

Informe Ejecutivo

Equipo de Tarea para la Revisión del Sistema Nacional de
Evaluación de Aprendizajes

Equipo de Tarea para la Revisión del Sistema Nacional de Evaluación de Aprendizajes

Informe Ejecutivo

1. Contexto de la convocatoria del Equipo de Tarea

El día 15 de julio de 2014 la Subsecretaria de Educación, Sra. Valentina Quiroga, conformó el *Equipo de Tarea para la Revisión del Sistema Nacional de Evaluación de Aprendizajes* con el propósito de generar un diagnóstico y recomendaciones respecto del mismo.

Un conjunto de antecedentes motivaron la conformación del Equipo de Tarea. Un primer antecedente relevante, fue la referencia que respecto del SIMCE se hiciera en el programa de Gobierno de la Presidenta Michelle Bachelet, donde se plantea “impulsar una comprensión más compleja e integral de calidad, para recuperar el sentido de la labor educativa”. Adicionalmente, el Ministerio de Educación debía presentar al Consejo Nacional de Educación un nuevo calendario de mediciones para el periodo 2016-2020, para lo cual determinó que solicitaría una revisión externa además de la asesoría y propuesta que recibiera de la Agencia de Calidad a inicios de 2014 al respecto¹. Por otra parte, desde 2013 se habían evidenciado diversas manifestaciones críticas al SIMCE por parte de movimientos sociales² y de un grupo de investigadores y académicos que firmaron una carta abierta cuestionando varias características de la medición nacional, como también sus políticas asociadas. Finalmente, en diciembre de 2014, la Agencia de Calidad debía dar inicio, en marcha blanca, a la ordenación o clasificación de las escuelas del país en cuatro categorías de desempeño en concordancia con la Ley de Aseguramiento de la Calidad (Ley 20.529). Esta Ley define un conjunto de consecuencias para las escuelas según las categoría en que quedan clasificadas y el tiempo que permanecen en dicha categoría. El criterio de mayor peso en esta categorización es la efectividad de los establecimientos en promover aprendizajes, según su rendimiento en el SIMCE.

En este contexto, y de acuerdo a los términos de referencia del Equipo, se espera que sus recomendaciones sean un insumo para que ‘el Ministerio de Educación proponga una política de evaluación de aprendizajes para los próximos años y formule el Plan Nacional de Evaluaciones a presentar al Consejo Nacional de Educación’³.

El Equipo de Tarea fue presidido por Lorena Meckes (investigadora del Centro de Estudios de Políticas y Prácticas en Educación, CEPPE) y conformado además por Paulina Bustos (Jefa del Departamento de Educación de Curicó); Alejandro Carrasco (Subdirector de CEPPE); Rafael

¹ De acuerdo a la Ley de Aseguramiento de la Calidad, el Ministerio de Educación debe someter al Consejo Nacional de Educación el plan de evaluaciones nacionales e internacionales, para lo cual debe coordinarse con y recibir la asesoría de la Agencia de Calidad, ‘especialmente en lo que se refiere a su viabilidad y requerimientos de implementación (Artículo 11 letra ñ, Ley 20.529).

² Organizaciones de secundarios que en 2013 llamaron a boicotear el SIMCE, Movimiento Alto al SIMCE y Educación 2020 que han desarrollado un análisis y propuestas referidas al SIMCE.

³ Términos de Referencia para el Equipo de Tarea.

Carrasco (Fundación Chile); Bárbara Eyzaguirre (Fundación Astoreca); Juan Eduardo García-Huidobro (Universidad Alberto Hurtado); Jorge Manzi (Director de Mide UC); Mariano Rosenzvaig (Director de Educación de la Municipalidad de Providencia); Guillermo Scherping (Asesor Colegio de Profesores); Ernesto Treviño (Director Ejecutivo del CPCE, Universidad Diego Portales); Xavier Vanni (investigador del CIAE, Universidad de Chile) y María José Zañartu (Belén Educa). En representación de la Subsecretaría de Educación participó Jaime Portales y en representación de la Agencia de la Calidad participaron Carlos Henríquez, su Secretario Ejecutivo, y Juan Bravo, Jefe de la División de Logros de Aprendizaje.

2. Metodología de trabajo del Equipo

Se realizaron veintidós sesiones de trabajo entre el 1 de agosto de 2014 y el 15 de enero de 2015. Para informar el diagnóstico a realizar, se destinaron inicialmente siete reuniones de audiencias para escuchar las experiencias y visiones de distintos actores respecto del Sistema Nacional de Evaluación de Aprendizajes. El Equipo de Tarea se reunió con profesores, directores de escuela, sostenedores (FIDE, CONACEP, Asociación Chilena de Municipalidades, Asociación Gremial de Corporaciones Municipales, Sociedad de Instrucción Primaria, Vicaría de la Educación), asociaciones de apoderados (AMDEPA, CORPADE)⁴, con el Colegio de Profesores, instituciones y movimientos sociales que habían desarrollado propuestas respecto de SIMCE (Educación 2020, Alto al SIMCE, UNICEF), y con investigadores y expertos, entre los cuales asistieron dos académicas que han recibido el Premio Nacional de Educación y que han tenido un rol destacado en relación al SIMCE⁵.

También se sostuvieron reuniones y se recibieron aportes del Consejo Nacional de Educación, del Consejo de la Agencia de Calidad y del Ministerio de Educación, específicamente de la Unidad de Curriculum, del Área de Educación Técnico Profesional y de la División de Educación General.

En la etapa siguiente, se destinaron dos reuniones a la revisión de la experiencia internacional con Daniel Koretz (investigador de Harvard quien ha estudiado distorsiones y efectos no deseados de mediciones con altas consecuencias en Estados Unidos) y Paulo Santiago (quien dirigió la revisión de la OCDE sobre sistemas nacionales de evaluación en 25 países, publicada en 2013). Asimismo, se revisaron los estudios y antecedentes nacionales referidos a la percepción, comprensión y uso del SIMCE en el país.

Finalmente, el Equipo de Tarea desarrolló sesiones temáticas de discusión y de revisión de borradores del diagnóstico y de las propuestas.

⁴ También fueron invitados la Asamblea Coordinadora de Estudiantes Secundarios (ACES) y a la Coordinadora Nacional de Estudiantes Secundarios (CONES), sin embargo no pudieron asistir a la sesión.

⁵ Erika Himmel, premio Nacional de Educación 2011, quien desarrolló el sistema nacional de evaluación de aprendizajes y luego presidió la primera Comisión de revisión del SIMCE en 2003, y Beatrice Avalos, premio Nacional de Educación 2013, quien lideró la firma de la carta abierta de investigadores y académicos en 2013 solicitando repensar el SIMCE y su aporte al sistema educativo. Otros investigadores que se recibieron en audiencia fueron Gregory Elacqua, Cristián Bellei, y Harald Beyer.

3. Resumen de audiencias con directivos, docentes y sostenedores

Sobre la base de las audiencias es posible identificar elementos positivos y negativos referidos al SIMCE en la visión de los establecimientos educacionales y sus sostenedores⁶.

Aspectos positivos:

Se considera un sistema de medición **técnicamente sólido**, en cuyos resultados se puede confiar. Prácticamente todos los actores invitados a las audiencias manifiestan que creen necesaria una evaluación de aprendizajes a nivel nacional, en particular para monitorear los logros y generar apoyos cuando se identifiquen debilidades. También se valora contar con información a nivel de cada establecimiento educacional que permita monitorear los resultados de aprendizaje en el tiempo, lo que resultaría inaccesible para la gran mayoría de las escuelas que no pueden contratar evaluaciones externas análogas al SIMCE.

Los datos que provee el SIMCE constituyen para los sostenedores que cuentan con la capacidad técnica suficiente para analizarlos, una ayuda para mejorar su gestión pedagógica y alinear los esfuerzos de las escuelas hacia el logro de los aprendizajes básicos. Estos sostenedores valoran una alta frecuencia de las mediciones externas.

También se valora **contar con un sistema de evaluación nacional de aprendizajes** que forme parte de un sistema más amplio de monitoreo de resultados y apoyo a la calidad educativa, en que se integran otras dimensiones de calidad, además de los logros cognitivos. El reporte de información sobre **otros indicadores de calidad** a nivel nacional en 2014, fue bien recibido, como un avance hacia una mirada más integral de la calidad educacional.

Asimismo, se valoran los **avances en la calidad de la información** que se entrega a los establecimientos educacionales, que desde 2007 permite conocer la proporción de estudiantes que alcanza diferentes **niveles de aprendizaje** que describen lo que los estudiantes saben y son capaces de hacer. En el caso de escuelas **con capacidades** para hacer aprovechamiento y uso de los datos, también se aprecia contar con información confiable de comparación con **escuelas similares o de mejores resultados, y la evolución de los resultados en el tiempo**.

Adicionalmente, se consideró positiva la incorporación de adecuaciones de acceso a la prueba para alumnos con necesidades educativas especiales, como es el caso de lo que ocurre con alumnos con discapacidad visual o pérdida auditiva parcial o total.

Las críticas:

Se considera que si bien los resultados de aprendizaje en las áreas evaluadas por el SIMCE son una parte muy relevante de la calidad de una escuela, no la abarcan completamente. Por tanto, el énfasis que la política educacional y los medios de comunicación han puesto en los resultados del SIMCE –de acuerdo a lo presentado en las audiencias– llevaría a un

⁶ Este informe ejecutivo se focaliza en las audiencias con actores del sistema escolar. Las audiencias con investigadores, organizaciones y movimientos sociales se integran a lo largo del texto.

estrechamiento **de la noción de “calidad educativa”** ya que los profesores y directivos perciben que los resultados en las asignaturas evaluadas es lo único que se valoraría. Se considera que el intenso calendario actual de pruebas genera una sobrecarga en las escuelas, preocupación expresada por la mayoría de los entrevistados.

Asimismo, se manifestó inquietud por las **consecuencias no previstas y no intencionadas** del sistema de evaluación de logros de aprendizaje. Se mencionó entre ellas, que las escuelas comienzan a focalizar su atención en las asignaturas evaluadas en desmedro de las restantes y destinan un tiempo excesivo a los ensayos para preparar el SIMCE. También se reportó que se utilizan los resultados del SIMCE para publicitar a las escuelas, lo que de acuerdo a algunas de las opiniones recogidas, incentivaría la competencia e inhibiría la cooperación entre ellas. Lo anterior se agudizaría cuando, a nivel local, se establecen premios o reconocimientos por parte de los sostenedores a escuelas con más alto SIMCE.

La **publicación y comparación de resultados** a nivel de escuelas es uno de los problemas frecuentemente mencionados. Hubo bastante coincidencia en la necesidad de revisar lo que publica la Agencia de Calidad y cómo lo hace, incluyendo los rankings. En las sesiones de audiencia, en los casos de escuelas o sostenedores con capacidades más precarias de gestión de la información, se destacó las consecuencias de **estigmatización** y consecuente fuga de matrícula asociada a la publicación de los resultados por escuela. Si bien algunos sostenedores valoraron contar con datos confiables y válidos que permitan comparaciones con escuelas de similares características socioeconómicas, esta comparación contextualizada de los resultados en general pareció ser poco visible para ellos, siendo más común las alusiones y rechazo a las comparaciones indebidas entre escuelas con muy diferentes recursos y composición social de sus estudiantes. En este sentido, se hizo notar que la comparación directa de resultados, que no consideran el nivel socioeconómico de los estudiantes, deriva en desánimo de los educadores y educadoras que trabajan en los sectores de mayor pobreza.

Finalmente, desde el punto de vista de los actores de las escuelas, aun es necesario mejorar la pertinencia de la información entregada, por ejemplo, reportar resultados por ejes de contenido o habilidad (ej. en Matemáticas, entregar información para Geometría, Datos y Azar, Álgebra, etc.), y liberar un mayor número de ítems de la prueba.

4. Diagnóstico del Equipo de Tarea

4.1 Síntesis de aspectos valorados por el Equipo de Tarea

A partir de la revisión de los antecedentes y estudios nacionales y de las presentaciones de invitados a sesiones de audiencia, el Equipo de Tarea valora positivamente variados aspectos del Sistema Nacional de Evaluación de Aprendizajes. A continuación se enumeran estos aspectos.

- a) Tener un Sistema Nacional de Evaluación de Aprendizajes ha permitido relevar la importancia de lograr aprendizajes fundamentales y generar un sentido de urgencia por mejorar.
- b) Una evaluación censal de logros de aprendizaje favorece políticas educacionales mejor informadas.
- c) Una evaluación censal permite dar información a cada establecimiento educacional sobre los logros de aprendizaje de sus alumnos en relación con referentes nacionales
- d) Se verifican importantes avances en materia de la calidad de la información provista a las escuelas.
- e) La incorporación del SIMCE a un sistema más completo de evaluación representa una oportunidad para lograr una evaluación integral de las escuelas.
- f) La existencia del SIMCE ha posibilitado desarrollar conocimiento e investigación educacional y evaluar la efectividad de políticas y programas.

A continuación se describe cada uno de los aspectos del SIMCE que son valorados en este diagnóstico:

- a) *Tener un Sistema Nacional de Evaluación de Aprendizajes ha permitido relevar la importancia de lograr aprendizajes fundamentales y generar un sentido de urgencia por mejorar.*

El Sistema Nacional de Evaluación de Aprendizajes ha permitido generar conciencia de la relevancia de que los estudiantes en el país, y en cada escuela, logren desarrollar competencias esenciales para integrarse y participar activamente en la sociedad.

Si bien el SIMCE no evalúa los aprendizajes en todas las asignaturas ni en todos los ámbitos involucrados en la formación integral de los estudiantes, es importante destacar que permite monitorear y poner foco en áreas que se consideran básicas y habilitantes para el aprendizaje en otras áreas y a lo largo de la vida.

La presencia de un cierto nivel de presión por mejorar logros en estos aprendizajes básicos, puede resultar favorable, cuando se acompaña de mecanismos de apoyo efectivos. Se sabe que los sistemas educativos y las escuelas están débilmente articulados y sobrecargados de iniciativas fragmentadas. En este contexto, hay mucho

espacio para la inercia, y para esfuerzos dispersos⁷. En este sentido, imprimir un sentido de urgencia por mejorar los niveles de aprendizaje en áreas que resultan básicas, favorece que las escuelas focalicen sus esfuerzos en forma sostenida.

b) Una evaluación censal de logros de aprendizaje favorece políticas educacionales mejor informadas.

Contar con información de cada establecimiento permite al Estado monitorear y dar cuenta del logro de aprendizajes básicos en cada uno de ellos. Esto favorece la responsabilización del Estado, que debiera proveer condiciones y dar garantías de acceso a una educación de calidad en cada escuela. También permite al Estado generar estrategias diferenciadas de apoyo a las escuelas y a los estudiantes; evaluar los efectos de sus políticas y programas en diferentes tipos de escuelas; e identificar buenas prácticas e innovaciones.

c) Una evaluación censal permite dar información a cada establecimiento educacional sobre los logros de aprendizaje de sus alumnos en relación con referentes nacionales

Contar con una evaluación censal, permite a cada escuela situar los logros de aprendizaje de sus alumnos en un contexto más amplio y utilizarla para la gestión curricular. Esto, además, contribuye a promover un compromiso profesional responsable de los docentes, directivos y sostenedores a cargo de los establecimientos educacionales.

La información provista ofrece la oportunidad de que -en las áreas evaluadas por el SIMCE- las escuelas diseñen, ajusten y evalúen sus estrategias a partir de datos y evidencia. La comparabilidad de las pruebas entre años les otorga información sobre su tendencia en el tiempo, mientras que la comparación con resultados de escuelas de similares características y estudiantes les permite aproximarse a una evaluación de cuán efectivas están siendo, en los aspectos que mide el SIMCE. Asimismo, los datos sobre la proporción de estudiantes que alcanzan los estándares nacionales, aporta información sobre el grado en que en cada establecimiento se están logrando las expectativas de aprendizaje establecidas en el currículum para las áreas evaluadas. Además, la distribución de los estudiantes en los diferentes niveles de aprendizaje aporta información sobre la equidad interna alcanzada, ya que evidencia la diversidad de logros al interior de las escuelas.

d) Se verifican importantes avances en materia de la calidad de la información provista a las escuelas.

Durante los últimos siete años, es posible constatar los siguientes avances en la información provista por el SIMCE:

⁷ Fullan (2010).

- La inclusión de reportes por nivel de aprendizaje, que desde 2007 describen los conocimientos y habilidades demostradas por los estudiantes en cada uno de ellos. Esto permite otorgar significado pedagógico a los resultados entregados por el SIMCE.
- Mayores apoyos y orientaciones para el procesamiento de los reportes de resultados por parte de las escuelas, a través de la provisión de materiales (guías para el desarrollo de talleres) y del tiempo para realizar jornadas de análisis de los resultados en los establecimientos⁸.

e) La incorporación del SIMCE a un sistema más completo de evaluación representa una oportunidad para lograr una evaluación integral de las escuelas.

La incorporación del SIMCE a un sistema de evaluación de las escuelas que comienza a integrar sus procesos y resultados en otros ámbitos de la formación de los estudiantes (como los Otros Indicadores de Calidad), constituye un avance significativo y una oportunidad para dar cuenta de su quehacer de un modo más completo. La instalación de la Agencia de Calidad en que se aloja actualmente el SIMCE y el desarrollo de metodologías y equipos profesionales para evaluar la gestión y procesos pedagógicos y formativos en las escuelas, a través de las visitas de evaluación y orientación, se orientan en esta dirección.

f) La existencia del SIMCE ha posibilitado desarrollar conocimiento e investigación educacional y evaluar la efectividad de políticas y programas.

Los resultados del SIMCE se han utilizado para una enorme variedad de estudios en educación, permitiendo, por ejemplo, conocer el impacto de intervenciones o programas tanto acotados como a nivel nacional (como P900, LEM y PAC), visibilizar temas tales como inequidad de género en los logros de aprendizaje en determinadas asignaturas, identificar fortalezas y debilidades en el sistema escolar y factores que inciden en los resultados de aprendizaje, etc. Solo en el año 2014, el Ministerio de Educación recibió 92 solicitudes de las bases de datos de resultados SIMCE para ser utilizadas en investigación, lo que ilustra el intensivo uso de sus resultados para producir conocimiento sobre el sistema educacional.

⁸ Entre 2005 y 2009 se suspendían las clases por un día para realizar estas jornadas de análisis. A partir de 2010 se hicieron voluntarias, dependiendo su realización de las definiciones del sostenedor.

4.2. Los principales problemas identificados por el Equipo de Tarea

Es importante destacar que los mayores problemas y desafíos identificados por el Equipo de Tarea se refieren a los usos dados al SIMCE por la política educacional, por los medios de comunicación y las escuelas y a los efectos de estos usos, más que al tipo de instrumentos o información entregada. Asimismo, se identifican problemas en el Sistema de Aseguramiento de la Calidad que determina el uso que se da al SIMCE.

En relación al Sistema Nacional de Aseguramiento de la Calidad el principal problema identificado es que:

- a) Existe un desequilibrio entre la expansión del sistema de evaluación y la presión hacia las escuelas asociada a este, por una parte, y la debilidad de los mecanismos y políticas para apoyarlas y desarrollar sus capacidades, por otra.

Los principales problemas asociados al SIMCE que fueron identificados, son los siguientes:

- b) Focalizar excesivamente la atención en los resultados en mediciones externas tipo SIMCE restringe la visión sobre la calidad de la educación.
- c) La publicación de rankings basados en puntajes que, entre otras, no consideran las características socioeconómicas de los estudiantes, induce a interpretaciones erróneas y se traduce en comparaciones injustas para las escuelas que trabajan en contextos vulnerables.
- d) Los usos y efectos, tanto deseados como indebidos de la medición de aprendizajes no se explicitan, no se estudian sistemáticamente ni se toman medidas respecto de ellos.
- e) Hay más pruebas e información disponible que preparación y oportunidades en las escuelas para aprovecharlas.
- f) La evaluación realizada por los profesores en las salas de clase ha sido, hasta ahora, invisible para el sistema nacional de evaluación de aprendizajes y desatendida por las políticas educativas.

Por último, el Equipo de Tarea identifica tanto riesgos como oportunidades asociadas a la Ordenación de Escuelas.

A continuación se explica cada uno de los problemas identificados.

- a) *Existe un desequilibrio entre la expansión del sistema de evaluación y la presión hacia las escuelas asociada a este, por una parte, y la debilidad de los mecanismos y políticas para apoyarlas y desarrollar sus capacidades, por otra.*

El Equipo de Tarea constata un claro contraste entre la expansión y evolución que ha experimentado la evaluación y la información disponible sobre sus resultados, y el débil desarrollo de los sistemas de apoyo y aseguramiento de las condiciones para que las escuelas puedan desplegar un trabajo de calidad. Este desequilibrio representa una seria limitación para el desarrollo de procesos de mejoramiento en las escuelas y un riesgo para la validación y eficacia del actual Sistema de Aseguramiento de la Calidad de la Educación.

Las siguientes son manifestaciones de este desequilibrio entre presión y apoyo:

- **Se han incrementado los incentivos y las consecuencias** que tanto la **política educacional** como los sostenedores y directivos a nivel local han asociado a los resultados del SIMCE durante las dos últimas décadas. Destacan los incentivos salariales colectivos por mejoramiento de resultados a través del “*Sistema Nacional de Evaluación del Desempeño de los Establecimientos Educacionales Subvencionados*” (SNED), las metas de los convenios de desempeño de los directores municipales, la categorización de escuelas en la Ley de Subvención Escolar Preferencial y en el nuevo Sistema de Aseguramiento de la Calidad que asigna consecuencias que van desde la frecuencia de las visitas evaluativas, a la posible revocación del reconocimiento oficial de las escuelas si no presentan mejorías dentro del plazo establecido.
Por su parte, algunos directivos y sostenedores utilizan los resultados para asignar incentivos individuales a los profesores (a pesar de que esta práctica no tiene fundamento técnico)⁹, y ofrecen premios o actividades recreativas a los estudiantes para motivarlos a rendir mejor en las pruebas SIMCE¹⁰.
- **Se ha expandido la medición y fortalecido la fiscalización sin un desarrollo equivalente de los mecanismos de apoyo.** El aumento de la frecuencia de las mediciones externas de aprendizaje y la progresiva inclusión de más asignaturas en ellas, se ha realizado a un ritmo acelerado. En un periodo de cinco años se ha duplicado el número de niveles evaluados (de tres a seis) y se han incrementado las pruebas censales administradas anualmente de ocho a quince pruebas¹¹.

Por otra parte, el nivel de desarrollo y fortalecimiento de las unidades ministeriales a cargo de asegurar el apoyo y asistencia técnica a las escuelas no presenta el mismo

⁹ Este antecedente se obtuvo a través de las Audiencias del Equipo de Tarea y es coincidente con los hallazgos de Manzi et al (2014). De acuerdo a este estudio, los directivos perciben que uno de los objetivos del SIMCE es evaluar a los docentes.

¹⁰ De acuerdo a estudio de Flórez (2014).

¹¹ De acuerdo a la fundamentación entregada por el Ministerio de Educación al Consejo Nacional de Evaluación para la aprobación del calendario de pruebas, este incremento buscaba generar más información para hacer más precisa la ordenación de escuelas, proveer condiciones para desarrollar valor agregado y evitar el estrechamiento curricular al incluir más asignaturas (Inglés, Educación Física) en la medición.

dinamismo que ha caracterizado la formación de las instituciones a cargo de la fiscalización y evaluación (Superintendencia y Agencia de la Calidad, respectivamente). Además es posible constatar que la Ley que regula el Sistema de Aseguramiento de la Calidad es mucho más específica y vinculante en los aspectos referidos a la *responsabilización* de las escuelas que en lo que se refiere al diseño del sistema de apoyo para que estas mejoren.

- **Los programas de apoyo a las escuelas implementados a la fecha muestran una efectividad limitada.** Las evaluaciones de los programas de apoyo promovidos por el Ministerio de Educación durante las últimas décadas y los programas de asistencia técnica externa (ATE) muestran que su efectividad ha sido moderada. En efecto, estos no han resultado suficientemente efectivos en el mejoramiento de resultados de aprendizaje evaluados por el SIMCE. Esto debilita las expectativas de que el ritmo de mejoramiento de resultados de las escuelas pueda ser el requerido por la Ley de Aseguramiento de la Calidad, la que actualmente establece un plazo de 4 años para que las escuelas en la categoría más baja de desempeño salgan de ella.
- **Las políticas referidas a docentes y directivos aun no ofrecen las condiciones para asegurar la disponibilidad y retención de profesores y directivos suficientemente preparados en todas las escuelas.** Si bien las políticas referidas a docentes y directivos han mejorado crecientemente sus condiciones y propiciado su formación permanente, ellas no han sido suficientes. Las actuales condiciones de la carrera docente, la escasez de tiempo no lectivo, el bajo nivel de remuneraciones de los profesores, la insuficiente regulación de la formación inicial, y la inequidad de la distribución de los docentes en el sistema, no permiten asegurar la disponibilidad de profesionales suficientemente preparados en las escuelas, en especial en aquellas más vulnerables. Si bien las inminentes modificaciones a la carrera docente buscarán cambiar esta situación, dicho cambio será gradual, al igual que el impacto que este tenga.

b) Focalizar excesivamente la atención en los resultados en pruebas nacionales restringe la visión sobre la calidad de la educación

El Ministerio de Educación, la Superintendencia y la Agencia de Calidad, recaban información relevante y muy variada sobre cada escuela. Además de la valiosa información sobre los logros de aprendizaje que alcanzan los estudiantes en las pruebas nacionales, se genera información sobre matrícula y asistencia de estudiantes, retención y promoción escolar, dotación y rotación docente, información sobre los planes de mejoramiento de cada establecimiento educacional, grado de ajuste a la normativa en aspectos tales como financiamiento, convivencia escolar, respeto a la ley de violencia escolar, equidad de género en los resultados de aprendizaje, calidad de la docencia y de la gestión escolar, entre otros. Sin embargo, dicha mayor disponibilidad de información no necesariamente se traduce en que se transmita una imagen completa e integral de la calidad de la educación a los distintos actores sociales. Una selección cuidadosa de algunos de estos datos permitiría conformar y comunicar un panorama más amplio y completo sobre la calidad y equidad de la educación, tanto en lo que se reporta a nivel de país, como a nivel de cada establecimiento.

A juicio del Equipo de Tarea, los siguientes elementos, entre otros, han contribuido a que se focalice la atención de modo excesivo en los resultados del SIMCE y a que se los interprete equivocadamente como sinónimo único de calidad:

- La amplia difusión de los resultados por los medios de comunicación masiva ha favorecido que la opinión pública centre la atención en éstos como indicadores únicos de la calidad de la educación brindada por cada establecimiento escolar y por el sistema escolar en su conjunto.
- El peso mayoritario que tienen los resultados de SIMCE en el índice de ordenación de escuelas, que según la Ley de Aseguramiento de la Calidad debe asignar al logro de estándares de aprendizaje al menos el 67%. La actual metodología, sujeta a marcha blanca, lleva a que los resultados SIMCE pesen entre el 73% y 77% del índice usado para la ordenación.
- La denominación de la evaluación nacional de aprendizajes (SIMCE: Sistema de Medición de la *Calidad* de la Educación) induce a pensar que los resultados en sus pruebas sintetizan y abarcan la calidad educativa de un establecimiento¹². Esto resulta particularmente confuso en la actualidad, en que este ‘Sistema de Medición’ se ha integrado a un sistema de evaluación mayor, encargado a la Agencia de Calidad, y será complementado con resultados de los Otros Indicadores de Calidad (convivencia escolar, participación, retención, equidad de género, entre otros), y con las visitas evaluativas que examinarán los procesos pedagógicos y de gestión al interior de las escuelas según los estándares indicativos de desempeño.
- La concentración de la medición en algunas áreas del currículum nacional y en las dimensiones evaluables confiablemente a través de pruebas de lápiz y papel y predominantemente de selección múltiple, también contribuye a una valoración mayor de estas por sobre las demás. En el caso de la Educación Media, además es necesario destacar la concentración de la evaluación de logros de aprendizaje en áreas curriculares de la educación general y científico-humanista, y la omisión de evaluaciones de aprendizaje en las áreas vinculadas a la educación técnico profesional.
- La baja relevancia que se ha dado a los procesos internos de las escuelas si se compara con la relevancia otorgada a la medición de logros de aprendizaje. Los resultados de las visitas evaluativas tienen menos consecuencias e incentivos asociados (por ejemplo no inciden en la categorización de las escuelas según la Ley de Aseguramiento de la Calidad) y solo recientemente se han comenzado a implementar.

De acuerdo a las audiencias del Equipo de Tarea, a veces este foco prioritario en puntajes SIMCE como indicador único de calidad tiene como consecuencia que en los establecimientos escolares se asignan mayores recursos y tiempo a preparar para lo que miden esas pruebas, provocando una reducción de las oportunidades de aprendizaje en otras áreas que no son medidas.

¹² El significado de la sigla corresponde a ‘Sistema de Medición de la Calidad de la Educación’, aun cuando este significado no se use en los documentos oficiales desde 2005.

- c) *La publicación de rankings basados en puntajes que, entre otras, no consideran las características socioeconómicas de los estudiantes, induce a interpretaciones equivocadas y se traduce en comparaciones injustas para las escuelas que trabajan en contextos vulnerables.*

La difusión pública de los resultados por escuela comenzó en 1995 dando cumplimiento a la Ley Orgánica Constitucional que así lo establecía. Desde entonces ha sido frecuente la comparación directa de puntajes entre las escuelas, sin importar las características socioeconómicas de la población escolar con la que trabajan. Este tipo de comparaciones adquirió más visibilidad en 2011 al publicarse el mapa de ‘semáforos’ basado en resultados SIMCE que tampoco consideraban el contexto socioeconómico de las escuelas¹³.

Los medios de comunicación de masas habitualmente construyen rankings o listados de establecimientos jerarquizándolos según el puntaje promedio obtenido en una sola medición. Estas listas o rankings de escuelas propician la interpretación errónea de que diferencias menores entre puntajes son diferencias relevantes (por ejemplo, que un colegio que está en el lugar 15 tiene mejores resultados que uno que está en el lugar 16, sin importar la diferencia entre sus puntajes).

Los puntajes entregan información sobre lo que los alumnos saben y pueden hacer, pero resultan engañosos si se interpreta que los que obtienen puntajes más altos están haciendo una mejor labor educativa, ya que se estimula así la falsa idea de que establecimientos de mejor composición socioeconómica son más efectivos, y viceversa. Esto no necesariamente es así, ya que los altos puntajes pueden deberse a las ventajas socioeconómicas de los estudiantes con que la escuela trabaja o a la selección previa de los que presentaban mayor potencial de buen rendimiento académico. En consecuencia, los rankings y comparaciones basadas en puntajes promedio pueden experimentarse como parciales e injustos, ya que contribuyen a prestigiar escuelas en contextos de alto capital cultural y socioeconómico y derivan en estigmatización de las escuelas que trabajan en contextos más vulnerables.

- d) *Los usos y efectos intencionados e indebidos de la medición de aprendizajes no se explicitan, no se estudian sistemáticamente, ni se toman medidas respecto de ellos.*

El Sistema Nacional de Evaluación de Aprendizajes no hace explícito de qué modo espera que los diferentes actores usen la información y actúen frente a la medición externa, es decir, no se establece públicamente de qué manera y a través de qué decisiones de los involucrados, se espera que la medición contribuya al mejoramiento de los aprendizajes. En otras palabras, no se clarifica cuál es la hipótesis sobre los cambios que se espera promueva la evaluación, o la teoría de la acción que le subyace. Tampoco se explicita cuáles serían usos indebidos e interpretaciones equivocadas o injustas de la información. ¿Qué decisiones se espera que adopten los directivos de un establecimiento al identificar un área curricular sistemáticamente más débil? ¿Se espera que los apoderados cambien a sus hijos de establecimiento al conocer los resultados del SIMCE o que se involucren para contribuir a su mejoramiento? ¿Qué se busca que

¹³ Aunque no es en rigor un ranking, en este mapa se asignaba un color verde a las escuelas de más altos puntajes, amarillo a las de puntajes intermedios y rojo a las de más bajos puntajes promedio, induciendo a concluir que las escuelas categorizadas en verde desempeñaban una “mejor” labor que las amarillas, y estas eran a su vez “mejores” que las rojas.

hagan los profesores con los datos sobre los errores más frecuentes de los alumnos en determinados contenidos? ¿Qué decisiones se espera que se adopten a nivel de política educacional?

En algunos casos, la Agencia de Calidad sí entrega información relativa a los cuidados que deben tenerse al usar la información como, por ejemplo, cuáles son las comparaciones que no se deben hacer (entre puntajes de distintas asignaturas) o cuándo interpretar un alza de puntaje como significativa. Sin embargo, en otros muy relevantes esto no se hace. Por ejemplo, no se establece explícitamente que responsabilizar y tomar decisiones respecto de profesores individuales por el resultado de un curso (recompensándolos o sancionándolos), no tiene fundamento alguno¹⁴. Tampoco se previene sobre la inconveniencia de presionar a los estudiantes de los cursos que rinden SIMCE por los resultados en las pruebas.

Es necesario establecer claramente cuáles son los usos y acciones que se espera motivar con la medición de aprendizajes, lo que permite luego ‘evaluar los efectos de la evaluación’ y verificar si los supuestos se cumplen, o si, por el contrario, predominan los efectos no intencionados. El Equipo de Tarea pudo recoger a través de las audiencias tanto experiencias positivas de análisis y uso de resultados del SIMCE, como testimonios de experiencias y preocupación respecto de los efectos y usos indeseados de la medición. A pesar de los 27 años de existencia del sistema de medición de aprendizajes en Chile, y de la recomendación explícita de la Comisión que revisó el SIMCE en 2003¹⁵, los estudios sobre sus usos y consecuencias son incipientes y no permiten establecer la magnitud y frecuencia con que se presentan los problemas identificados.

Tampoco se observa que las instituciones correspondientes (Ministerio de Educación, Agencia de Calidad y Superintendencia) asuman un papel activo en educar a la ciudadanía sobre los usos indebidos o no intencionados del SIMCE y ni en establecer medidas respecto de ellos. Esto es muy relevante debido a que la investigación a nivel internacional en Estados Unidos e Inglaterra, indica que frecuentemente las mediciones que tienen altas consecuencias para las escuelas producen efectos indeseados que es necesario monitorear y evitar. Por ejemplo, comportamientos estratégicos de corto plazo para elevar los puntajes en las pruebas sin mejorar los procesos educativos, o el estrechamiento curricular.

e) *Hay más pruebas e información disponible que preparación y oportunidades en las escuelas para aprovecharlas*

El Equipo de Tarea pudo constatar que los informes que SIMCE entrega a directivos y docentes sobre los resultados de sus establecimientos son bastante ricos y pormenorizados e incluyen orientaciones y guías para realizar talleres de análisis de los datos. Asimismo, se pudo constatar que, dado el incremento de mediciones SIMCE durante los últimos años, los establecimientos reciben actualmente hasta cinco o seis informes diferentes (para 2°, 4°, 6° y 8° básico, para II y

¹⁴ Esto debido a que son varios los profesores que han contribuido al nivel alcanzado por el curso y a que tampoco se cuenta con medidas de progreso o valor agregado para cada año y asignatura.

¹⁵ “La Comisión plantea la necesidad de considerar estudios periódicos que permitan evaluar el impacto de las medidas implementadas, de manera de establecer la conveniencia de la permanencia de las mismas, su modificación o intensificación. Este tipo de evaluaciones debe considerar (...) eventuales consecuencias no deseadas” (p.84)

III medio). A través de las audiencias y los estudios nacionales revisados¹⁶, el Equipo de Tarea pudo constatar una gran variabilidad en las capacidades instaladas en escuelas y sostenedores para analizar y sacar provecho de la información aportada por el SIMCE, y limitaciones de tiempo para realizar jornadas colectivas de análisis orientadas al buen uso de ella. En las audiencias del Equipo de Tarea se hizo evidente que aquellas escuelas con capacidades técnicas para hacerlo, analizan y utilizan los datos para hacer un seguimiento de sus resultados y tomar decisiones, mientras que en las escuelas en que se carece de estas capacidades, el SIMCE no se experimenta como una fuente de información y se vive más bien como un elemento de agobio.

- f) *La evaluación realizada por los profesores en las salas de clase ha sido, hasta ahora, invisible para el Sistema Nacional de Evaluación de Aprendizajes y desatendida por las políticas educativas.*

Por sus características y propósitos, la información que entregan mediciones externas estandarizadas como SIMCE resulta más apropiada para que los docentes y directivos hagan un balance periódico y análisis global de los resultados de la escuela, que para informar cotidianamente decisiones pedagógicas puntuales respecto de los alumnos específicos que rindieron la prueba. En este sentido, la evaluación externa no puede sustituir la evaluación de aprendizajes realizada por los profesores en las escuelas.

La evaluación interna realizada por los profesores resulta vital para un sistema nacional de evaluación de aprendizajes, ya que tiene impacto cotidiano para los estudiantes, marca las expectativas respecto de su desempeño, determina decisiones como la promoción, e incide en el proceso de selección para la educación superior. Adicionalmente, la investigación educacional ha mostrado que los cursos cuyos profesores utilizan evaluaciones sumativas de mayor calidad y amplitud, pueden mejorar la motivación de los estudiantes y potenciar el desarrollo de habilidades cognitivas más complejas¹⁷. Asimismo, usar la evaluación formativamente, para dar retroalimentación a los estudiantes y para modificar la enseñanza, es una de las prácticas pedagógicas con mayor efecto positivo sobre los aprendizajes¹⁸.

Sin embargo, a pesar de su relevancia, en Chile la evaluación de aprendizajes que realizan los profesores parece ser una pieza invisible y desestimada por las políticas educacionales si se contrasta con la importancia que se le otorga a la evaluación externa (SIMCE). La evaluación interna y la medición externa de logros de aprendizajes deben articularse y complementarse. Los resultados en mediciones externas favorecen que las escuelas y profesores ajusten sus expectativas según los estándares nacionales a la hora de evaluar cotidianamente a sus alumnos, y las evaluaciones internas en las escuelas permiten un seguimiento permanente de los aprendizajes de los estudiantes, como también cubrir áreas y habilidades que no pueden ser evaluadas a través de mediciones masivas.

¹⁶ Según el estudio de Manzi et al. (2014), los directivos declaran inseguridades respecto de la interpretación de los datos y limitaciones de tiempo para realizar jornadas de análisis de ellos.

¹⁷ Jensen, McDaniel, Woodard & Kummer (2014)

¹⁸ Black & William (1998); Hattie (2003)

Varios son los síntomas de debilidad y de invisibilidad de la evaluación de aprendizajes realizada por las escuelas y por los docentes. De acuerdo a los resultados reportados persistentemente por el sistema de Evaluación Docente del sector municipal y la evaluación para la asignación de excelencia pedagógica (AEP), la evaluación de aprendizajes es uno de los ámbitos del desempeño de los profesores que más requiere mejorar¹⁹, y de los más abandonados por las políticas y orientaciones entregadas por el Ministerio de Educación; además se verifica una débil conexión entre la evaluación externa estandarizada y sus resultados, y la evaluación realizada por las escuelas.

Riesgos y oportunidades de la Ordenación de escuelas

La Ley de Aseguramiento de la Calidad establece que las escuelas deben ser clasificadas u ordenadas en cuatro categorías de desempeño según sus resultados referidos a los estándares de aprendizaje y a Otros Indicadores de Calidad educativa, considerando las características socioeconómicas de su alumnado. Desde diciembre de 2014, la Agencia de Calidad ha comenzado a implementar en marcha blanca esta ordenación, según recomendación del Consejo Nacional de Educación. A través de la marcha blanca se podrá estudiar, antes de entrar en régimen, cómo reciben y entienden los establecimientos educacionales esta categorización y qué efectos desencadena en las escuelas. La categoría asignada a cada establecimiento no será de conocimiento público en esta etapa.

Dado que la marcha blanca ha comenzado muy recientemente, el Equipo de Tarea no cuenta aún con una evaluación de ella. Sin embargo, como la Ordenación es uno de los usos relevantes de la medición nacional de aprendizajes²⁰, y existen en el Equipo de Tarea distintas apreciaciones sobre los riesgos y oportunidades que abre, se aborda este punto a continuación. En este caso, no solo se incluye las apreciaciones sobre riesgos y oportunidades, también se adelantan las recomendaciones, algunas de las cuales se propone sean probadas en la marcha blanca.

La Ordenación toma en cuenta las características socioeconómicas de los estudiantes de las escuelas, por lo que el Equipo de Tarea estima que esta representa un avance respecto a comunicar solamente los puntajes promedio del SIMCE, lo que, como ya se ha señalado, induce a interpretaciones erróneas sobre la efectividad de ellas. Un reporte de resultados basado en la Ordenación comunicaría en qué medida los establecimientos están logrando niveles de resultados superiores, equivalentes, inferiores o muy inferiores a escuelas de similares características socioeconómicas. También se valora que todas las escuelas del país reciban información sobre su aporte a los resultados de sus estudiantes, para movilizar a todas ellas a superarse.

Sin embargo, se advierten algunos riesgos asociados a la ordenación que es necesario prevenir, y respecto de los cuales el Equipo de Tarea ofrece recomendaciones específicas.

¹⁹ Manzi, González y Sun (2011)

²⁰ El 67% del índice de calidad corresponde a estándares de aprendizaje evaluados por el SIMCE y la metodología de ordenación actual incluye el puntaje SIMCE en la categoría de otros indicadores de calidad llevando a que el peso del SIMCE en el índice de calidad sea entre 73% y 77 % si se considera equidad de género.

En primer lugar, en la medida que la Ordenación es una clasificación pública de las escuelas realizada por el Estado, implica una rotulación oficial de ellas (*‘Escuela de Desempeño Alto’, ‘Escuela de Desempeño Insuficiente’,* etc.). Esto conlleva el riesgo de acentuar la desmoralización precisamente de las escuelas que más necesitan mejorar y de intensificar las consecuencias no deseadas que se han descrito anteriormente.

Para potenciar las oportunidades que abre la ordenación y evitar sus riesgos, el Equipo de Tarea recomienda que se comunique la información sobre la situación de la escuela evitando el uso público de rótulos que poco aportan para un proceso de mejora. Por ejemplo, en lugar de *‘Escuela de Desempeño Insuficiente’* se debiera comunicar su significado: *‘sus resultados son marcadamente más bajos que los de escuelas similares’*, y *‘Escuela de desempeño Alto’* debiera sustituirse por *‘Sus resultados son muy superiores a los de escuelas de similares características socioeconómicas’* (*).

Adicionalmente, se recomienda que para las escuelas situadas en las categorías más bajas de desempeño, esta descripción se acompañe de los compromisos y acciones de mejoramiento que tanto estas como sus sostenedores y el Ministerio de Educación, han comprometido o ya han emprendido. Por ejemplo, es posible indicar en estos casos que *“En esta escuela se ha iniciado un programa de mejoramiento en el cual se ha comprometido el sostenedor, la escuela y el Ministerio de Educación”*, poniendo así el énfasis en los usos de la evaluación para el mejoramiento más que en el señalamiento público de las escuelas.

El Equipo de Tarea, valora que se consideren los Otros Indicadores de Calidad (OIC), ya que esto constituye un avance al destacar otro tipo de resultados²¹ de los establecimientos educacionales, y dar una señal de que la evaluación de su calidad es integral. Sin embargo, se manifestaron diferencias respecto de los beneficios de incorporar los otros indicadores en un índice único de calidad. Algunos integrantes²² apoyan la inclusión de los Otros Indicadores en el índice único de calidad que se utiliza para la ordenación de escuelas, considerando que un índice global facilita identificar a aquellas que necesitan ayuda prioritaria y tomar decisiones sobre medidas tales como su re-estructuración. Desde esta perspectiva, es de todos modos imprescindible monitorear si los datos recogidos para los otros indicadores conservan su validez en el tiempo y entregar los resultados de todos los indicadores desagregados para asegurar una adecuada comprensión de las diferentes fortalezas y debilidades de los establecimientos escolares.

En cambio, otros integrantes²³ consideran que la inclusión de los otros indicadores en un solo índice de calidad tiene varios problemas. Paradojalmente, invisibiliza las otras dimensiones de calidad consideradas, al subsumirlas en un solo índice que pierde significado al referirse a una

²¹ Como clima de convivencia escolar, participación y formación ciudadana, equidad de género, etc.

²² Bárbara Eyzaguirre, Jorge Manzi, Xavier Vanni y María José Zañartu.

(*) Recomendación con implicancias de modificación legal.

²³ Juan Eduardo García Huidobro, Paulina Bustos, Guillermo Scherping, Mariano Rosenzvaig, Rafael Carrasco, Alejandro Carrasco y Lorena Meckes. Jorge Manzi y Xavier Vanni comparten los cuestionamientos metodológicos que se expresan acerca de los otros indicadores de calidad, pero consideran necesario el cálculo de un índice global en la medida que el sistema de aseguramiento de la calidad requiera una ordenación de las escuelas y contar con un índice que permita tomar decisiones al respecto de ellas.

mezcla de factores. Además, se propone una visión única sobre lo que compone la calidad educacional y sobre el peso que deben tener los diferentes elementos que la conforman. Esto resulta problemático porque los OIC tienen una ponderación muy limitada en la actual metodología de ordenación (entre 23% y 27%)²⁴, por lo que resulta cuestionable que este índice compuesto logre reflejar integralmente el desempeño de las escuelas. Al ser medidos los OIC a través de cuestionarios de percepción y opinión, es probable que, en un escenario en que se asocien consecuencias a los resultados en ellos, su validez y confiabilidad se reduzca, como también la variabilidad²⁵ en sus respuestas. Por lo mismo, tampoco resulta recomendable elevar esta ponderación. Una categorización del desempeño de las escuelas basada en este índice combinado, se traduciría en un juicio global sobre la calidad de ellas, acentuando el riesgo de estigmatización de las que queden clasificadas en la categoría de más bajo desempeño. Desde esta perspectiva, se recomienda estudiar una modificación de la metodología de ordenación y separar los resultados de aprendizaje medidos por el SIMCE de los ‘Otros Indicadores de Calidad’ (*). Esto podría dar origen no solo a una sino a diversas categorizaciones de las escuelas, según la dimensión de calidad que se considere. Ello permitiría hacer más visibles los OIC, evitando concentrar la información en un solo índice con ponderaciones de sus componentes que son necesariamente arbitrarias.

El Equipo de Tarea recomienda utilizar la marcha blanca para estudiar la metodología planteada en el decreto de ley que la establece y analizar en qué medida incluir los Otros Indicadores de Calidad hace o no una diferencia importante en la categorización de las escuelas si se la compara con una clasificación basada únicamente en los resultados de aprendizaje. Así también, la marcha blanca debiera monitorear el significado que le asignan los actores escolares a diferentes formas de reportar los resultados de la Ordenación, como también su comprensión y usos potenciales.

Es necesario precisar que parte del Equipo de Tarea²⁶ propone no hacer una ordenación de todas las escuelas (*), y distinguir sólo la proporción de ellas que en cada región presente mayores dificultades para lograr que sus estudiantes alcancen niveles básicos de aprendizaje y que por lo tanto requieran medidas especiales de apoyo.

²⁴ Estos porcentajes corresponden a la ponderación nominal, Si se considera los indicadores de equidad de género, que también usa los puntajes SIMCE para su cálculo, el peso nominal del SIMCE es de 77% y el de los OIC un 23%. Si además se considera la variabilidad de cada indicador y el rango explicativo dentro del índice de calidad, es muy probable que el peso efectivo del SIMCE supere el 85%.

²⁵ Esto es importante porque si se reduce la variabilidad de respuestas, el peso efectivo de los otros indicadores será aun menor.

(*) Recomendación con implicancias de modificación legal.

²⁶ Juan Eduardo García Huidobro, Mariano Mariano Rosenzvaig, Guillermo Scherping.

(*) Recomendación con implicancias de modificación legal.

Aspectos metodológicos de la Ordenación que es necesario estudiar

Después de cerca de dos años de desarrollo, la metodología para categorizar a los establecimientos educacionales fue aprobada por el Consejo de la Agencia de la Calidad y fue sometida a revisión del Consejo Nacional de Educación. Si bien la metodología de ordenación atravesó por un extenso proceso de revisión y cumplió con los procedimientos que establece la Ley, dada su relevancia, el Equipo de Tarea la revisó con el propósito de identificar los aspectos que aun debieran ser objeto de mayor estudio. Además de lo que ya se mencionó respecto de la medición de los OIC y el monitoreo de su validez, entre otros aspectos, se sugiere estudiar lo siguiente:

- a) *La inclusión de filtros.* En la actual metodología, se mezcla en la misma categoría de desempeño, escuelas clasificadas en ella por su efectividad con otras incluidas en dicha categoría por sus resultados absolutos. Esto sucede en las categorías extremas de la ordenación, por la consideración de ‘filtros’ que dejan en la categoría más alta a escuelas de elevado SIMCE, sin importar su efectividad y en la categoría más baja a las que tienen resultados absolutos muy bajos, también sin considerar su efectividad. En ambos casos, a las escuelas se les reporta que su ‘desempeño’ es *Alto* o *Insuficiente*, aunque estén en esa categoría por sus resultados y no por su aporte a ellos.
- b) *La inestabilidad de la categorización* y su actual sensibilidad ante variaciones en el número de mediciones que rinden las escuelas (la categorización de escuelas es muy sensible al número de niveles medidos que se incluyan en el índice).
- c) *Las implicancias de ponderar la distribución de los estudiantes* de las escuelas en los niveles de aprendizaje de los estándares y otorgarle a dicha ponderación un peso mayoritario en la ordenación. En particular, interesaría verificar si esto genera una atención exclusiva de las escuelas en el grupo de estudiantes de más bajos resultados en desmedro de los demás, o si por el contrario, estas dirigen sus esfuerzos a los estudiantes en todos los niveles de logro.

5. Recomendaciones

Respecto de los propósitos del sistema nacional de evaluación de aprendizajes, el Equipo de Tarea establece propósitos de responsabilización y de desarrollo o de mejoramiento para las evaluaciones externas y propósitos formativos y sumativos para la evaluación de aprendizajes que realizan internamente los profesores en las escuelas.

Respecto de las mediciones externas, el Equipo de Tarea considera que es necesario mantener los siguientes aspectos del SIMCE:

- a) Las evaluaciones censales de los logros de aprendizaje en cada establecimiento escolar.
- b) El acceso público a información sobre resultados de la evaluación en diferentes ámbitos y considerando el contexto de las escuelas.
- c) El rigor técnico y preocupación permanente por mejorar los instrumentos y procedimientos de evaluación.
- d) La entrega de información útil para la gestión curricular de las escuelas.
- e) Las acomodaciones para evaluar estudiantes con necesidades educativas especiales permanentes.

Los aspectos que se necesita cambiar o mejorar son:

- a) Diseñar e implementar una política amplia que permita potenciar la evaluación de aprendizajes que realizan los profesores y directivos al interior de las escuelas y complementarla con las mediciones externas.
- b) Equilibrar el apoyo y condiciones ofrecidas a las escuelas con la presión derivada de las mediciones externas, reduciendo significativamente el número de mediciones censales y fortaleciendo las políticas de apoyo y generación de capacidades.
- c) Desarrollar un sistema de evaluación que promueva un compromiso responsable de las comunidades educativas y asigne responsabilidades de modo justo.
- d) Desarrollar un sistema que dé cuenta de la calidad de la educación de modo integral y permita diversas lecturas de ella, sin que ello diluya la importancia de que todos los estudiantes logren aprendizajes fundamentales.
- e) Explicitar qué interpretaciones y usos se espera que los distintos actores hagan de los resultados y cuáles son erróneos o indebidos, y monitorearlos.
- f) Avanzar hacia una cultura de uso equilibrado de la medición reconociendo transparentemente sus límites.
- g) Dar mayor énfasis a las visitas evaluativas y de orientación realizadas por la Agencia de Calidad.

5.1. Recomendaciones sobre los propósitos del sistema nacional de evaluación de aprendizajes

En esta sección el Equipo de Tarea identifica los propósitos del Sistema Nacional de Evaluación de Aprendizajes. Dado que concebimos que un sistema completo de evaluación de aprendizajes se compone de la evaluación estandarizada externa de las escuelas (SIMCE) y también de la evaluación que realizan cotidianamente los profesores en ellas (evaluación interna), definiremos los propósitos para cada una de ellas.

Propósitos de la evaluación externa de aprendizajes (SIMCE)

El objetivo último del sistema de evaluación externa es contribuir al mejoramiento de la calidad y equidad de la educación a través de los propósitos de *desarrollo* y de *responsabilización*²⁷. Al reportar resultados y pedir cuenta de ellos a las escuelas y al Estado, el sistema de evaluación cumple un propósito de *responsabilización*. Al identificar las fortalezas y debilidades a nivel del sistema educacional y de cada una de sus escuelas, genera información para alimentar decisiones de mejoramiento o *desarrollo* en todos los niveles del sistema.

A continuación se exponen los propósitos que debieran orientar el Sistema Nacional de Evaluación de Aprendizajes.

Propósitos de desarrollo del SIMCE:

1. Proveer información para monitorear el estado y avance de los logros de aprendizaje a nivel nacional, para orientar decisiones de política educacional.
2. Proveer información para orientar los dispositivos del Sistema de Aseguramiento de la Calidad para apoyar los establecimientos educacionales, identificando en forma precisa a los que más lo necesitan.
3. Proveer información para guiar decisiones de gestión curricular²⁸ por parte de los equipos directivos y técnico-pedagógicos en los establecimientos educacionales y sus sostenedores.
4. Proveer información global sobre los niveles de aprendizaje de los estudiantes en cada escuela y sobre las dificultades y fortalezas que en general presenta el aprendizaje en determinadas asignaturas en el país, para informar las prácticas pedagógicas de los docentes**.

Propósitos de responsabilización del SIMCE:

1. Contribuir a promover el compromiso profesional responsable de las comunidades educativas con el aprendizaje de sus estudiantes.

²⁷ En concordancia con la distinción que hace la OCDE (2013).

²⁸ De acuerdo con los Estándares Indicativos (p. 71), la gestión curricular se refiere a las políticas, procedimientos y prácticas que lleva a cabo el director, el equipo técnico-pedagógico y los docentes del establecimiento para coordinar, planificar, monitorear y evaluar el proceso de enseñanza-aprendizaje. Incluye las acciones tendientes a asegurar la cobertura curricular y mejorar la efectividad de la labor educativa.

2. Responsabilizar al Estado como garante de que en cada una de las escuelas se logren aprendizajes de calidad.
3. Informar a los padres para contribuir a que se comprometan con la educación de sus hijos, se involucren en las escuelas y puedan tomar decisiones informadas.

**Las mediciones externas de resultados de aprendizaje aportan información a los docentes, pero esta es global y no está referida específicamente a cada uno de los estudiantes o a unidades del programa de estudios que estén abordando, por ejemplo. Sí puede entregar información sobre los errores de conceptos o de procedimientos que en general son más frecuentes en los alumnos en una determinada área, o el nivel alcanzado al término de cada ciclo de educación básica de un establecimiento. En este sentido, el Equipo de Tarea considera que la función formativa, de informar cotidianamente las decisiones pedagógicas de los docentes respecto de estudiantes específicos, se debe cumplir a través de la evaluación realizada internamente en las escuelas.

Propósitos de la evaluación de aprendizajes realizada internamente en las escuelas

1. Monitorear el progreso de cada uno de los estudiantes y retroalimentar tanto su aprendizaje como la enseñanza que se les ofrece (función formativa).
2. Certificar²⁹ el nivel de aprendizaje alcanzado por cada uno de los estudiantes (función sumativa).

Estos propósitos guían las recomendaciones de lo que es necesario mantener, mejorar o modificar en el sistema nacional de evaluación de aprendizajes, incluyendo las políticas referidas a la evaluación de aprendizajes realizada internamente en las escuelas.

²⁹ A través de las calificaciones o notas.

4.2 Lo que es necesario mantener

A continuación se describen los aspectos del SIMCE que el Equipo de Tarea considera que es necesario mantener:

a) Las evaluaciones censales de los logros de aprendizaje de los estudiantes en cada establecimiento escolar

El Equipo de Tarea considera que es relevante poder seguir contando con información confiable sobre los resultados de aprendizaje a nivel de cada escuela. Esto permite monitorear y dar cuenta del logro de aprendizajes básicos en cada una de ellas. Además se favorece así tanto la responsabilización del Estado, que debe proveer condiciones y dar garantías de acceso a una educación de calidad en cada escuela. Además puede contribuir al desarrollo de un compromiso profesional responsable de los docentes, directivos y sostenedores a cargo de ellas.

Adicionalmente, a diferencia de una evaluación muestral, una medición censal proporciona un diagnóstico fino que permite focalizar esfuerzos diferenciados según las necesidades particulares de los establecimientos. La evidencia acumulada a través de años de medición indica que las áreas curriculares de mayor debilidad no siempre son las mismas en todas las escuelas, que las brechas al interior de las mismas presentan variaciones, y que en muchos casos estas obtienen resultados diferentes de lo que sería predecible dados los antecedentes socioeconómicos de sus estudiantes.

Proporcionar periódicamente a cada establecimiento escolar información sobre sus resultados de aprendizaje en relación con referentes nacionales, permite además que ellos monitoreen sus avances y tomen decisiones de gestión pedagógica basadas en evidencia.

b) El acceso público a información sobre resultados de la evaluación en diferentes ámbitos y considerando el contexto de las escuelas.

La información de resultados de las evaluaciones debe entregarse a las comunidades educativas para informar sus procesos de mejora. También se considera que debe haber acceso público a la información de resultados en evaluaciones nacionales de aprendizaje. Sin embargo, se enfatiza que estos datos nunca deben exponerse bajo la forma de listados o rankings que se basen en indicadores aislados o en puntajes obtenidos en las pruebas estandarizadas sin considerar las características de la población que atienden. Para evitar comunicar visiones estrechas sobre la calidad de la educación, tampoco debiesen reportarse resultados de las pruebas aisladamente, es decir sin otra información de la calidad de las escuelas.

c) El rigor técnico y preocupación permanente por mejorar los instrumentos y procedimientos de evaluación.

Para potenciar el constante mejoramiento que realiza la Agencia de Calidad respecto de sus propios procedimientos y métodos de trabajo, el Equipo de Tarea sugiere mantener la práctica

de someterlos a revisiones externas Asimismo, se recomienda mantener el informe técnico que se comenzó a publicar en 2014.

Dadas los crecientes desafíos metodológicos y las complejidades técnicas involucradas en la medición y también en la Ordenación de escuelas, se recomienda que la Agencia de Calidad incorpore un Consejo Técnico Asesor de expertos internacionales (*)³⁰ con una agenda sistemática de revisiones externas para asegurar una adecuada continuidad y actualización de los procesos de mejora.

d) La entrega de información útil para la gestión curricular de los establecimientos escolares.

El Equipo de Tarea reconoce que ha habido grandes avances en torno a la manera de entregar información sobre los logros de las escuelas, y sugiere mantener un programa de desarrollo continuo de la comunicación de resultados, diferenciando la información entregada según las necesidades de cada actor. Además se sugiere incluir un estudio sistemático de la valoración, comprensión y los usos reales que cada tipo de actor le da a esta información, de manera de ir contrastando los supuestos con lo que realmente sucede en la práctica, e informar así qué mejoramientos es necesario realizar.

e) Las acomodaciones para evaluar estudiantes con necesidades educativas especiales permanentes

La incorporación de pruebas especiales para evaluar alumnos con discapacidad visual y auditiva, es uno de los avances relevantes valorados por el Equipo de Tarea. Se recomienda que además de esto, se estudie de qué modo hacer más efectivos los resguardos para asegurar que la medición no induzca comportamientos estratégicos y no desincentive políticas de inclusión de las escuelas. También se sugiere evaluar acomodaciones para las necesidades educativas transitorias como por ejemplo, más tiempo para rendir la prueba, entre otros.

(*) Recomendación con implicancias de modificación legal.

4.3. Lo que es necesario modificar

A continuación se describe cada modificación sugerida.

- a) *Diseñar e implementar una política amplia que permita potenciar la evaluación de aprendizajes que realizan los profesores y directivos al interior de las escuelas y complementarla con las mediciones externas.*

A juicio del Equipo de Tarea, el aporte de un Sistema Nacional de Evaluación de Aprendizajes al mejoramiento de la calidad dependerá en parte importante de cuánto contribuya al desarrollo de capacidades profesionales de los docentes y de los directivos. La evaluación de los aprendizajes, en todos sus ámbitos y formas, es una de las competencias profesionales claves en la docencia, ya que a través de ella los profesores comunican a los estudiantes lo que esperan de su desempeño, obtienen evidencias sobre las áreas en que su enseñanza ha sido más o menos eficaz, y proporciona oportunidades para que puedan dar a sus alumnos retroalimentación sobre lo que necesitan mejorar.

Forma también parte de las competencias profesionales de docentes y directivos su capacidad para utilizar la información que proveen las mediciones nacionales e internacionales para el mejoramiento de sus prácticas. Consecuentemente, un sistema completo de evaluación debe incluir y fortalecer la evaluación de aprendizajes que hacen los profesores cotidianamente al interior de las escuelas, y robustecer sus capacidades y las de sus directivos para interpretar y utilizar la información que proviene de las mediciones externas.

Es necesario establecer una política de evaluación que incluya y desarrolle la evaluación de aprendizajes en las escuelas. Para ello es necesario definir responsabilidades y asignar recursos a las instituciones (Ministerio de Educación y Agencia de Calidad) y unidades o divisiones al interior de ellas a cargo de los diferentes componentes de una política integrada de evaluación de aprendizajes que incluya al menos, los siguientes componentes:

- Potenciar el uso de evaluaciones externas, optimizando la información que ellas proveen de modo que respondan a las necesidades de los establecimientos escolares.
- Generar orientaciones sobre evaluación formativa y sumativa realizada en las escuelas y el modo de articular la información que proveen con la que entregan las mediciones externas.
- Proveer instrumentos para ser aplicados en las escuelas de modo autónomo para verificar en qué medida se están alcanzando los estándares de aprendizaje.
- Entregar ilustraciones, tareas y ejemplos de actividades de evaluación, para enriquecer la evaluación de aprendizajes asociada a los programas de estudio. Adicionalmente, se debiera ilustrar, a través de ejemplos de interacción en sala de clases, cómo es posible hacer un uso cotidiano de la evaluación formativa.
- Establecer una estrategia de formación inicial y continua de docentes y directivos en el área de evaluación de aprendizajes, incluyendo la interpretación y uso de las mediciones externas.

- Generar las condiciones (por ejemplo, tiempo docente y generación de capacidades) en las escuelas para hacer viable el uso de la evaluación tanto interna como externa para decisiones pedagógicas y de gestión.

b) Equilibrar el apoyo y condiciones ofrecidas a las escuelas con la presión derivada de las mediciones externas.

La exigencia y presión por mejorar los niveles de aprendizaje de los estudiantes puede constituirse en un aliciente para las escuelas cuando estas cuentan con las capacidades y las instancias de desarrollo profesional para responder a tales exigencias. Sin embargo, cuando no es así, intensificar la presión por la vía de asociar incentivos y consecuencias a las pruebas, en algunos casos puede motivar comportamientos estratégicos de corto plazo y efectos no intencionados, especialmente en las escuelas con capacidades más precarias y de resultados previos más bajos en las mediciones³¹.

Como se señaló en la sección de diagnóstico, el Equipo de Tarea advierte sobre el desequilibrio entre la creciente presión que experimentan las escuelas y la debilidad de los mecanismos de apoyo para lograr los resultados esperados. Considerando lo anterior, se proponen las siguientes medidas tendientes a equilibrar la presión ejercida y a robustecer las estrategias y dispositivos para el fortalecimiento de las capacidades de docentes y directivos de las escuelas proveyendo las condiciones necesarias.

Para fortalecer las estrategias de apoyo y construcción de capacidades:

- Impulsar una carrera docente capaz de incorporar, formar y retener profesores adecuadamente preparados y con el tiempo disponible para trabajar en equipo.
- Desarrollar instancias descentralizadas que articulen las estrategias de apoyo a las escuelas y promover el desarrollo de redes estables de colaboración entre ellas.
- Potenciar y fortalecer, al interior del Ministerio de Educación, una unidad responsable de monitorear los mecanismos de apoyo a las escuelas (incluyendo la supervisión de las ATE) y de promover y articular las políticas de apoyo y construcción de capacidades a nivel central y descentralizado.
- Fortalecer las visitas evaluativas de la Agencia de Calidad para mejorar las capacidades de autodiagnóstico de las escuelas y para focalizar la atención en los procesos que han dado origen a los resultados observados proporcionando así información diagnóstica que oriente la mejora escolar.

³¹ Si bien en Chile no hay suficiente investigación al respecto, una revisión de los estudios realizados en Estados Unidos identifica que existe un efecto diferencial de las pruebas con altas consecuencias según los resultados previos de las escuelas y las capacidades instaladas (conocimiento y habilidades de sus profesores y directivos), entre otros factores (Faxon Mills et al, 2013).

Para reducir la excesiva presión externa:

- Impedir o prohibir³² (*) la divulgación de rankings basados en puntajes promedio que no consideren, entre otras, las características socioeconómicas de los establecimientos. De no ser esto legalmente viable, al menos es recomendable desincentivar y reprobar públicamente la divulgación y uso de dichos rankings.
- En las escuelas municipales, sustituir la revocación del reconocimiento oficial que contempla la Ley de Aseguramiento de la Calidad para los establecimientos de bajo desempeño sostenido, por su re-estructuración³³ o alternatively por la fusión con otro establecimiento. El Equipo de Tarea concuerda en que no es admisible que en el país hayan escuelas en que los estudiantes no estén logrando aprendizajes básicos. Sin embargo, considera que para garantizar la provisión pública, para las escuelas municipales la medida más apropiada es la intervención o fusión. En el caso de las escuelas particulares subvencionadas el procedimiento de re-estructuración debe ser sugerido por la Agencia de Calidad. De no aceptar el sostenedor, se desencadenaría la revocación del reconocimiento oficial.
- Reemplazar la revocación automática del reconocimiento oficial de las escuelas de bajos resultados sostenidos por una re-estructuración como primera medida. El Equipo de Tarea considera que decisiones relevantes respecto de las escuelas requieren un juicio profesional y no se pueden gatillar automáticamente basándose únicamente en los resultados de algoritmos cuantitativos como establece actualmente la Ley de Aseguramiento de la Calidad, que determina que se deberá cerrar las escuelas que después de un máximo de 5 años no logren salir de la categoría de más bajo desempeño. El Equipo de Tarea propone que en el caso de las escuelas municipales, se sustituya la revocación del reconocimiento oficial que contempla la Ley para los establecimientos de bajos resultados sostenidos, por su re-estructuración, o alternatively por su fusión con otras escuelas, asegurando así la provisión de educación pública de calidad en todo el territorio. Para las escuelas particulares subvencionadas de persistente bajo desempeño, se sugiere que la Agencia de Calidad proponga, en una primera etapa, una re-estructuración del equipo profesional a cargo. En caso que el sostenedor no emprenda la re-estructuración, o que esta no logre revertir el desempeño del establecimiento, se revocaría su reconocimiento oficial (*).
- Reducir significativamente las mediciones censales externas resguardando que su número y frecuencia sea el estrictamente necesario para cumplir con los propósitos de mejoramiento y responsabilización. Esta medida busca tres objetivos: reducir la carga de mediciones externas para lograr un mejor balance entre presión y apoyo, dejar mayor tiempo a las escuelas para analizar y utilizar la información, y confiar a sus equipos profesionales la aplicación y

³² Bárbara Eyzaguirre no está de acuerdo con prohibir la publicación de rankings dado que sostiene que una medida de esta naturaleza vulnera la libertad de expresión. Propone en cambio, que se informe a la población respecto de sus alcances y limitaciones. Específicamente, que los rankings de puntajes solo comunican lo que los alumnos saben y pueden hacer en promedio y que no son un buen indicador de efectividad escolar.

(*) Recomendación con implicancias de modificación legal.

³³ Se entiende re-estructuración como el reemplazo de parte del equipo profesional de la escuela.

análisis de las pruebas en los cursos y asignaturas que ya no serían evaluadas por mediciones externas tipo SIMCE.

Para el corto plazo se plantean dos alternativas diferentes al interior del Equipo de Tarea para materializar esta reducción en las mediciones externas:

1. La primera alternativa³⁴ recomienda reducir a la mitad los cursos evaluados y las pruebas aplicadas anualmente, bajando el número de mediciones censales de 15 a 8 pruebas anuales en promedio, y de 6 niveles o cursos a 3. El calendario diseñado (en anexo) propone evaluar 4° básico y 2° medio todos los años, acompañados de 6° y 8° alternadamente, año por medio. Se propone dejar de evaluar censalmente 2° básico y 3° medio. Esta propuesta de calendario permite seguir algunas generaciones de estudiantes y evaluarlas más de una vez para verificar cuánto progresan en su aprendizaje (por ejemplo de 4° a 6° básico) y cuánto aportan las escuelas a este avance. Lo anterior ampliaría y mejoraría la información disponible para las escuelas, favoreciendo una responsabilización más justa al identificar los avances que son atribuibles a ellas. Además, quienes adscriben a esta propuesta de calendario valoran obtener datos confiables sobre cada escuela (que dependan menos de la generación particular que fue evaluada cada vez) y proveerles información suficientemente periódica y oportuna para monitorear sus avances en los planes de mejora y orientar sus decisiones de gestión curricular.
2. La segunda propuesta³⁵ sugiere conservar los mismos niveles y asignaturas medidas en la actualidad, pero reduciendo desde seis a dos los niveles medidos cada año³⁶, y evaluando cada nivel o curso o cada tres años. Esto implicaría reducir a un tercio los niveles evaluados cada año. Adicionalmente, se sugiere comenzar el calendario otorgando un periodo de dos años sin pruebas censales para implementar las recomendaciones del Equipo de Tarea que sean acogidas por el Ministerio de Educación, especialmente las referidas a fortalecer las capacidades internas de las escuelas para la evaluación de aprendizajes y para hacer buen uso de los datos de las pruebas SIMCE disponibles hasta ahora. Quienes adscriben esta alternativa, proponen priorizar el uso de la información para identificar y dar un apoyo especial a aquellas escuelas de cada región que están teniendo mayores dificultades en hacer que sus estudiantes logren el dominio de los aprendizajes básicos. Este uso de la información haría innecesario evaluar más de una vez a la misma generación de estudiantes para verificar cuánto avanzan los alumnos en su aprendizaje y poder establecer a cuál es el aporte de las escuelas al mismo. Desde esta perspectiva, una periodicidad trianual para cada nivel evaluado es adecuada para que las escuelas dispongan del tiempo requerido para llevar adelante sus planes de mejora sin desviar su atención de este esfuerzo. También se desestima evaluar un curso en básica y uno en media anualmente para la ordenación, ya que quienes respaldan esta

³⁴ Esta alternativa de calendario fue propuesta por Paulina Bustos, Alejandro Carrasco, Rafael Carrasco, Bárbara Eyzaguirre Jorge Manzi; Lorena Meckes; Ernesto Treviño; Xavier Vanni y María José Zañartu.

³⁵ Esta alternativa fue propuesta por Juan Eduardo García-Huidobro, Guillermo Scherping y Mariano Rosenzvaig.

(*) Esta alternativa de calendario requiere modificación legal.

³⁶ Si además se prescindiera de 2° básico y de III medio, en promedio se aplicarían pruebas a 1,5 niveles o cursos cada año.

propuesta de calendario estiman que no se debiera hacer tal ordenación de las escuelas (*).

Las dos posiciones planteadas coinciden en recomendar que sus respectivos calendarios comiencen a aplicarse lo antes posible.

También hay coincidencia en recomendar que en los niveles y asignaturas que dejen de ser evaluados por el SIMCE (por ejemplo 2° básico), la Agencia de Calidad entregue orientaciones e instrumentos para ser aplicados, corregidos colaborativamente y analizados por los equipos profesionales de las escuelas, para que ellos puedan evaluar autónomamente los logros de aprendizaje de sus estudiantes en relación con los estándares nacionales. Adicionalmente, para estos niveles y asignaturas, se propone realizar estudios muestrales que permitan monitorear su evolución a nivel nacional y al mismo tiempo calibrar externamente los resultados que obtienen las escuelas en las pruebas administradas internamente.

Se propone mantener muestral y voluntaria la prueba de 3° medio de Inglés para los establecimientos que reciben subvención, ya que da la posibilidad de obtener un certificado de competencia en el idioma a estudiantes que se inscriban.

Pruebas Muestrales

Ambas alternativas de calendario coinciden en complementar las pruebas censales aplicadas a todas las escuelas con pruebas muestrales para monitorear los logros de aprendizaje en áreas difíciles o innecesarias de evaluar en pruebas masivas de lápiz y papel. Específicamente, se considera relevante evaluar la comunicación oral de los estudiantes como también su formación ciudadana y competencias genéricas para el mundo del trabajo en el caso de la Educación Media Técnico Profesional. También se recomienda realizar estudios en profundidad sobre convivencia escolar y sobre habilidades no cognitivas y habilidades cognitivas de orden superior.

c) Desarrollar un sistema de evaluación que promueva un compromiso profesional responsable de las comunidades educativas y asigne responsabilidades de modo justo.

El Equipo de Tarea entiende que la *responsabilización* debe ser justa y debidamente adjudicada a los distintos actores y niveles del sistema educativo (escuelas, sostenedores, Ministerio de Educación). En este sentido, se recomienda que así como se hace para las escuelas, haya un reporte público de los resultados por sostenedor (cuando tiene más de una escuela), incluyendo los informes de las visitas de evaluación para revisar el cumplimiento de los estándares indicativos. Asimismo, resulta recomendable que los sostenedores deban presentar un plan de desarrollo y apoyo a sus escuelas, con metas de procesos y resultados, respecto de los cuales debiera dar cuenta anualmente. Finalmente, se sugiere revisar los estándares indicativos actuales, para incorporar criterios y estándares más claramente pertinentes para sostenedores de un conjunto de escuelas, ya que los actuales están más bien concebidos para sostenedores de una única escuela.

Del mismo modo, el Ministerio de Educación debiera dar cuenta periódicamente del avance y de la eficacia del plan de Aseguramiento de la Calidad, como también del resultado de las evaluaciones de las políticas de mejoramiento implementadas.

A nivel de las escuelas, el sistema debiera equilibrar su actual énfasis en las consecuencias externas asociadas a los resultados de la medición de aprendizajes, con el desarrollo de un compromiso profesional responsable con las metas educativas nacionales. Un énfasis muy marcado en premios y sanciones motiva conductas estratégicas que desvían la atención del objetivo central que es el mejoramiento del aprendizaje de los estudiantes. Para avanzar hacia una cultura de responsabilidad interna por los resultados, y construir una cultura evaluativa, se propone fortalecer prácticas como el análisis colectivo de la información de las pruebas externas, la aplicación y análisis de evaluaciones realizadas internamente, y dar cuenta a la comunidad escolar de los compromisos de la escuela con las metas educacionales del país.

d) *Desarrollar un sistema de evaluación que dé cuenta de la calidad de la educación de modo integral y permita diversas lecturas de ella, sin que ello diluya la importancia de que todos los estudiantes logren aprendizajes fundamentales.*

Tomando en cuenta el diagnóstico realizado anteriormente, el Equipo de Tarea propone:

- *Reportar los resultados del SIMCE en contexto y siempre acompañado de información sobre otras dimensiones de calidad.* Se sugiere dejar de reportar puntajes del SIMCE aislados y entregar información contextualizada a nivel de cada escuela y del sistema en su conjunto. Además, se sugiere entregar información más completa y para esto se propone hacer visibles los resultados en los otros indicadores de calidad junto con el resultado de las visitas evaluativas que aportarán información sobre los procesos pedagógicos y de gestión en las escuelas.
- *Mayor reconocimiento público de escuelas por una diversidad mayor de buenas prácticas.* Se recomienda que haya mayor reconocimiento público de escuelas a partir de una amplia diversidad y tipo de buenas prácticas. Dado que las visitas de aprendizaje contempladas en la Ley para las escuelas de Alto Desempeño buscan destacar buenas prácticas educativas, se propone que se aproveche esta oportunidad para que la Agencia destaque experiencias enriquecedoras en distintos ámbitos que puedan ser orientadoras para otros establecimientos.
- *Cambiar la denominación actual del SIMCE (Sistema de Medición de la Calidad de la Educación) para clarificar que su foco es la medición de logros de aprendizaje y comunicar que no pretende ser una evaluación de la calidad completa del quehacer de una escuela.*
- *Ampliar los tipos de metas e indicadores solicitados obligatoriamente a los establecimientos con motivo de convenios de desempeño incluyendo metas de proceso y en otras áreas de resultados.*
- *Generar información sobre la Educación Media Técnico-Profesional*
Dada la complejidad de los aprendizajes a evaluar, se sugiere diseñar un programa de evaluación muestral que se centre en competencias genéricas

claves tanto para el mundo del trabajo como para acceder a niveles post secundarios de formación³⁷.

e) *Explicitar qué interpretaciones y usos se espera que los distintos actores hagan de los resultados y cuáles son erróneos o indebidos, y monitorearlos.*

Para avanzar en este sentido, el Equipo de Tarea propone:

- Generar un marco conceptual del Sistema Nacional de Evaluación que explicita los propósitos, las interpretaciones y los usos propuestos para los resultados de este por parte de los diferentes actores involucrados. Con estas definiciones será posible posteriormente recoger datos sobre si efectivamente es este el tipo de acciones que se movilizan en las escuelas.
- Asignar presupuesto para estudiar de modo permanente si efectivamente se dan los usos y efectos intencionados o los no deseados, y de este modo ‘evaluar la evaluación’.
- Identificar los efectos beneficiosos de las evaluaciones.
- Identificar los usos y prácticas indebidas e interpretaciones erróneas de la información y adoptar un papel activo en difundir cuáles son estos.
- Implementar acciones para prevenir o reprobar usos indebidos, tales como excluir estudiantes de la medición, asignar incentivos positivos o negativos a profesores individuales basados en los resultados de un curso, presionar a los estudiantes, descuidar las asignaturas no evaluadas, etc.

f) *Avanzar hacia una cultura de uso equilibrado de la medición reconociendo transparentemente sus límites.*

Si bien las medidas sugeridas ayudarán a que el sistema de medición cumpla con su rol de contribuir al mejoramiento de los resultados de aprendizaje, el Equipo de Tarea considera importante hacer presente el riesgo de descansar exclusivamente en la evaluación para asegurar la calidad de los mismos. Aunque resulta imposible avanzar sin información sobre lo que se requiere modificar, evaluar no es sinónimo de mejorar. La evaluación debe ser parte integral de un proceso que también involucra planificación, implementación y ajustes de las acciones implementadas. Por lo tanto, el valor de la evaluación para la mejora está determinado por la calidad de la evaluación misma, pero también por las capacidades de sostenedores, directivos y docentes para analizar los resultados y transformarlos en un diagnóstico y un conjunto de medidas que se orienten a la mejora. Sin ambos componentes, difícilmente la evaluación llevará a la mejora.

³⁷ En el mediano plazo, se podría complementar lo anterior con estudios también muestrales de competencias específicas por grupos de especialidades.

Valoramos la toma de decisiones y la gestión de las escuelas basadas en información. Sin embargo, es necesario considerar que dicha información debe ser lo más útil posible e integrar los datos provenientes de mediciones con los antecedentes que provienen del juicio profesional. En este sentido nos parece que decisiones tales como la revocación del reconocimiento oficial o la re-estructuración de escuelas, nunca debieran basarse solo en algoritmos y resultados automáticos de medición sino que en fuentes variadas de información, entre otras, las visitas de escuelas (*). De acuerdo a la actual Ley de Aseguramiento de la Calidad, la revocación del reconocimiento oficial se deriva automáticamente de permanecer en la última categoría de desempeño por un período de 4 años. El Equipo de Tarea recomienda que decisiones de este tipo, como por ejemplo la re-estructuración de las escuelas, incorporen un juicio profesional fundamentado por escrito y apelable al Consejo Nacional de Educación.

Quienes ven que la Ordenación puede ser un aporte, valoran que la Agencia de Calidad haya acogido la recomendación del Consejo Nacional de Educación de realizar una marcha blanca de la instalación del Sistema de Aseguramiento de la Calidad, que permita estudiar y verificar la validez de la Ordenación. Asimismo, recomiendan extender la marcha blanca por un año adicional para asegurar dicha validez³⁸ (*). Se entiende aquí validez en dos sentidos: (i) corroborar que el desempeño de las escuelas, según sus categorías de efectividad, se corresponde sistemáticamente con la apreciación derivada de las visitas de evaluación; y (ii) que las acciones que la Ordenación estimula en las escuelas y en sus sostenedores son las que se buscaba motivar³⁹.

Sin duda la Ley debe resguardar que exista cierta estabilidad del Sistema de Aseguramiento de la Calidad en el tiempo (de modo que las escuelas no se vean sometidas, por ejemplo, a nuevas exigencias de modo imprevisto). Sin embargo, es igualmente necesario que haya preocupación por la validez del Sistema, tanto en términos de la precisión, confiabilidad y estabilidad de los métodos utilizados para categorizar el desempeño de las escuelas, como por los efectos que tienen las consecuencias definidas para los actores y el sistema educativo en su conjunto. Conforme a lo anterior, el Equipo de Tarea sugiere que debe explicitarse en la Ley, la posibilidad de ajustar la metodología durante la marcha blanca y también posteriormente, si se identifican debilidades en cualquiera de estos ámbitos. Hoy la Ley de Aseguramiento impide esto ya que define que no es posible modificar la metodología antes de cuatro años (*).

Por último, se recomienda hacer explícitos los límites de la medición tanto en aquello que es evaluado como en la precisión de sus resultados. Por lo mismo, se sugiere explicitar claramente

³⁸ Bárbara Eyzaguirre no apoya la extensión de la Marcha Blanca más allá de 2015. Sostiene que en la ley de Subvención Preferencial publicada el año 2008 ya se estipulaba que los establecimientos que no cumplen en forma reiterada con los objetivos de logro de aprendizaje enfrentarían la revocación del reconocimiento oficial. El sistema de ordenación perfecciona y reemplaza la clasificación SEP, y por lo tanto, en su opinión no es presentable seguir aplazando la entrada en vigencia de esta medida que protege a los alumnos que asisten a establecimientos que no ofrecen la oportunidad de aprender.

(*) Recomendación con implicancias de modificación legal.

³⁹ El Equipo de tarea reconoce que será difícil estudiar estos efectos en un contexto de bajas consecuencias como es la marcha blanca, por lo que se recomienda extender el estudio sobre las consecuencias de la medición y ordenación después de que entre en vigencia.

(*) Recomendación con implicancias de modificación legal.

las habilidades que no son cubiertas por las pruebas y el error o imprecisión asociada a la ordenación o a categorizaciones basadas en resultados de la medición.

En Chile, contamos con un sistema de evaluación de logros de aprendizaje de calidad técnica reconocida, pero tenemos un gran desafío en término de los usos de la información que nos entrega. Esperamos que la implementación de las recomendaciones y resguardos sugeridos por el Equipo de Tarea contribuyan a que las escuelas del país ofrezcan más y mejores oportunidades a nuestros estudiantes.

Anexo

Calendario de de Pruebas Censales Opción 1 ⁴⁰

Niveles	Calendario Vigente			Propuesta de Calendario				
	2013	2014	2015	2016 (Año 1)	2017 (Año 2)	2018 (Año 3)	2019 (Año 4)	2020 (Año 5)
2° básico	LEC	LEC	LEC					
4° básico	LEC MAT	LEC MAT	LEC MAT	LEC MAT	LEC MAT	LEC MAT **	LEC MAT	LEC MAT
	CNAT	H G y CSOC	CNAT					
6° básico	LEC MAT	LEC MAT	LEC MAT		LEC MAT		LEC MAT	
	ESC	CNAT ESC	H G y CSOC ESC		CNAT ESC		H G y CSOC ESC	
8° básico	LEC MAT	LEC MAT	LEC MAT	LEC MAT		LEC MAT **		LEC MAT
	CNAT	H G y CSOC	CNAT					
2° medio	LEC MAT	LEC MAT	LEC MAT	LEC MAT	LEC MAT	LEC MAT*	LEC MAT	LEC MAT
		CNAT	H G y CSOC	CNAT	H G y CSOC	CNAT	H G y CSOC	CNAT
3° medio		Inglés						
Número de Niveles	5	6	5	3	3	3	3	3
Número de Pruebas	12	15	14	7	9	7	9	7

* Esta cohorte rinde PISA el 2018 junto con SIMCE en 2° medio, y rindió 4° básico en 2012.

** Esta cohorte rinde TIMSS en 2018, lo que permite comparar sus resultados de MAT en la prueba nacional con resultados a nivel internacional.

*** La prueba de Comprensión Lectora y Escritura de 2° básico, referidas a los estándares nacionales de aprendizaje, la proveerá la Agencia de Calidad a las escuelas para su aplicación autónoma en un sistema de moderación externa de corrección.

⁴⁰ Los colores indican que hay seguimiento de una misma cohorte, por tanto podrían generarse indicadores de progreso y valor agregado.

Códigos de asignaturas en calendario: LEC = Lectura, MAT = Matemática, ESC = Escritura, CNAT = Ciencias Naturales, H G y CSOC = Historia, Geografía y Ciencias Sociales.

Referencias

- Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2): 139-148. Disponible en: <http://www.pdkintl.org/kappan/kbla9810.htm>.
- Faxon-Mills, S., Hamilton, L., Rudnick, M. & Stecher, B. (2013). *New assessments, better Instruction? Designing assessment systems to promote instructional improvement*. Washington, D.C.: Rand Corporation.
- Flórez, M. T. (2013). Análisis crítico de la validez del sistema de medición de la calidad de la educación (SIMCE). Oxford University Centre for Educational Assessment: Oxford.
- Fullan, M. (2010). Positive Pressure. En A. Hargreaves, A. Lieberman, M. Fullan, D. Hopkins (Eds.) *Second International Handbook of Educational Change* (pp. 119-130), Springer International Handbooks of Education 23, Springer Science+Business Media B.V.
- Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London: Routledge.
- Jensen, J. L., McDaniel, M. A., Woodard, S. M., & Kummer, T. A. (2014). Teaching to the test ... or testing to teach: Exams requiring higher order thinking skills encourage greater conceptual understanding. *Educational Psychology Review*, 26, 307-329.
- Madaus, G. F., & Clarke, M. (2001). The adverse impact of high-stakes testing on minority students: Evidence from one hundred years of test data. In G. Orfield & M. L. Kornhaber (Eds.), *Raising standards or raising barriers? Inequality and high-stakes testing in public education* (pp. 85-106). New York: Century Foundation Press.
- Manzi, J., González, R. & Sun, Y. (2011). *La evaluación docente en Chile*. Santiago: MIDE UC.
- Manzi, J., Bogolasky, F., Gutiérrez, G., Grau, V., Volante, P. (2014). Análisis sobre valoraciones, comprensión y uso del SIMCE por parte de directores escolares de establecimientos subvencionados. Informe preliminar FONIDE (no publicado).
- Mineduc (2014). *Estándares indicativos de desempeño para los establecimientos educacionales y sus sostenedores*. Santiago: Mineduc. Disponibles en: <http://www.agenciaeducacion.cl/visitas-evaluativas/estandares-indicativos-de-desempeno/>.
- OECD (2013). *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*. OECD Reviews of Evaluation and Assessment in Education. Paris: OECD Publishing. Disponible en: <http://dx.doi.org/10.1787/9789264190658-en>.