padres de cada establecimiento.

Todas estas recomendaciones se han venido cumpliendo desde entonces, en diversos grados. Las mediciones en 4° básico se realizan todos los años a partir de 2005, en vez

⁴ Ver: Comisión para el Desarrollo y Uso del SIMCE (2003): *Evaluación de Aprendizajes para una Educación de Calidad*. Ministerio de Educación, Santiago.

de cada tres; la periodicidad de las mediciones en 8° básico y 2° medio también aumentó, de cada tres a cada dos años. La inclusión de nuevas áreas curriculares se materializó en la medición del desempeño en escritura de los alumnos y alumnas de 4° básico en los dos últimos años. La entrega de resultados en relación a estándares de desempeño comenzó a hacerse efectiva en 2008 en 4° básico y se agregó 8° básico en la última medición, siendo tal vez la innovación informativa más importante y a la cual la comisión le había otorgado la primera prioridad de sus recomendaciones. Las modificaciones en los mecanismos de entrega de la información de logros escolares a los usuarios están viendo la luz este año, no sin controversia, con instrumentos que amplían la accesibilidad de los padres a los resultados comparados de los establecimientos.

¿Ha mejorado el uso pedagógico de los resultados SIMCE?

En general, la importancia otorgada al SIMCE por los docentes para realizar acciones de mejoramiento de los aprendizajes en los establecimientos ha crecido con el paso del tiempo. De acuerdo a la VII Encuesta a Actores del sistema Educativo, realizada por el CIDE en 2008, en el año 2000 la mitad de los profesores le daba poca o ninguna importancia al SIMCE para promover acciones de mejoramiento del aprendizaje, en cambio en los últimos años esa proporción ha disminuido a la cuarta parte, aproximadamente. Los directores le asignan más importancia al SIMCE que los profesores, pues frente a la misma pregunta solo el 6% de ellos contestó “poca” o “ninguna” en 2008. Estos patrones de resultados son muy similares entre las tres dependencias de los establecimientos.

Profesores					
¿Qué importancia le asigna usted al SIMCE para realizar acciones de mejoramiento de los aprendizajes en el establecimiento?					
	2000	2001	2003	2006	2008
Mucha	11%	14%	26%	23%	31%
Mediana	39%	45%	58%	51%	46%
Poca o ninguna	50%	40%	15%	26%	23%
Total	100	100	100	100	100

La pregunta formulada a los docentes y directores no se hace cargo de cuáles son las “acciones de mejoramiento”, ni distingue entre acciones del director y del docente, las que pueden ser muy diversas y de diversos grados de generalidad, como sustituir profesores, poner presión a los actores, aumentar las horas de las asignaturas medidas, o aumentar la cantidad de textos de lectura exigidos a los estudiantes, por ejemplo. El cuadro siguiente permite conocer algo más, pero sin detalles pedagógicos, cuáles son los usos más relevantes de la información aportada por el SIMCE.

De los siguientes usos que se pueden hacer de los resultados del SIMCE, ¿cuál sería el más relevante en su opinión?				
	Profesores 2004	Directores 2004	Profesores 2008	Directores 2008
Ayuda a definir las metas a alcanzar en el establecimiento	20,9%	24,1%	11,3%	12,9%
Permite identificar áreas de aprendizajes con problemas	54,5%	54,9%	34,2%	31,9%
Ayuda a definir posibles acciones para mejorar aprendizajes	14,3%	13,1%	35,5%	40,3%
Permite una mejor planificación del año escolar	1,0%	1,3%	4,2%	5,7%
Sirve a los padres para conocer la calidad del establecimiento	1,3%	1,7%	3,6%	2,7%
Es útil para la evaluación de las políticas educativas	2,5%	,8%	3,2%	3,2%
Sirve para producir un ranking entre los establecimiento	4,4%	2,5%	4,4%	0,9%
No utilizo los resultados del SIMCE, no son relevantes	1,0%	1,7%	3,6%	2,5%
Total	100,0%	100 %	100,0%	100 %

El cuadro muestra que: i) profesores y directores tienen opiniones muy parecidas, ii) el año 2004 los usos más relevantes eran la definición de metas y la identificación de áreas de aprendizaje con problemas, los cuales son de una gran generalidad; iii) en el año 2008 hay un aumento importante de las preferencias por “definir posibles acciones para mejorar aprendizajes”; iv) los resultados SIMCE no se usan para mejorar la planificación del año escolar. Es posible que los cambios entre el 2004 y el 2008 se deban a las innovaciones en la información reportada por el SIMCE a partir de 2007.

Pondremos ahora el foco del análisis en la apreciación que tienen los docentes de las novedades informativas que el SIMCE les hace llegar. La idea es que los docentes reciban resultados que vayan más allá del puntaje y tengan mayor significado y utilidad pedagógica, describiendo de qué son capaces los alumnos del sistema y de cada establecimiento, y cuántos han alcanzado o superado las expectativas que se establecen. En efecto, los docentes están recibiendo información de esta naturaleza, bajo la denominación *niveles de logro*, para Lectura, Escritura⁵, Matemática y Comprensión del Medio Social Natural y Cultural de 4° básico; y par a Lectura y Matemática de 8° básico. Se trata de una información cualitativa y cuantitativa a la vez, en la medida en que los niveles de logro (avanzado, intermedio e inicial) son descripciones de los desempeños que son capaces de hacer los estudiantes, y se acompañan del porcentaje de ellos que

⁵ La Prueba de Escritura ha sido aplicada solo a una muestra representativa de estudiantes, pese a ello el análisis de los desempeños de los estudiantes, informado en el documento de resultados, puede ser utilizado por los docentes para evaluar los escritos de sus propios alumnos con relativa facilidad.

está en cada nivel. En principio, permite a los establecimientos fijar metas pedagógicas más específicas que si se cuenta solo con el puntaje (por ejemplo 260) como información.

En su encuesta 2008, el Centro de Investigación y Desarrollo de la Educación preguntó a los docentes y a los directores por el principal aporte de los resultados con niveles de logro, después de la segunda vez que los establecimientos habían recibido esta clase de información para 4° básico. Los docentes tienen una impresión positiva de la utilidad de los niveles de logro.

El cuadro muestra que el principal aporte para los docentes⁶ y para los directores es que la nueva información ha permitido hacer cambios en la planificación de los aprendizajes del primer ciclo básico; en menor medida, también la planificación de 5° básico (es decir, el curso al que se refieren los resultados). Al contrastar estos datos con los anteriores referidos a la utilidad de los resultados SIMCE en general, encontramos que el aporte de los niveles de logro para la planificación es una novedad interesante.

El cuadro revela, de paso, que los docentes y directivos preferirían, razonablemente, focalizar sus esfuerzos estratégicos en acciones que rendirían frutos en el corto plazo en vez del mediano y largo plazo. Si no, ¿por qué las preferencias se inclinan por hacer cambios en la planificación del primer ciclo en vez de hacerlo en el 5° año? La información proporcionada por los niveles de logro describe las fortalezas y debilidades de los estudiantes de 4° básico del año anterior, que ahora están en 5° año; sería de la mayor pertinencia hacer cambios en la planificación de los aprendizajes de estos alumnos y alumnas, una vez recibida la información cualitativa aportada por los niveles de logro. Pero los docentes y directivos se inclinan por hacer cambios en la planificación del primer ciclo, probablemente porque desean prepararse mejor para la prueba SIMCE de 4° básico que tendrá lugar ese mismo año.

Aporte de los resultados SIMCE con niveles de logro. Docentes de básica y directores, 2008		
	Docentes de básica	Directores
Precisar cuántos alumnos están bajo lo esperado	24,3%	19,3%
Conocer qué saben hacer los alumnos de 5° año	12,2%	10,1%
Hacer cambios en la planificación de la enseñanza del 5° año	8,1%	7,8%
Hacer cambios en la planificación de la enseñanza 1° ciclo	39,7%	56,9%
No han sido más útiles que los puntajes brutos	5,9%	3,7%
Han complicado la comprensión de los resultados SIMCE	5,1%	0,5%
No sé	4,6%	1,8%
Total	100,0%	100,0%

Si bien se trata de indicios favorables, es necesario conocer con más precisión de qué manera la enseñanza se está beneficiando con esta manera de informar los resultados SIMCE; las preguntas de la encuesta solo dejan ver aspectos muy generales. Por de pronto, la expresión “hacer cambios en la planificación de la enseñanza” no aclara si se

⁶ Incluye solo a los docentes de básica, por ser los concernidos con la pregunta.

refiere a la planificación general del año, o a la planificación clase a clase, las que refieren a dos niveles de concreción muy diferentes.

Lo que está en juego con el uso pedagógico de los resultados SIMCE es su *validez* como sistema de medición. De todos los criterios de validez de un instrumentos de medición educacional, el principal debería ser su sensibilidad pedagógica⁷, es decir, hasta qué punto la prueba tiene una afinidad tal con la enseñanza, que la varianza de sus resultados no solo refleja la varianza en la calidad de la enseñanza, sino que además tales resultados permiten orientar la misma de manera *específica*. Un sistema de evaluación externa sensible a la enseñanza hace posible, a partir de la calidad de la retroalimentación que entrega, intervenir pedagógicamente en las deficiencias y, subsecuentemente, *mejorar* los resultados. Así como los resultados de los exámenes de sangre del paciente permiten tomar decisiones clínicas precisas al especialista, y monitorear su efecto terapéutico en el examen siguiente.

Si bien la calidad de la retroalimentación depende, en cierta medida, de cómo ésta es interpretada por los docentes⁸, la riqueza informativa intrínseca de los datos reportados, es fundamental. Hasta antes de los niveles de logro, podríamos decir que la validez del SIMCE era al menos discutible, si no francamente cuestionable. Por una parte, este sistema ha tenido serios problemas para demostrar que la varianza de sus resultados refleja la varianza de la calidad de la enseñanza; por el contrario, ha puesto en evidencia que esta varianza depende principalmente del capital cultural y económico de los estudiantes. Por otra, no ha sido capaz de mostrar que la retroalimentación que entrega a los establecimientos ha sido aprovechada por los docentes para mejorar sus prácticas y los aprendizajes de los estudiantes.

Cuando las mediciones del rendimiento sirven como criterio principal para el juicio sobre la eficacia educativa en los sistemas de rendición de cuentas, se requiere evidencia que muestre una relación defendible entre los puntajes de la prueba y la calidad de las experiencias educativas ofrecidas en la escuela (Baker, 2009 op cit). De otro modo, las decisiones asociadas a la rendición de cuentas se revisten de arbitrariedad, lo que termina por despojar de toda validez a la medición.

⁷ Seguimos planteamientos de Eva Baker (2009): "Consideraciones de validez prioritaria para la evaluación formativa y de rendición de cuentas". *Revista de Educación* N° 348, pp 91-109. (Recuperado en www.revistaeducación.mec.es, en julio 2010). La validez "instruccional" es un tema importante en el debate sobre evaluación y rendición de cuentas en USA.

⁸ La Comisión antes aludida piensa que existe un déficit de formación que dificulta que los docentes comprendan y utilicen los resultados de la medición (op cit pag.11)