



**PLAN DE EVALUACIONES NACIONALES
E INTERNACIONALES
2016-2020**



Tabla de contenido

I.	Introducción.....	5
II.	Antecedentes.....	6
	a) Elementos del contexto educativo nacional.....	6
	b) Definiciones del marco legal.....	8
	c) La evaluación externa en el marco de un sistema nacional integrado de evaluación.....	9
	d) Elementos diagnósticos que sustentan el Plan.....	14
	e) Decisiones y acciones en función del diagnóstico.....	16
	f) Propósitos de la evaluación externa de aprendizajes a nivel de establecimientos y de sistema escolar.....	17
III.	Evaluaciones externas de aprendizajes para el periodo 2016–2020.....	19
	Criterios utilizados para diseñar el Plan Nacional de Evaluaciones.....	19
	Propuesta y fundamentación de evaluaciones censales.....	23
	Tabla N° 1: Calendario de evaluaciones censales.....	23
	Fundamentación de evaluaciones censales.....	24
	Propuesta y fundamentación de evaluaciones muestrales.....	25
	Tabla N° 2: Calendario de evaluaciones muestrales.....	25
	Fundamentación de evaluaciones muestrales.....	25
	Tabla N° 3: Calendario de evaluaciones internacionales.....	30
	Fundamentación de evaluaciones internacionales.....	30
	Tabla N° 4: Relación entre evaluaciones nacionales y participación de Chile en Estudios Internacionales.....	32
	Propuesta de desarrollos y estudios exploratorios.....	32
	Análisis del plan según los criterios definidos.....	33
	Tabla N° 5: Calendario de implementación curricular.....	34
	Tabla N° 6: Vigencia de los Estándares de Aprendizaje.....	35
	Tabla N° 7: Plan de elaboración de Estándares de Aprendizaje.....	37
IV.	Elementos que complementan la evaluación externa de aprendizajes.....	41
	Propuesta y fundamentación de evaluaciones externas para uso interno.....	41

Fundamentación de evaluaciones externas usadas internamente para Lectura en 2° básico.....	42
Diseño e implementación de evaluaciones externas para ser usadas internamente	45
Otros Indicadores de Calidad Educativa.....	48
V. Comunicación y uso de resultados	50
VI. Aseguramiento de la calidad técnica de las pruebas.....	55
VII. Monitoreo de la evaluación.....	56
VIII. Análisis de la factibilidad técnica y económica del Plan	58
IX. Referencias.....	59
X. Anexos.....	64

I. Introducción

El presente documento presenta el Plan de Evaluaciones Nacionales e Internacionales para los años 2016 a 2020, en conformidad con la Ley General de Educación (Ley N° 20.370) y la Ley del Sistema de Aseguramiento de la Calidad de la Educación (Ley N° 20.529), el cual fue evaluado y aprobado por el Consejo Nacional de Educación (CNED)¹. El Plan considera un importante trabajo que refleja acuerdos nacionales sobre cómo y qué evaluar, y cómo y con qué fines utilizar la información que proveen las evaluaciones externas.

La garantía de una educación de calidad para todas y todos, de la cual el Estado es responsable, no es posible sin la información que provee la evaluación. La evaluación es el proceso que permite obtener información sobre los aprendizajes de las y los estudiantes para tomar decisiones que apunten a fortalecer los procesos de enseñanza-aprendizaje. Esta evidencia de los aprendizajes es recogida por medio de múltiples fuentes en los sistemas nacionales de evaluación, considerando tanto procesos externos a los establecimientos escolares como información que se genera dentro de ellos, y es utilizada con diferentes propósitos, todos los cuales buscan como fin último la mejora de la calidad de la educación.

El presente Plan fue desarrollada por el Ministerio de Educación en coordinación con la Agencia de Calidad de la Educación, que es la institución encargada, entre otras funciones, de la implementación de la evaluación externa de los establecimientos escolares, la que incluye la evaluación de los logros de aprendizaje como un elemento fundamental. El presente Plan fue elaborado sobre la base de una propuesta que la Agencia de Calidad presentó al Ministerio de Educación, y fue discutido y acordado entre ambas instituciones.

El documento considera tanto el plan de evaluaciones externas aprobado por el CNED como el marco de antecedentes que lo sustentan. El documento está estructurado en siete secciones: la primera consta de antecedentes y en ella se brinda un marco referencial para el Plan en términos del contexto educativo actual y el lugar de la evaluación externa en el marco de sistemas de evaluación comprensivos a nivel nacional y los propósitos que la guían; en la segunda se describen y fundamentan los criterios que sirvieron de guía para el plan, y se presentan y justifican los calendarios de pruebas censales, muestrales e internacionales para los próximos cinco años; en la tercera se explican elementos que complementan las evaluaciones externas del aprendizaje; en la cuarta se analizan las estrategias para promover la comunicación y el uso de los resultados; y en las tres secciones finales se revisan antecedentes relativos al aseguramiento de la calidad técnica de las pruebas, al monitoreo de la evaluación y al análisis de la factibilidad técnica y económica del Plan.

¹ Resolución de acuerdo 002/2016.

II. Antecedentes

a) *Elementos del contexto educativo nacional*

La pertinencia de los procesos de evaluación que se deciden implementar a nivel nacional obedece a las necesidades, desafíos y oportunidades que presenta el contexto educativo en distintos momentos de la política educacional de un país. Esta es una de las principales razones que fundan la definición de la legislación en cuanto a revisar y presentar a deliberación del CNED un plan de evaluaciones con una duración de cinco años.

En esta sección se describen tres aspectos que se consideran significativos desde la perspectiva del contexto específico en el que el presente plan es propuesto. El primer aspecto es la creación e instalación en el país del Sistema Nacional de Aseguramiento de la Calidad; el segundo es la ampliación del concepto de *calidad* a un concepto más integral, que busca relevar ámbitos tradicionalmente menos considerados en la política pública, y el tercero es la reforma educativa que impulsa el gobierno de la presidenta Bachelet, que concibe la educación como un derecho y, por lo tanto, atribuye al Estado un rol garante sobre este derecho. Estos tres aspectos serán analizados como parte del contexto del Plan.

El sistema educacional de nuestro país se encuentra en un momento de cambio. Hace pocos años surgió el Sistema Nacional de Aseguramiento de la Calidad (SAC), creado al alero de la ley 20.529, conforme a lo dispuesto en la Ley General de Educación (20.370), cuyo objeto es propender a asegurar la equidad, entendida como la igualdad de oportunidades de recibir una educación de calidad para todos los y las estudiantes. El Sistema implica el trabajo coordinado de cuatro instituciones: el Ministerio de Educación (órgano rector del sistema, responsable de la coordinación de los órganos del Estado para un trabajo eficaz y eficiente, y de generar las definiciones normativas, las políticas educativas, las orientaciones y el apoyo necesarios para el sistema escolar), la Agencia de Calidad de la Educación (que evalúa a los establecimientos educacionales y los informa y orienta según los resultados obtenidos), la Superintendencia de Educación (que fiscaliza el cumplimiento de requisitos para el reconocimiento oficial, la normativa educacional y el uso de recursos por parte de los sostenedores de los establecimientos) y el Consejo Nacional de Educación (que aprueba las Bases Curriculares, los Planes y Programas de Estudio, el Plan de Evaluaciones Nacionales e Internacionales, las normativas de calificación y promoción, y los Estándares de Aprendizaje, entre otros). El Sistema de Aseguramiento actúa en el ámbito de la educación formal, de acuerdo a los objetivos generales señalados en la Ley General de Educación N° 20.370 y sus respectivas Bases Curriculares, y opera mediante un conjunto de políticas, estándares, indicadores, evaluaciones, información pública y mecanismos de apoyo y fiscalización a los establecimientos, para lograr la mejora continua de los

aprendizajes de los alumnos y las alumnas y fomentar el desarrollo de capacidades en los establecimientos y sus cuerpos directivos, docentes y asistentes de la educación.

Tanto en la Ley General de Educación como en la Ley SAC hay una noción amplia de la calidad educativa, con un énfasis en la formación integral. A partir de esto se desarrollaron herramientas para evaluar procesos y resultados de los establecimientos educativos que se consideran como indicadores de la calidad de la educación. Por una parte, los Estándares Indicativos de Desempeño permiten analizar información sobre los procesos de gestión interna de los establecimientos escolares que, conforme a la evidencia, son relevantes tanto para la formación integral de los y las estudiantes como para el desarrollo de experiencias y espacios educativos de calidad. Por otra parte, la evaluación SIMCE referida a Estándares de Aprendizaje, y la información proveniente de los llamados Otros Indicadores de Calidad permiten recoger evidencia acerca de la calidad de estos procesos formativos en términos de resultados de aprendizaje y experiencia escolar de los estudiantes. En congruencia con estos énfasis, el presente Plan Nacional de Evaluaciones busca generar información que evidencie distintas dimensiones de una educación integral de calidad.

En este marco de definiciones estructurales del sistema, el Ministerio de Educación ha emprendido el desafío de una reforma educativa que responde a una concepción de educación como un derecho que el Estado debe garantizar, priorizando el fortalecimiento de la educación pública como espacio común de resguardo de este derecho. Teniendo como principios centrales la inclusión, la calidad y la equidad, el derecho a la educación concibe a las escuelas como espacios inclusivos que valoran y gestionan la diversidad de sus estudiantes, ofreciendo a cada uno lo que necesita para lograr aprendizajes que le permitan desarrollar sus proyectos de vida, participar activa y críticamente como ciudadanos en la sociedad y aportar al desarrollo del país. En este enfoque, las escuelas son el espacio privilegiado y estratégico donde se juega la calidad de la educación, por lo que los esfuerzos de las distintas instituciones del SAC deben orientarse a apoyar los procesos de gestión curricular y pedagógica de los establecimientos, de modo de favorecer el desarrollo de capacidades para la mejora continua. A su vez, el contexto de reforma y fortalecimiento de la educación pública en el que se encuentra el país pone de manifiesto la necesidad de mantener una perspectiva longitudinal en la información que se recoge por medio de la evaluación, para así monitorear los aprendizajes en el mediano y largo plazo mediante estos procesos de cambio y, a partir de ello, brindar información, orientar y apoyar a los establecimientos y sectores que más lo necesiten en función de una mejora escolar continua, en el marco de esta concepción de Estado frente a la educación.

En este contexto, es importante contar con la experiencia aprendida y la reflexión desarrollada luego de haber implementado las primeras evaluaciones en el contexto del Sistema Nacional de Aseguramiento de la Calidad.

Reconociendo esta necesidad, el Ministerio de Educación conformó un Equipo de Tarea² para proponer recomendaciones que fortalecieran la evaluación externa en función de apoyar los procesos de mejora. El presente Plan se construye sobre la base de esa experiencia y las sugerencias emanadas del proceso de discusión de dicho equipo.

Este plan considera también, dentro de sus antecedentes, el marco normativo del sistema educativo nacional, el que se explica sucintamente a continuación.

b) Definiciones del marco legal

La elaboración del plan nacional de evaluaciones se enmarca en la Ley N° 20.370, que establece la Ley General de Educación, y la Ley N° 20.529, que establece el Sistema Nacional de Aseguramiento de la Calidad de la Educación parvularia, básica y media y su fiscalización. Al respecto, la Ley N° 20.370 en su artículo 7° plantea: “El Ministerio de Educación y la Agencia de Calidad de la Educación velarán, de conformidad a la ley, y en el ámbito de sus competencias, por la evaluación continua y periódica del sistema educativo, a fin de contribuir a mejorar la calidad de la educación”.

Además, esta misma ley en su artículo 37 establece:

“Le corresponderá a la Agencia de Calidad de la Educación diseñar e implementar el sistema nacional de evaluación de logros de aprendizaje. Esta medición verificará el grado de cumplimiento de los objetivos generales a través de la medición de estándares de aprendizaje referidos a las bases curriculares nacionales de educación básica y media. La Agencia deberá contar con instrumentos válidos y confiables para dichas evaluaciones, que se apliquen en forma periódica a lo menos en un curso, tanto en el nivel de educación básica como en el de educación media, e informar los resultados obtenidos. Estas mediciones deberán informar sobre la calidad y equidad en el logro de los aprendizajes a nivel nacional.

Las evaluaciones nacionales e internacionales se desarrollarán de acuerdo a un plan de, a lo menos, cinco años, elaborado por el Ministerio de Educación, aprobado previo informe favorable del Consejo Nacional de Educación, emitido conforme al procedimiento del artículo 53. Este plan deberá contemplar las áreas curriculares que son objeto de evaluación, los grados de educación básica y media que son

² 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 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 del 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 (Secretario Ejecutivo) y Juan Bravo (Jefe de la División de Logros de Aprendizaje).

medidos, la periodicidad de la evaluación y las principales desagregaciones y modos de informar resultados.

Las evaluaciones nacionales periódicas serán obligatorias y a ellas deberán someterse todos los establecimientos educacionales de enseñanza regular del país”.

Por su parte, la Ley N° 20.529 establece en su artículo 11:

“El sistema nacional de medición del grado de cumplimiento de los estándares de aprendizaje de los alumnos, así como la medición del grado de cumplimiento de los otros indicadores de la calidad educativa, será de aplicación obligatoria para todos los establecimientos educacionales reconocidos oficialmente por el Estado. La Agencia podrá realizar las mediciones respectivas directamente o por medio de terceros.

Las mediciones del grado de cumplimiento de los estándares de aprendizaje de los alumnos, referidos a los objetivos generales señalados en la ley y sus respectivas bases curriculares y de los otros indicadores de calidad educativa, se realizarán mediante instrumentos y procedimientos estandarizados, válidos, confiables, objetivos y transparentes. En el caso de los instrumentos referidos a la medición del cumplimiento de los estándares de aprendizaje, estos se aplicarán en forma periódica en distintos cursos y sectores de aprendizaje, en forma censal a lo menos en algún curso, tanto del nivel de enseñanza básica, como de enseñanza media”.

Considerando lo estipulado en ambas leyes, los requisitos legales con los que debe cumplir el plan de evaluaciones son los siguientes:

- Entregar información sobre la calidad y equidad en el logro de los aprendizajes a nivel nacional.
- Incorporar la medición del cumplimiento de los Estándares de Aprendizaje en forma periódica en distintos cursos y sectores de aprendizaje, en forma censal a lo menos en algún curso, tanto del nivel de enseñanza básica como de enseñanza media.

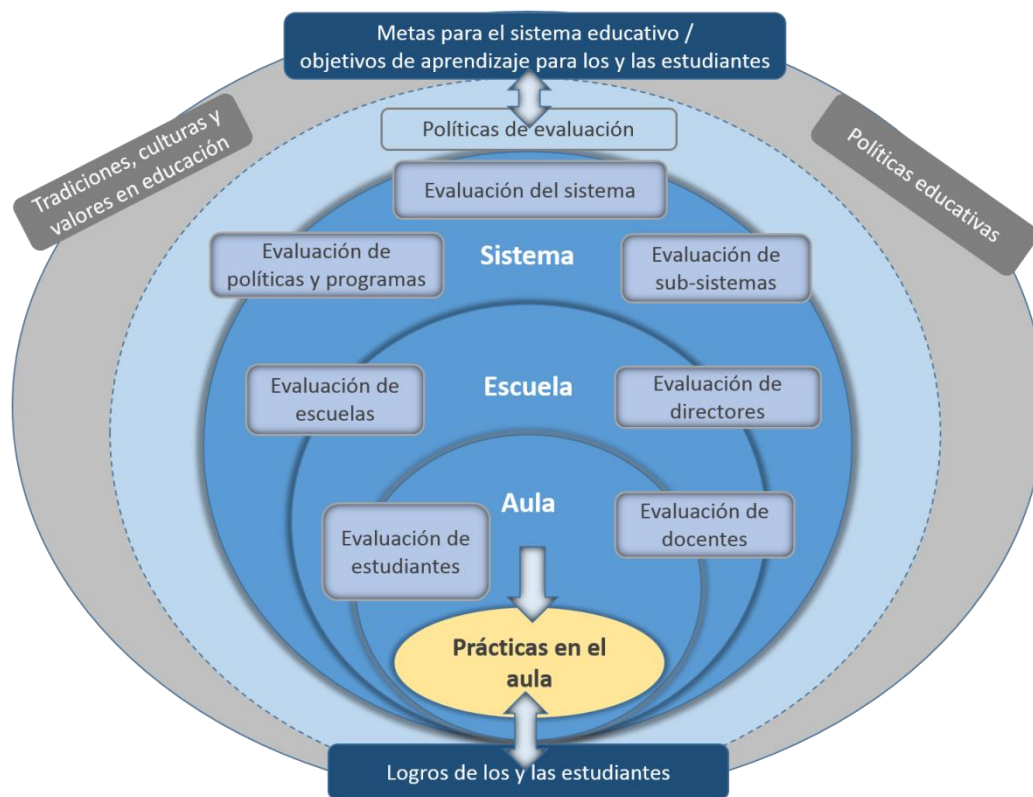
Considerando estas definiciones desde la perspectiva normativa, tanto el Ministerio de Educación como la Agencia de Calidad reconocen la importancia de generar acciones y estrategias coherentes y articuladas para apoyar los procesos de mejora de los establecimientos educativos. Por lo tanto, el presente plan, aunque se centra en la evaluación externa de aprendizajes, asume el desafío de conceptualizarse como parte de un sistema más integrado de evaluación.

c) La evaluación externa en el marco de un sistema nacional integrado de evaluación

La meta fundamental de la evaluación en un sistema educativo es proveer información para tomar decisiones que permitan mejorar. Tal como muestra el siguiente diagrama, un sistema nacional de evaluación comprende

diferentes componentes que cumplen distintos propósitos específicos o funciones dentro de este, incluyendo evaluación a nivel de sistema, evaluación de políticas o programas, evaluación del desempeño de docentes, evaluación de directivos, evaluación de establecimientos educacionales y evaluación de aprendizajes de los y las estudiantes (a nivel individual, realizada al interior de los establecimientos). El Plan de Evaluaciones Nacionales e Internacionales se centra en el diseño de la evaluación externa de aprendizajes³ como un componente central de la evaluación de establecimientos educacionales y del sistema educativo.

Figura Nº 1: Componentes y propósitos de un sistema nacional de evaluaciones



Adaptación de OECD (2013)

Para la evaluación del sistema educativo y monitoreo a nivel de políticas y programas, en Chile se participa en evaluaciones internacionales y se desarrollan evaluaciones nacionales tanto censales como muestrales, cada una de las cuales provee información diferente y valiosa para la mejora del sistema. Las definiciones

³ La evaluación es externa a los establecimientos cuando se trata del SIMCE, y externa al país cuando se trata de evaluaciones internacionales.

principales de estas evaluaciones las presenta el Ministerio de Educación con la asesoría de la Agencia de Calidad de la Educación, y su diseño e implementación está a cargo de la Agencia de Calidad.

Participar en evaluaciones o estudios internacionales permite ir monitoreando los logros en comparación con otros países, dando una mirada exterior a las políticas públicas nacionales que promueve la reflexión y la mejora, y entregando información acerca de factores que se asocian con altos logros de los y las estudiantes. Este tipo de evaluaciones permite también recoger valiosa información de insumo para decisiones de desarrollo curricular, basadas en análisis comparados de currículos y políticas educativas en distintos países y su logro en estos estudios.

Las evaluaciones externas nacionales, tanto muestrales como censales, brindan importante información para el monitoreo, evaluación y toma de decisiones de política pública, como, por ejemplo, asignación de recursos y estrategias de apoyo. Además de lo anterior, las evaluaciones censales ponen a disposición de cada establecimiento información sobre sus logros en distintas áreas y niveles, aportando datos para la toma de decisiones de gestión de los equipos directivos y docentes.

Por último, cabe mencionar que, a diferencia de otros países, Chile no cuenta con evaluaciones o exámenes externos de carácter certificativo (de promoción) de los y las estudiantes, es decir, con consecuencias a nivel individual. La evaluación de los aprendizajes, tanto en su función certificativa (también llamada sumativa) como en su función propiamente de fortalecimiento del aprendizaje (también llamada formativa), está en manos de los equipos docentes de los establecimientos. En este contexto, el Mineduc reconoce que la información que los mismos establecimientos educacionales generan para tomar decisiones de gestión y pedagógicas es muy relevante para los distintos actores, pero ha sido poco visibilizada y apoyada, por lo que se estima fundamental desarrollar capacidades en los equipos directivos y docentes para la obtención, interpretación y uso de información sobre los aprendizajes para definir e impulsar mejoras. Esta necesidad de fortalecer y apoyar la evaluación que realizan los y las docentes en aula -y la reflexión y uso que se hace de ella- es reconocida por ellos y ellas, y ha sido evidenciada por los resultados de la evaluación docente en nuestro país. Aún más, esta necesidad fue una de las principales conclusiones del Equipo de Tarea (2014), el cual recomendó “diseñar e implementar una política amplia para potenciar la evaluación de aprendizajes que realizan docentes y directivos al interior de los establecimientos escolares y complementarla con las mediciones externas”.

Lo anterior ha llevado al Mineduc y la Agencia de Calidad a desarrollar una **Política y Plan de Fortalecimiento de la Evaluación en Aula** que brinde un marco orientador y articulador de los distintos actores del sistema que participan en procesos relacionados con la evaluación de aprendizajes en aula, y que genere orientaciones, estrategias y herramientas para reforzar estas capacidades en los equipos docentes. Se espera con esto fortalecer las competencias evaluativas de los equipos, tanto para recoger información cada vez más pertinente y analizarla de modo más preciso y pedagógicamente útil, como para poner en juego, de

forma complementaria, la información externa y la interna, de modo de tomar decisiones bien fundadas y con pertinencia local, lo que ha sido ampliamente sustentado como un aspecto crítico para la sustentabilidad de la mejora educativa (OECD, 2013). En esta línea, se espera apoyar las capacidades que tienen los y las docentes para recoger evidencias de los aprendizajes logrados por cada estudiante y, a partir de ellas, hacer ajustes pedagógicos oportunamente (es decir, usarlas formativamente), algo que se ha identificado como una práctica docente crucial para favorecer el progreso del aprendizaje de los y las estudiantes (Black & Wiliam, 1998; Hattie, 2003, Griffin, 2014).

La literatura indica la importancia de la evaluación formativa como una de las acciones más efectivas para la mejora de los aprendizajes de las y los estudiantes. Black y Wiliam (1998) informan que el típico tamaño efecto está en los rangos de 0.4 a 0.7, mientras que Hattie y Timperley (2007) reportan un gran efecto en los aprendizajes de los y las estudiantes (tamaño efecto promedio de 0.95 desviaciones estándar en cerca de 4.000 estudios). Así, se ha identificado el uso formativo de la evaluación como una de las prácticas más potentes que son utilizadas en las pedagogías efectivas.

Es por esto que el Ministerio de Educación y la Agencia de Calidad proponen avanzar gradualmente en una línea de desarrollo que promueva la evaluación formativa en el sistema educacional. Para fortalecer las capacidades evaluativas de las y los docentes y otros actores de la comunidad educativa, se está trabajando en las siguientes líneas de acción:

- Orientaciones para la evaluación del currículum por parte de docentes, con fines principalmente formativos (es decir, con foco en retroalimentar el aprendizaje de los y las estudiantes, y las prácticas pedagógicas), pero también considerando orientaciones para que los procesos de calificación – sumativos o certificativos– sean más válidos, tengan mayor significado y apoyen de mejor manera el aprendizaje.
- Actualización de las normas de calificación y promoción, y ejemplos de reglamentos de evaluación con sustento y utilidad pedagógica.
- Convenios de colaboración con instituciones de formación continua y un programa de desarrollo profesional docente en evaluación en aula que contemple cursos de creciente profundización.
- Orientaciones y acciones para el fortalecimiento de la formación inicial docente (por ejemplo, por medio de Convenios de Desempeño).
- Trabajo con Seremías y Deprov sobre las orientaciones y recursos de apoyo que perciben como necesarias para los sostenedores y los establecimientos educacionales.
- Generación de espacios de trabajo colaborativo entre docentes para el diseño, análisis y uso de evaluaciones internas y externas.
- Coordinación de las instituciones del SAC en torno a las iniciativas para apoyar el uso de la evaluación para la mejora.
- Seminarios y otras instancias de formación, intercambio y difusión.

- Herramientas y recursos para apoyar la evaluación que llevan a cabo los y las docentes. Entre ellos, la Agencia desarrollará lo siguiente:
 - o Actividades evaluativas con foco en lo formativo: si bien, por un lado, todos los resultados de una medición podrían usarse de manera formativa, hay ciertas actividades de aprendizaje que son más propicias para hacer un uso formativo de los datos obtenidos. Las actividades serán dispuestas libremente a los y las docentes para que puedan elegir aquellas que consideren necesarias para el contexto del establecimiento y de acuerdo a la planificación realizada, de manera de facilitar su uso, y tendrán diferentes formatos tanto de instrumentos como de preguntas.
 - o Guías y materiales interpretativos para las y los docentes: independientemente de la calidad de las actividades evaluativas entregadas, es necesario apoyar y orientar el uso formativo. Para esto, se propone entregar guías y materiales que orienten a la profesora o el profesor en el uso pedagógico de los datos obtenidos.

Entre las herramientas y recursos de apoyo que pondrá a disposición el Mineduc se encuentran las siguientes:

- o Recursos audiovisuales y materiales auténticos comentados de prácticas e interacciones en aula en que se evidencia un uso formativo de la evaluación, para diferentes cursos y asignaturas.
- o Ejemplos del uso de la retroalimentación en aula.
- o Actividades de evaluación asociadas a los Objetivos de Aprendizaje de las Bases Curriculares, considerando una diversidad de instrumentos y momentos del proceso de enseñanza-aprendizaje (por ejemplo, actividades o preguntas para levantamiento de conocimientos previos y actividades para evaluar el logro de cierto objetivo).
- Incentivos a la investigación en este tema y realización de estudios, con el fin de, entre otras cosas, monitorear el desarrollo de las capacidades de evaluación de los y las docentes y levantar información para focalizar y mejorar los apoyos. En este sentido, tanto la Agencia de Calidad como el Mineduc desarrollarán estudios para indagar aspectos tales como:
 - o qué tipos de mediciones, constructos o conocimiento disciplinar necesitan las y los docentes para profundizar sus prácticas de evaluación formativa;
 - o qué tipo de información necesitan los y las docentes para interpretar y analizar los datos de manera formativa;
 - o qué tipo de prácticas docentes son las facilitadoras para implementar la evaluación formativa en la sala de clases;
 - o qué tipo de retroalimentación necesitan los y las docentes para mejorar la calidad y frecuencia de evaluaciones formativas;
 - o qué tipo de experiencias favorecen el desarrollo de habilidades de evaluación en el proceso de formación inicial de docentes;

- cuáles son las prácticas de evaluación declaradas por docentes y cómo estas son más o menos favorables al desarrollo de evaluación formativa;
- qué aspectos perciben como necesarios de mejorar respecto de las orientaciones y recursos de apoyo con los que cuentan para facilitar el trabajo evaluativo, entre otras preguntas.

Las respuestas a estas preguntas apuntan a mejorar las acciones de orientación y apoyo desplegadas y también a entregar recomendaciones para los cursos de desarrollo profesional que se diseñen.

Por último, otro de los desafíos a los que se enfrentan los sistemas nacionales de evaluación es resguardar la coherencia o articulación entre la evaluación realizada externamente y aquella llevada a cabo al interior de los establecimientos educativos, articulando ambas de modo de facilitar que los y las docentes comprendan en mayor profundidad las expectativas curriculares nacionales y tengan acceso a diferentes tipos de evidencia de los aprendizajes que permiten inferir distintos grados de logro de estas expectativas, de modo de poder ir generando oportunidades de aprendizaje y procesos de evaluación al interior del aula que favorezcan el aprendizaje de cada estudiante. El presente Plan propone un nuevo tipo de herramienta de evaluación para comenzar a construir este puente entre la evaluación externa y lo que se realiza al interior de los establecimientos, ofreciendo al sistema pruebas externas para ser usadas internamente.

d) Elementos diagnósticos que sustentan el Plan

En el marco de una comprensión integral de evaluación en sus distintos niveles –como ha sido planteado en el punto anterior–, el Ministerio de Educación valora contar con un sistema de evaluación de aprendizajes a nivel del sistema educativo y de establecimientos educacionales, con una trayectoria en la que se han ido estableciendo mejoras tanto en su diseño como en la entrega de resultados y despliegue de estrategias cada vez más facilitadoras de procesos de reflexión y toma de decisiones para el mejoramiento educativo, y que cuenta con una solidez que es ampliamente reconocida por distintos actores. Efectivamente, y como parte de la discusión del Equipo de Tarea que sesionó durante el año 2014, se recogió, en las distintas audiencias y entre los miembros de la Mesa, una visión bastante consensuada respecto a la valoración de la solidez técnica del SIMCE y a la relevancia de su rol para la evaluación de aprendizajes a nivel nacional, especialmente, para monitorear los logros y priorizar apoyos cuando sea necesario. Al mismo tiempo, se identifica una visión también compartida en cuanto a la preocupación por los usos y por el énfasis en el control más que en el soporte de la información que genera la evaluación para orientar los procesos de mejora.

En términos de la valoración de una evaluación de aprendizajes a nivel nacional, se reconoce que esta se funda en la convicción profunda de que todos los y las estudiantes pueden aprender, y en el marco de garantizar una formación de calidad con equidad, por lo que contar con evaluación externa nacional ha

permitido relevar la importancia de resguardar que todos y todas aprendan. En esta misma línea, se considera positivo para el sistema educativo incluir evaluaciones censales que permiten generar información sobre y para cada establecimiento escolar. Este monitoreo a nivel nacional ha posibilitado, a su vez, que existan evaluaciones de la efectividad de iniciativas y políticas estatales, así como el desarrollo de investigación y conocimiento que aporta al mejoramiento del sistema educativo. Además, ha provisto de información a los establecimientos educativos y sostenedores que complementa la que ellos mismos generan, para tomar decisiones de gestión curricular en relación con referentes nacionales.

El Ministerio de Educación también reconoce que ha habido importantes avances en la forma que se entregan los resultados, pues se provee información cada vez más pertinente para los diferentes usuarios, en particular, para quienes trabajan en los establecimientos escolares. La complementación del SIMCE con herramientas como los Otros Indicadores de Calidad y los Estándares Indicativos de Desempeño en la evaluación externa de los establecimientos se considera una oportunidad para entregar información más diversa en términos tanto de resultados como de procesos educativos que forman parte de una educación integral y de calidad.

Sin embargo, el Mineduc identifica áreas que se deben mejorar, en particular, respecto de los usos que se hace de los resultados de la evaluación externa en el país, tanto a nivel de política educativa como a nivel del sistema escolar mismo y el público general. El Mineduc coincide con el diagnóstico del Equipo de Tarea, que revisó el SIMCE en 2014, respecto del desequilibrio entre la expansión del sistema de evaluación y la presión a los establecimientos (debido a sus consecuencias y los usos indebidos que se han ido identificando), por una parte, y la necesidad de fortalecer las políticas y estrategias de apoyo y desarrollo de capacidades, por otra. Los principales problemas identificados por el Equipo de Tarea (2014) en relación con el SIMCE son los siguientes:

- a. El excesivo foco en los resultados de las mediciones externas estrecha la visión sobre la calidad de la educación.
- b. Hay más pruebas externas e información disponible que preparación y oportunidades en los establecimientos escolares para aprovecharlas.
- c. La publicación de *rankings* basados en puntajes que no consideran las características socioeconómicas de los y las estudiantes, entre otras variables que afectan los aprendizajes, induce a interpretaciones erróneas y comparaciones injustas sobre la calidad del trabajo realizado por los establecimientos, particularmente, para establecimientos que trabajan en contextos vulnerables. Además, esto fomenta lógicas competitivas por sobre colaborativas (esenciales para el mejoramiento).
- d. Existe falta de estudio y monitoreo sistemático de los usos reales (deseados e indeseados) que se le da al SIMCE, así como de medidas para evitar aquellos indebidos.

- e. La evaluación realizada por los y las docentes en las salas de clase ha sido, hasta ahora, invisible para nuestro sistema nacional de evaluación de aprendizajes y desatendida por las políticas educativas.

En la sección siguiente se presentan las decisiones y acciones que ha tomado el Ministerio de Educación y la Agencia de Calidad basándose en este diagnóstico.

e) Decisiones y acciones en función del diagnóstico

Como respuesta al diagnóstico anterior respecto de aquellos aspectos a mejorar, el Ministerio de Educación y la Agencia de Calidad proponen llevar a cabo las siguientes acciones:

- a. Desarrollar un Plan de Evaluaciones de Aprendizajes que considere diferentes tipos de información (censal, muestral y evaluaciones externas para ser aplicadas internamente) en función de sus propósitos y usuarios particulares, basándose en una visión integradora y articulada de un sistema nacional de evaluación de aprendizajes que genere información a distintos niveles y con diferentes fines, para apoyar la mejora continua.
- b. Seguir mejorando la forma de entregar la información a los establecimientos educativos, para facilitar su uso en la toma de decisiones de gestión curricular, y a otros actores (instituciones formadoras de docentes, Mineduc, secretarías regionales y provinciales de educación, sostenedores, medios de comunicación, entre otros), evitando el uso de *rankings* y explicitando interpretaciones y usos intencionados e indeseados de la información.
- c. Desarrollar estrategias para el monitoreo de los usos que se hacen de la información y las herramientas que se entregan.
- d. Generar mayores apoyos para el desarrollo de capacidades tanto para utilizar la información que proviene de las mediciones externas como para fortalecer los procesos de evaluación internos de los establecimientos, a nivel de la gestión escolar y de los aprendizajes de los y las estudiantes. Como parte de esta línea de trabajo, se está desarrollando una Política de Fortalecimiento de la Evaluación en Aula, además de los avances de la Agencia respecto de estrategias de comunicación de resultados que facilitan su interpretación y uso para la toma de decisiones en función de la mejora.
- e. Incentivar una cultura de evaluación en la que haya un compromiso y responsabilización por los aprendizajes de los y las estudiantes por parte de todos los actores educativos involucrados y en la que no se actúe solamente por la búsqueda de recompensas o la evitación de consecuencias negativas.

f) *Propósitos de la evaluación externa de aprendizajes a nivel de establecimientos y de sistema escolar.*

Como se mencionó, la revisión realizada por la OECD en 2013 sobre sistemas de evaluación en el mundo reconoce como meta última de estos el mejoramiento educativo. El progreso hacia esta meta puede ser logrado a través de dos grandes mecanismos o vías, la responsabilización o rendición de cuentas (*accountability*) y el desarrollo. La rendición de cuentas o responsabilización se concibe como una forma de asignar o tomar responsabilidad por los resultados educativos, específicamente, por los aprendizajes que logran los y las estudiantes. En general, se utiliza el término para atribuir responsabilidad a los establecimientos educacionales, pero es posible responsabilizar también a otros actores del sistema escolar, como sostenedores de agrupaciones de establecimientos o el Estado (en términos de rendir cuentas respecto de la efectividad de políticas, programas u otras estrategias). Los sistemas de aseguramiento de la calidad basados en la responsabilización suelen enfatizar la rendición de cuentas pública de los resultados obtenidos y frecuentemente se le asocian, además, otro tipo de consecuencias por los resultados, generalmente, en la forma de sanciones o incentivos (por ejemplo, amonestar o cerrar escuelas por bajos resultados y entregar recursos por buenos resultados). Aquellos sistemas que les asocian sanciones o recompensas a los establecimientos educacionales por sus resultados se conocen como *sistemas de evaluación de altas consecuencias*. Por su parte, el desarrollo se entiende como el despliegue de mecanismos de apoyo y fortalecimiento de las capacidades de los equipos de los establecimientos educacionales en función de mejorar las prácticas pedagógicas y de gestión que llevan a mayores logros de aprendizaje. Así, al identificar las fortalezas y debilidades a nivel nacional y de cada uno de sus establecimientos escolares, se genera información para alimentar decisiones de mejoramiento o desarrollo en todos los niveles del sistema (Equipo de Tarea, 2014).

En el informe de la OECD (2013), Chile aparece como uno de los dos países (de los 25 revisados) en que se pretende que la información de la evaluación externa nacional tenga alto uso tanto para propósitos de desarrollo como para propósitos de rendición de cuentas⁴. Dado que las evaluaciones con uno y otro fin suelen tener requerimientos técnicos y prácticos diferentes, es difícil que un solo tipo de instrumento o estrategia de evaluación sirva de buena forma para ambos fines. Reconociendo esto, y en línea con las recomendaciones del Equipo de Tarea (2014), a continuación se describen los propósitos del sistema de evaluación externa de aprendizajes para, en secciones subsiguientes, especificar cómo dicho sistema propone hacerse cargo de ambos propósitos de forma integrada. Estos propósitos se enmarcan en la necesidad de ampliar la mirada de la calidad, reducir la presión y recarga sentida por la cantidad de evaluaciones censales, promover una cultura de evaluación en que esta se entienda como un proceso de generación de información para la mejora, e incentivar un compromiso responsable por parte de las comunidades educativas, sostenedores y el Estado.

⁴ Sin embargo, en entrevista con Paulo Santiago, investigador a cargo de este análisis para Chile, señala que predomina la rendición de cuentas o *accountability*.

Propósitos de *responsabilización* del SIMCE:

1. Contribuir a promover el compromiso profesional de las comunidades educativas con el aprendizaje de sus estudiantes.
2. Responsabilizar al Estado como garante de que en cada uno de los establecimientos escolares se logren aprendizajes de calidad.
3. Informar a los padres y madres, para contribuir a que se comprometan con la educación de sus hijos e hijas, se involucren en las escuelas y puedan tomar decisiones con conocimiento.

Propósitos de *desarrollo* del SIMCE:

1. Proveer información para monitorear el estado y avance de los logros de aprendizaje a nivel nacional, y así orientar decisiones de política educacional.
2. Proveer información para orientar los dispositivos del Sistema de Aseguramiento de la Calidad para apoyar a los establecimientos escolares, e identificar en forma precisa a aquellos que más lo necesitan.
3. Proveer información para guiar decisiones de gestión curricular⁵ por parte de los equipos directivos y técnico-pedagógicos en los establecimientos escolares y sus sostenedores.
4. Proveer información sobre los niveles de aprendizaje de los y las 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 las y los docentes⁶.

En la siguiente sección, se presenta el Plan de Evaluaciones Nacionales (censales y muestrales) e Internacionales.

⁵ Los Estándares Indicativos de Desempeño hacen referencia a la gestión curricular como las políticas, procedimientos y prácticas que llevan a cabo el director, el equipo técnico-pedagógico y los y las docentes del establecimiento para coordinar, planificar, monitorear y evaluar el proceso de enseñanza-aprendizaje.

⁶ Las mediciones externas de resultados de aprendizaje aportan información a los y las docentes sobre los errores conceptuales o procedimentales más comunes de los y las estudiantes en ciertas áreas, o sobre el nivel logrado al término de un ciclo escolar. Sin embargo, esta información es global y no está referida específicamente a cada estudiante o a unidades del Programa de Estudio que estén abordando. En este sentido, se considera que la función más propiamente formativa de informar cotidianamente las decisiones pedagógicas de los y las docentes respecto de los avances de cada estudiantes se cumple, principalmente, por medio de la evaluación realizada internamente en los establecimientos escolares, mientras que el SIMCE entrega información para apoyar la gestión curricular general a nivel escuela.

III. Evaluaciones externas de aprendizajes para el periodo 2016-2020

En esta sección se presenta el Plan de Evaluaciones. Con el fin de cumplir con todos los propósitos, se ha diseñado un plan de evaluaciones que pretende responder a los distintos propósitos por medio de diferentes herramientas de evaluación. En este sentido, se presentan los criterios utilizados para diseñar el Plan y, luego, los calendarios y la fundamentación de las evaluaciones censales, las evaluaciones muestrales nacionales y la participación de Chile en estudios internacionales⁷. En la siguiente sección, y de modo complementario al Plan, se presenta una nueva herramienta de evaluación basada en Estándares de Aprendizaje que se pone a disposición de los establecimientos y se entrega información sobre la evaluación de los Otros Indicadores de Calidad. Todas estas instancias evaluativas ayudan a cumplir diferentes propósitos del sistema, aportando información relevante para la mejora a distintos niveles.

Criterios utilizados para diseñar el Plan Nacional de Evaluaciones

Con el propósito de promover que el Plan de Evaluaciones, además de cumplir con los requerimientos legales establecidos en las leyes N° 20.370 (Ley General de Educación) y N° 20.529 (Sistema Nacional de Aseguramiento de la Calidad de la Educación parvularia, básica y media y su fiscalización), favorezca efectivamente el mejoramiento de la calidad de la educación, se han definido los siguientes criterios para orientar su definición:

1. El Plan de Evaluaciones debe ser coherente con la propuesta curricular vigente. Tal como se ha planteado anteriormente en este documento, la definición de un plan nacional de evaluaciones tiene como finalidad contar con un sistema articulado y consensuado que provea información respecto de los aprendizajes desarrollados por los y las estudiantes, de modo de orientar la toma de decisiones, tanto a nivel de políticas públicas como a nivel de gestión escolar.

El instrumento que, por definición, operacionaliza los Objetivos Generales de la Educación y, por tanto, define los aprendizajes a desarrollar por todos y todas las estudiantes durante su escolaridad es el Currículum Nacional. En esta lógica, el Plan de Evaluaciones debe ser coherente con lo planteado por el Currículum Nacional, especialmente, al definir las asignaturas o áreas a evaluar y los momentos en

⁷ En el caso de las evaluaciones internacionales, el plan fue aprobado para el periodo 2011-2020. En el presente documento, se presentan las fechas de aplicación actualizadas de acuerdo a información reciente de los consorcios internacionales.

que es conveniente tener información respecto de los aprendizajes que se están logrando. Además, el Plan debe ser consistente con el concepto de calidad de la educación que se encuentra en la Ley General de Educación y se cristaliza en el currículum vigente, de manera de dar señales claras y consistentes al sistema escolar y de entregar información relevante para la toma de decisiones orientadas a mejorar la calidad de la educación. Esto implica ir progresivamente diversificando los distintos tipos de evaluaciones (censales, muestrales y otras), de modo de ir generando información de una amplia gama de áreas entendidas como parte de una formación integral de calidad. Asimismo, en términos de los procesos de diseño y construcción de los instrumentos, implica fortalecer los procesos de trabajo entre la Agencia de Calidad y la Unidad de Currículum y Evaluación del Ministerio de Educación.

2. El Plan de Evaluaciones debe ajustarse al ciclo de implementación curricular. En este sentido, debe proveer información respecto de cuánto están aprendiendo los y las estudiantes en relación con lo que se espera que aprendan de acuerdo con el Currículum Nacional. Por ello, el Plan debe considerar el estado de implementación de ajustes del currículum, de manera de aplicar evaluaciones, idealmente, a las cohortes que han tenido la oportunidad de aprender con dicho currículum en la mayor parte de los cursos incluidos en la evaluación. Lo anterior implica el alineamiento del Plan de Evaluaciones con la calendarización de la implementación de ajustes y cambios curriculares, lo que es especialmente sensible en momentos de transición curricular. En aquellos casos de transición curricular en los que se deba aplicar una evaluación, ésta se elaborará incorporando solo la evaluación de los conocimientos y habilidades que son comunes a ambos currículos.
3. El Plan de Evaluaciones debe articularse con las definiciones adoptadas respecto de la elaboración de Estándares de Aprendizaje. El Sistema Nacional de Aseguramiento de la Calidad de la Educación incorpora los Estándares de Aprendizaje como una herramienta de política pública que permite evaluar qué tan adecuados, elementales o insuficientes son los aprendizajes alcanzados por las y los estudiantes en relación con lo planteado en el currículum vigente, en diferentes momentos de su trayectoria escolar. Los Estándares de Aprendizaje se elaboran para aquellas asignaturas y grados en los que se considera fundamental contar con información respecto de los aprendizajes alcanzados, de manera que los establecimientos puedan responsabilizarse por los logros de sus estudiantes, ya sea por medio de la Ordenación u otra vía. En esta lógica, es necesario que el Plan de Evaluaciones entregue al sistema información respecto del logro de los Estándares de Aprendizaje, por lo que debe articularse con las asignaturas y grados que cuentan con este tipo de estándares y con el calendario de elaboración de futuros estándares.

4. El Plan de Evaluaciones debe permitir que se lleve a cabo una clasificación en categorías de desempeño de establecimientos de educación básica y media que sea confiable. El actual Sistema Nacional de Aseguramiento de la Calidad introduce la Ordenación de los establecimientos, la cual consiste en el proceso de categorización de las escuelas y liceos tomando en cuenta los resultados de aprendizaje de todas las áreas evaluadas censalmente en las mediciones nacionales y la distribución de los resultados obtenidos en relación con los Estándares de Aprendizaje, junto con el grado de cumplimiento de los Otros Indicadores de Calidad educativa, considerando el nivel de vulnerabilidad de los y las estudiantes. Lo anterior se traduce en que el Plan de Evaluaciones debe contar con una cantidad, variedad y periodicidad de evaluaciones censales que entreguen información suficiente para llevar a cabo una clasificación u ordenación confiable de los establecimientos. Además, considerando que es necesario velar por el buen uso de los recursos públicos y privilegiar espacios para que los establecimientos puedan aprovechar de mejor manera la información que se entrega desde las evaluaciones externas, se ha definido que el Plan debe incorporar la menor cantidad de pruebas censales que se requieran para poder realizar una clasificación u ordenación confiable de los establecimientos.
5. El Plan de Evaluaciones debe permitir una continuidad con el plan histórico, de manera de no perder la capacidad de monitoreo. El SIMCE se ha aplicado por más de 25 años, por lo que en la actualidad se cuenta con una serie de datos históricos que permiten visualizar los logros de aprendizaje a lo largo del tiempo. Contar con este tipo de información ha sido altamente valorado, por lo que el Plan debe permitir continuidad en la capacidad de recoger información para las series de datos históricos de 4° básico, 8° básico y 2° medio. En lo que respecta a la participación de nuestro país en estudios internacionales, Chile ha participado históricamente en algunos de ellos, por lo que el Plan también debe resguardar la continuidad respecto de obtener información que permita realizar análisis de política comparada.
6. El Plan de Evaluaciones debe considerar la evaluación de estudiantes con necesidades educativas especiales⁸. La inclusión es considerada fundamental dentro del rol garante del Estado de una educación de calidad, por lo que el Plan de Evaluaciones debe ser inclusivo y contemplar la evaluación con acomodaciones para las y los estudiantes que poseen necesidades educativas especiales.
7. El Plan de Evaluaciones debe permitir la realización de estudios de indicadores de progreso y/o valor agregado, en línea con la recomendación de la Comisión que revisó el SIMCE en 2003 y lo establecido

⁸ Esto fue relevado por la Comisión que revisó el SIMCE en el año 2003.

en la Ley N° 20.529. La posibilidad de incorporar el estudio de estos indicadores se ha venido evaluando largamente en cuanto a su factibilidad técnica. La realización de este tipo de estudios se ha considerado como algo deseable, por lo cual se ha definido que el Plan de Evaluaciones sea diseñado de manera que deje abierta la posibilidad de incorporarlos.

8. El Plan de Evaluaciones debe articular el calendario de las evaluaciones nacionales con la participación de Chile en estudios internacionales, para así aprovechar de mejor manera los recursos y la información que se puede obtener de ambos tipos de estudios.

A continuación se presenta el diseño de evaluaciones censales, muestrales e internacionales para los años 2016 a 2020, junto con la justificación de cada uno de estos diseños. Considerando el Plan global, es decir, las pruebas censales, las muestrales y las internacionales, se desarrolla una fundamentación respecto a la forma en que este da cuenta de los criterios planteados.

Propuesta y fundamentación de evaluaciones censales

A continuación se presenta el calendario de evaluaciones censales, aprobado por el Consejo Nacional de Educación, y su fundamentación. Para este se ha considerado la estructura de enseñanza básica y media vigente en la actualidad (8 + 4), la que, conforme a lo establecido en la Ley General de Educación, debería cambiar a modalidad de 6 + 6 años, respectivamente, a partir del año 2018. El Ministerio de Educación se encuentra evaluando la postergación de este cambio, lo que, de suceder, no implicaría modificaciones significativas para el calendario de evaluaciones propuesto (a excepción de la prueba muestral de Educación Física y Salud, que debería cambiarse a 6° básico).

Tabla N° 1: Calendario de evaluaciones censales⁹

CURSO	ÁREA	2016	2017	2018	2019	2020
4° básico	Lectura*	X	X	X	X	X
	Matemática*	X	X	X	X	X
6° básico	Lectura*	X		X		X
	Escritura	X		X		X
	Matemática*	X		X		X
	Cs. Naturales			X		
	Cs. Sociales	X				X
8° básico	Lectura*		X		X	
	Matemática*		X		X	
	Cs. Naturales		X			
	Cs. Sociales				X	
2° medio	Lectura	X	X	X	X	X
	Matemática	X	X	X	X	X
	Cs. Naturales	X		X		X
	Cs. Sociales		X		X	

(*) Pruebas que incluirán acomodación para estudiantes con necesidades educativas especiales permanentes (discapacidad visual y auditiva).

⁹ Cada año la Agencia realiza su solicitud presupuestaria sobre la base de poder implementar el plan de evaluaciones vigente; no obstante, la asignación real de recursos es la que se aprueba en el contexto de la Ley de Presupuestos.

Fundamentación de evaluaciones censales

Como se mencionó anteriormente, las evaluaciones censales de aprendizaje permiten contar con información tanto a nivel del sistema educativo en general como de los establecimientos, en particular. Así, brindan evidencia a las escuelas sobre los logros de aprendizaje de sus estudiantes, de forma de identificar qué áreas hay que fortalecer.

A modo de descripción general, el calendario mantiene las cinco áreas curriculares evaluadas en cuatro niveles, reduciendo su frecuencia (para disminuir la sobrecarga a los establecimientos, para incrementar las probabilidades de que se use la información y para evitar efectos indeseados de sobreentrenamiento). Se mantienen evaluaciones anuales para 4° básico y 2° medio, asegurando así que todas las cohortes de estudiantes cuenten con información al menos una vez en enseñanza básica y en enseñanza media. En este Plan, la prueba censal de Lectura en 2° básico implementada en el periodo 2011-2015 se reemplaza por una evaluación muestral periódica y una nueva herramienta de evaluación externa para uso interno, referida a Estándares de Aprendizaje, que se pondrá a disposición de los establecimientos (para mayor detalle, ver apartado "Propuesta y fundamentación de evaluaciones externas para uso interno", en esta sección). Se alterna la evaluación de Ciencias Naturales y Ciencias Sociales en 6° y 8° básico, de modo que toda cohorte de estudiantes sea evaluada al menos una vez en su trayectoria escolar en cada área. Además, se amplía la cantidad de niveles en que se evalúan a estudiantes con discapacidad sensorial (visual y auditiva), para así fomentar una mayor inclusión.

Propuesta y fundamentación de evaluaciones muestrales

A continuación se presentan las evaluaciones muestrales y su fundamentación.

Tabla Nº 2: Calendario de evaluaciones muestrales

CURSO	ÁREA	2016	2017	2018	2019	2020
2° básico	Lectura		X		X	
8° básico	Educación Física y Salud			X		
	Formación Ciudadana		X			
3° medio	Inglés ¹⁰		X			X
4° medio	Competencias genéricas EMTP					X

Fundamentación de evaluaciones muestrales

Las evaluaciones muestrales son una herramienta esencial del sistema nacional de evaluación de aprendizajes, puesto que generan información precisa y diversa para monitorear los logros del sistema educativo, lo que permite al Estado rendir cuentas sobre la efectividad de sus políticas y programas, fomentando así su responsabilización por estas y su compromiso con una constante mejora en la calidad. Además, estas evaluaciones permiten focalizar recursos, posibilitan identificar las fortalezas y aspectos a mejorar (por ejemplo, errores comunes que cometen los y las estudiantes en cada área) a nivel nacional, lo que puede informarse a instituciones formadoras de docentes, a instituciones y organismos encargados de desarrollo profesional continuo docente, y a sostenedores y escuelas, para así entregar orientaciones sobre hacia dónde dirigir sus esfuerzos, así como proponer estrategias para fortalecerlos.

El calendario de evaluaciones muestrales incorpora evaluaciones en áreas que se consideran relevantes para la formación de los y las estudiantes. Algunas de ellas se evaluaban de modo censal en el plan anterior y en el presente plan se incorporan como evaluación muestral, de forma de dar continuidad a la información sobre estas áreas en los niveles evaluados y de reducir la sobrecarga de evaluaciones a los establecimientos.

¹⁰ Los instrumentos de evaluación utilizados para realizar este estudio son diseñados por instituciones externas a la Agencia con amplia experiencia y trayectoria en la evaluación del inglés como idioma extranjero.

Considerando que el propósito de las evaluaciones muestrales es contar con información para el monitoreo del sistema educativo, se propone una periodicidad general trienal, de tal forma de poder dejar un tiempo suficiente –pero no muy extenso– entre una aplicación y otra para evidenciar posibles cambios.

A continuación se describen las evaluaciones del calendario muestral.

Lectura en 2º básico

La prueba SIMCE (censal) de Lectura para 2º básico se reemplaza por una prueba muestral y un instrumento de evaluación externa referido a Estándares de Aprendizaje que pasará a ser administrado, analizado y utilizado autónomamente por los equipos profesionales de los propios establecimientos educacionales (ver descripción de este instrumento en la sección sobre evaluaciones externas para uso interno). Este cambio se sustenta en los siguientes argumentos:

1. *La necesidad de avanzar hacia un mejor balance entre la presión asociada a las mediciones externas y las estrategias para el desarrollo de capacidades de evaluación de los actores educativos, especialmente, los y las docentes.* Para esto, se reduce la frecuencia y número de pruebas externas, comenzando precisamente con 2º básico, en atención a que este es el nivel en que el Equipo de Tarea (2014) observó mayor ambivalencia y reparos, por tratarse de estudiantes menores que participan en un procedimiento de aplicación que incluye examinadores externos a las escuelas, desconocidos para los niños y las niñas.
2. *La importancia de aportar al desarrollo de una cultura evaluativa al interior de las escuelas.* El cambio propuesto permite al cuerpo docente hacerse responsable de aplicar esta evaluación –definiendo autónomamente los periodos más oportunos para administrarla–, interpretar, discutir y analizar sus resultados, compartir con sus colegas los criterios de evaluación en el caso de las preguntas abiertas, formular hipótesis sobre las estrategias pedagógicas más y menos efectivas y definir cursos de acción cuyos resultados podrá monitorear con nuevas aplicaciones de estos instrumentos u otros.
3. *La necesidad de que las escuelas cuenten con información precisa y oportuna sobre el nivel alcanzado por cada estudiante en relación al Estándar de Aprendizaje,* en este caso, en Lectura, a partir de la aplicación que realicen autónomamente los equipos profesionales de los establecimientos. Al ser el mismo equipo docente el que analiza los resultados, sería posible contar inmediatamente con datos para informar sus decisiones pedagógicas. Adicionalmente, esta información puede ser reportada a la comunidad escolar, especialmente, como retroalimentación temprana a las y los estudiantes y como información a los padres y apoderados. El Ministerio de Educación trabajará conjuntamente con la Agencia de Calidad en el desarrollo de orientaciones para el análisis de los resultados y para la comunicación de estos a la comunidad escolar.

4. *La importancia de mantener información externa y periódica sobre los aprendizajes a este nivel, sin consecuencias.* Dado que se ha evidenciado que las evaluaciones externas con consecuencias pueden conllevar efectos indeseados –como, por ejemplo, el estrechamiento curricular, la excesiva preparación para la prueba, la reducción de las formas de evaluar, entre otros–, se estima que contar con este tipo de evaluaciones a edades tempranas puede poner en riesgo la integralidad de la formación y la calidad de las experiencias educativas –por ejemplo, mermándose las oportunidades para desarrollar la creatividad (Kim, 2011)–. Así, para evitar dichos usos y acciones indeseadas sin dejar de contar con información externa, se traslada la evaluación censal a una muestral sin consecuencias y se complementa con una herramienta de uso interno que permita fomentar el compromiso de las comunidades educativas con el logro de los Estándares de Aprendizaje.

La evaluación muestral de Lectura en 2º básico será aplicada bianualmente a una muestra nacional de estudiantes, de modo que se cuente con información temprana a nivel de sistema.

Con las evaluaciones muestrales, en conjunto con el nuevo instrumento referido a Estándares de Aprendizaje, se mantendrá la continuidad de la información a nivel de sistema educativo, y todas las escuelas podrán obtener información con un nivel de riqueza y detalle mayor y en un lapso de tiempo menor, lo que les permitirá tomar decisiones de gestión pedagógica más pertinentes y oportunas.

Educación Física y Salud

Reconociendo que el desarrollo de la Educación Física y Salud es central para fomentar una vida sana para las y los estudiantes, se continuará con el Estudio Nacional de Educación Física realizado en 8º básico¹¹ (establecido en el artículo 5º, inciso 4, de la Ley del Deporte N° 19.712, de febrero de 2001), aplicándolo de forma periódica y muestral, con ciertas modificaciones para un mayor alineamiento curricular, es decir, ampliando lo que se evalúa.

El actual estudio evalúa principalmente el componente de la Condición Física, por lo que se propone ampliarlo para incorporar la evaluación de habilidades motrices y hábitos de vida saludable.

Además, esta evaluación proveerá información complementaria a la que entrega el indicador Hábitos de Vida Saludable, el cual busca informar, en términos generales, respecto del grado en que la escuela o liceo promueve este tipo de hábitos en su gestión.

Formación Ciudadana en 8º básico

¹¹ Para seguir dando cumplimiento con lo estipulado en la Ley del Deporte N° 19.712 esta evaluación será trasladada a 6º básico cuando entre en vigencia la estructura 6 + 6.

Considerando la relevancia que se le asigna a la formación ciudadana en las escuelas, el Plan incorpora la aplicación de una prueba muestral que evalúa este tipo de formación en estudiantes de 8° básico el 2017.

La relevancia de la formación ciudadana en las escuelas se refuerza con los siguientes antecedentes:

- El 2009 se promulga la Ley General de Educación, la cual incorpora Objetivos Generales en el ámbito personal y social, que se relacionan fuertemente con este tipo de formación, tanto para enseñanza básica como para enseñanza media.
- En el currículum vigente la formación ciudadana cuenta con una importante presencia tanto a nivel de los Objetivos de Aprendizaje y Objetivos Fundamentales Transversales como a nivel de los Objetivos de Aprendizaje, Objetivos Fundamentales y Contenidos Mínimos Obligatorios de distintas asignaturas¹².
- El 2013 se introducen los Otros Indicadores de Calidad Educativa, los cuales incluyen el indicador: "Participación y formación ciudadana".
- El 2015 el Ministerio de Educación, siguiendo las recomendaciones del Consejo Asesor Presidencial contra los Conflictos de Interés, el Tráfico de Influencias y la Corrupción, ha propuesto un proyecto de Ley que establece que todos los establecimientos educacionales reconocidos por el Estado deberán incluir un Plan de Formación Ciudadana en los niveles de enseñanza parvularia, básica y media.

En este contexto, contar con información respecto de la formación ciudadana de las y los estudiantes es fundamental para tomar decisiones de política pública orientadas a fortalecer los conocimientos, habilidades y actitudes que se desarrollan en la escuela. Lo anterior, además, es clave dentro del actual contexto en el que se amplía el concepto de calidad de la educación, se considera a la educación de calidad como un derecho y se reconoce el rol garante del Estado en ello.

En esta línea, la evaluación muestral permitirá recoger información que complementará la que se obtiene mediante la aplicación de los cuestionarios SIMCE que evalúan el indicador de calidad "participación y formación ciudadana". Con ello, se podrá contar con información más rica para orientar la toma de decisiones. Además, esta evaluación permitirá profundizar la información que se obtenga de la participación de Chile en el Estudio Internacional para la Educación Cívica y Ciudadana (ICCS).

Inglés en 3° medio

Considerando que el aprendizaje del inglés es fundamental para una ciudadanía global como la que se requiere hoy, el Plan de Evaluaciones contempla una evaluación muestral de Inglés en 3° medio para los años

¹²Referencias a partir de los instrumentos curriculares vigentes en los distintos ciclos.

2017 y 2020, continuando con la evaluación periódica que se realiza desde 2010. Esto permite mantener información sobre esta área en el tiempo, siendo un importante referente para evaluar políticas y programas.

Competencias Genéricas EMTP en 4º medio

En el marco de la inclusión educativa y una mirada amplia de la calidad, se considera relevante generar información para política pública en áreas que han estado desprovistas de ella. En línea con la recomendación del Equipo de Tarea (2014), se incorporará una evaluación muestral de la Educación Media Técnico-Profesional centrada en competencias genéricas claves, tanto para el mundo del trabajo como para acceder a niveles postsecundarios de formación.

Esta evaluación muestral se aplicará cada tres años y se definirá el mejor momento del año para su aplicación, de modo de no interferir con los procesos educativos de los y las estudiantes de esta modalidad de enseñanza.

En términos curriculares, las competencias de carácter general o genéricas son aquellas que todo y toda estudiante de la Educación Media Técnico-Profesional debiese desarrollar, y están relacionadas con las habilidades y actitudes que son necesarias para un desarrollo integral y para insertarse exitosamente en el mundo laboral. Estas son comunes a todas las especialidades, ya que son requeridas en los desempeños de todos los técnicos -y de todos los trabajadores-, independientemente del sector productivo al que esté vinculada la especialidad.

Propuesta y fundamentación de participación en evaluaciones internacionales

A continuación se presentan las evaluaciones internacionales en las que Chile participará y su fundamentación.

Tabla Nº 3: Calendario de evaluaciones internacionales

CURSO	ESTUDIO	2016	2017	2018	2019	2020
3° básico	ERCE: Estudio Regional Comparativo y Explicativo			X		
4° básico	PIRLS: Progress in International Reading Literacy Study	X				
	TIMSS: Trends in International Mathematics and Science Study				X	
6° básico	ERCE: Estudio Regional Comparativo y Explicativo			X		
8° básico	ICCS: International Civic and Citizenship Study	X				X
	ICILS: International Computer and Information Literacy Study			X		
	TIMSS: Trends in International Mathematics and Science Study				X	
2° medio	PISA: Programme for International Student Assessment (Lectura, Matemática y Ciencias Naturales, con foco en Lectura) ¹³			X		

ERCE: Estudio Regional Comparativo y Explicativo (UNESCO).

PIRLS: Estudio Internacional de Progreso en Comprensión Lectora (IEA).

TIMSS: Estudio Internacional de Tendencias en Matemáticas y Ciencias (IEA).

ICCS: Estudio Internacional sobre Educación Cívica y Formación Ciudadana.

ICILS: Estudio Internacional sobre Alfabetización Computacional y Manejo de Información (IEA).

PISA: Programa Internacional para la Evaluación de Estudiantes (OCDE).

Fundamentación de evaluaciones internacionales

Los estudios internacionales cumplen una función importante de monitoreo externo para el sistema educativo, puesto que sus resultados permiten comparar los logros de las y los estudiantes de nuestro país con estándares internacionales y los resultados de otros países, tanto dentro como fuera de Latinoamérica. Al mantener una participación constante en estos estudios, Chile además ha podido ir monitoreando la

¹³ Esta evaluación se aplica a estudiantes de 15 años.

tendencia y los cambios en el tiempo. Adicionalmente, estos estudios generan información sobre los factores asociados a altos logros, lo que permite investigar la relación entre los resultados obtenidos y las políticas educativas, las prácticas de gestión escolar y las prácticas pedagógicas, entre otros. En esta misma línea, posibilitan analizar sistemas educativos que son similares, de forma de identificar qué factores aportan más directamente a la calidad del sistema y al logro de aprendizajes.

Este tipo de monitoreo del sistema permite, además, contar con referentes externos para contrastar y analizar tanto el currículum de nuestro país como la forma en que este se evalúa nacionalmente. Los marcos de evaluación de los estudios internacionales han sido útiles para inspirar las definiciones curriculares en distintos países, incluyendo el nuestro. Por ejemplo, PISA ha sido un fuerte inspirador de reflexiones sobre la actualización del currículum. Actualmente, un trabajo colaborativo entre la UCE y la Agencia de Calidad está permitiendo analizar los ítems de TIMSS 2015 y saber cuáles de ellos son abordados por el currículum nacional (en Matemática y Ciencias Naturales en 4° y 8° básico). De este análisis se podrían derivar investigaciones que provean información valiosa para los procesos nacionales de desarrollo curricular.

El presente plan de evaluaciones internacionales permite dar continuidad al monitoreo del sistema en las áreas y niveles evaluados desde que el país comenzó su participación en los estudios internacionales a fines del siglo pasado. Chile participa en estudios internacionales de diversa índole, tanto con referencia curricular (como los estudios regionales ERCE del LLECE en Lenguaje y Matemática, y las pruebas TIMSS y PIRLS de la IEA) como sin base curricular (en el caso de PISA de la OECD, ICCS e ICILS). La diversidad de áreas en las que se evalúa el país (Lectura, Matemática, Ciencias Naturales, Formación Ciudadana, Alfabetización Computacional, entre otras) permite contar con información rica para monitorear el sistema educativo, sus políticas y programas. Ya sea que se trate de áreas curriculares o no, este tipo de evaluaciones permite conocer y comparar los rendimientos de los y las estudiantes del país en el logro de resultados considerados relevantes por muchos países para participar activamente como ciudadanos.

En el caso del Estudio Internacional de Alfabetización Computacional y Manejo de Información (ICILS, por sus siglas en inglés), se prepara la implementación de un segundo ciclo: ICILS 2018. Chile participó en este estudio en 2013, y se pudo constatar una gran debilidad en el desarrollo de habilidades complejas relacionadas con la alfabetización computacional y manejo de información, en contraste con un fuerte interés y seguridad por parte de las y los estudiantes en el uso de computadores.

Al participar en ICILS 2018, Chile tiene la posibilidad de medir la tendencia de sus resultados por medio de un instrumento desarrollado por expertos mundialmente reconocidos y contrastar estos resultados con los de otros sistemas educativos. Considerando que un ciclo de cinco años es un periodo razonable de tiempo para la medición de una habilidad que se aborda de manera transversal en el currículum, el ciclo de la evaluación internacional resulta adecuado para Chile. Teniendo esta consideración y entendiendo la importancia de evaluar en un contexto internacional habilidades relativas al manejo de información, se propone sustituir la

prueba nacional (muestral) de habilidades TIC para el aprendizaje por el estudio internacional ICILS. Dado que las evaluaciones internacionales coinciden en los cursos en que también se aplica SIMCE (tal como muestra la tabla N°4), se pueden hacer estudios que vinculen ambos tipos de evaluaciones, lo cual es posible gracias a que, desde hace al menos cinco años, se recoge sistemáticamente el RUT de los y las estudiantes en cada aplicación (manteniendo confidencialidad de su identidad). Esto permite parear información en las áreas, los cursos y los años de aplicación en que coincidan evaluaciones nacionales e internacionales. Así, los estudios internacionales contribuyen a enriquecer el tipo de evidencia de los aprendizajes de los y las estudiantes así como los análisis que se hacen sobre ellos y otros factores relacionados, y son una importante fuente de validación externa de nuestras evaluaciones nacionales.

Tabla N° 4: Relación entre evaluaciones nacionales y participación de Chile en Estudios Internacionales

NIVEL	2016	2017	2018	2019	2020
4° básico	SIMCE	SIMCE	SIMCE TIMSS	SIMCE	SIMCE PIRLS
6° básico	SIMCE		SIMCE ERCE		SIMCE
8° básico		SIMCE		SIMCE	
2° medio	SIMCE	SIMCE	SIMCE PISA¹⁴	SIMCE	SIMCE

Con el plan de evaluación de SIMCE 2016–2020 y la calendarización dada por los consorcios que organizan las pruebas internacionales, es posible evidenciar diversas instancias en las cuales se pueden parear los datos, por ejemplo, PIRLS 2021 –que es aplicado en Chile en 2020¹⁵– se puede parear con la evaluación SIMCE de Lectura de 4° básico del año 2020, ERCE –contemplado para 2018– se puede analizar conjuntamente con SIMCE de 6° básico, y TIMSS se puede relacionar con la evaluación SIMCE de Matemática en 4° básico el año 2018.

Propuesta de desarrollos y estudios exploratorios

¹⁴ Esta evaluación se aplica a estudiantes de 15 años.

¹⁵ Las evaluaciones TIMSS, PIRLS e ICCS deben aplicarse en nuestro país el año anterior al que corresponde formalmente.

El desafío de contar con un sistema de evaluación comprensivo demanda un esfuerzo permanente de búsqueda y exploración de nuevas herramientas evaluativas que permitan tanto dar señales más claras respecto a un concepto amplio e integral de calidad como reportar información más diversa y completa sobre el proceso formativo de los y las estudiantes. Por esto, el Ministerio de Educación y la Agencia de Calidad tienen prevista una agenda de desarrollo de nuevos instrumentos de evaluación en áreas curriculares que no han sido abordadas hasta la fecha desde la perspectiva de la evaluación externa.

En primer lugar, se explorarán distintas alternativas metodológicas para la evaluación de la formación artística de las y los estudiantes en Educación Básica, que permitan el monitoreo de los aprendizajes en esta importante área y retroalimenten el Currículo de Artes Visuales y Música. En segundo lugar, se explorarán distintas vías de desarrollo de instrumentos para evaluar otras áreas del aprendizaje escolar, tales como la creatividad, el trabajo colaborativo y otras habilidades en Educación Media, para recoger evidencia acerca de la calidad de los aprendizajes en estos ámbitos transversales del Currículo, que son esenciales para el desarrollo de los y las estudiantes, tanto para su futuro desempeño en lo académico y personal como en lo laboral. En tercer lugar, se espera que estas iniciativas sirvan de insumo para delinear el desarrollo futuro del sistema nacional de evaluación, incorporando nuevas dimensiones que permitan dar cuenta efectiva de una educación integral de calidad.

Análisis del plan según los criterios definidos

A continuación, se describe con mayor detalle la forma en que el Plan de Evaluaciones presentado cumple con los criterios antes definidos.

En relación con el **primer criterio** (coherencia con la propuesta curricular vigente), el Plan permite generar información respecto de los aprendizajes alcanzados por los y las estudiantes en una amplia gama de áreas entendidas como parte de una formación integral de calidad, tanto para el ciclo de enseñanza básica como para el de enseñanza media. Específicamente, el Plan contempla evaluaciones censales para Lectura, Matemática, Ciencias Naturales e Historia, Geografía y Ciencias Sociales en ambos ciclos, así como una para Escritura en el ciclo básico. Lo anterior se complementa con evaluaciones muestrales que integran otras áreas, como Educación Física y Salud y Formación Ciudadana en el ciclo de enseñanza básica, e Inglés y competencias genéricas técnico-profesionales en el ciclo de enseñanza media. Además, se incluye la participación en evaluaciones internacionales para Educación Cívica y Ciudadanía, y para Alfabetización Computacional y Manejo de Información.

Como se describe anteriormente, el Plan incorpora evaluaciones, ya sea censales, muestrales o internacionales, que cubren la mayoría de los Objetivos Generales estipulados en la Ley y operacionalizados en el currículo vigente.

En lo que respecta a los cursos evaluados, el Plan de Evaluaciones contempla evaluaciones en los cursos considerados como fin de ciclo, tanto para la estructura vigente (8 + 4) como para la estructura futura (6 + 6 + 2). Es importante hacer notar que si hubiera aplazamiento de la implementación de la nueva estructura esto no afectaría el Plan de Evaluaciones.

Con respecto al **segundo criterio** (ajuste al ciclo de implementación curricular), el Plan cumple con él en su articulación. Las transiciones curriculares son un desafío de implementación para las escuelas, las que tardan un tiempo en apropiarse de las nuevas propuestas. En este contexto, el Plan de Evaluación se articula con las lógicas de cambio curricular, evitando incorporar evaluaciones durante el primer año de transición curricular en los cursos y asignaturas que corresponda. Esto es especialmente relevante para el caso de las pruebas censales, las que están asociadas a una alta consecuencia para las escuelas.

Para analizar el cumplimiento de este criterio, a continuación se presenta el cuadro de entrada en vigencia de las nuevas Bases Curriculares, las que se comenzarán a implementar el año 2016 para 7° y 8° básico.

Tabla N° 5: Calendario de implementación curricular

Cursos	2014	2015	2016	2017	2018	2019	2020			
NT1	BC	BC	BC	BC Act	BC Act	BC Act	BC Act			
NT2	BC	BC	BC	BC Act	BC Act	BC Act	BC Act			
1° básico	BC 439/433	BC 439/433	BC 439/433	BC 439/433	BC 439/433	BC 439/433	BC 439/433			
2° básico										
3° básico										
4° básico										
5° básico										
6° básico										
7° básico	MC/240	MC/240	BC 614/*	BC 614/*	BC 614/*	BC 614/*	BC 614/*			
8° básico	AC/256	AC/256								
1° medio	MC/220 AC/254	MC/220 AC/254	MC 220 AC 254	BC 452	AC 254 MC 220	BC 452	AC 254 MC 220	BC 452	**FD	**FD
2° medio										
3° medio										
4° medio										

Considerando este ciclo de implementación de un nuevo currículum nacional, el Plan no cuenta con evaluaciones, durante 2016, en ninguna asignatura de 7° u 8° básico, ya que los y las estudiantes de estos

cursos se encontrarán en el primer año de estudio con esta nueva definición curricular. El año 2017, por su parte, los alumnos de 8° básico serán evaluados en Lectura, Matemáticas y Ciencias Sociales, en relación a las bases curriculares vigentes desde el comienzo de su 7° básico. Una situación distinta se dará, sin embargo, en el año 2018, cuando los alumnos de 2° medio sean evaluados en las asignaturas de Lectura, Matemáticas y Ciencias Naturales en relación a un currículum que se encuentra por primera vez vigente en ese curso de la trayectoria escolar. Para abordar esta situación la prueba SIMCE estará referida a aquellos conocimientos y habilidades del currículum que son comunes entre la nueva y la previa definición, conforme a lo explicado en la presentación de este criterio.

En relación con el **tercer criterio** (articulación con las definiciones respecto de la elaboración de Estándares de Aprendizaje), el Plan de Evaluaciones cumple con estar articulado con el plan de desarrollo de Estándares de Aprendizaje.

Actualmente se cuenta con Estándares de Aprendizaje vigentes para 2° básico Lectura, 4° básico Lectura, Matemática y Ciencias Naturales, 8° básico Lectura, Matemática, Ciencias Naturales e Historia, Geografía y Ciencias Sociales, y 2° medio Lectura y Matemática.

A continuación se presenta una tabla que muestra el periodo de vigencia de los Estándares de Aprendizaje que se encuentran actualmente en el sistema.

Tabla N° 6: Vigencia de los Estándares de Aprendizaje

Estándares de Aprendizaje	Fecha de promulgación del decreto	Término de vigencia¹⁶
2° básico: Lectura	Febrero 2014	Febrero 2020
4° básico: Lectura, Matemática y Ciencias Naturales	Mayo 2013	Mayo 2019
8° básico: Lectura, Matemática, Ciencias Naturales e Historia, Geografía y Ciencias Sociales	Mayo 2013	Mayo 2019 ¹⁷
2° medio: Lectura y Matemática	Junio 2015	Junio 2021 ¹⁸

¹⁶ Los Estándares de Aprendizaje tienen una vigencia de seis años. Posterior a esa fecha, se entienden por renovados si no se han decretado nuevos Estándares que los remplacen.

¹⁷ Los Estándares de Aprendizaje de 8° básico deberán ser modificados cuando entren en vigencia las Bases Curriculares de Educación Media para dicho curso.

Tomando en cuenta lo estipulado en la Ley N° 20.529 respecto de la evaluación del cumplimiento de los Estándares de Aprendizaje, el Plan permite una evaluación articulada del cumplimiento con estos Estándares, ya que introduce evaluaciones censales periódicas para la mayoría de los cursos en los que existen Estándares (4°, 8° y 2° medio) y para todas las asignaturas o áreas (Lectura, Matemática, Ciencias Naturales e Historia, Geografía y Ciencias Sociales). De esta manera, el Plan cumple con evaluar de forma censal a lo menos en algún curso, tanto del nivel de enseñanza básica como de enseñanza media.

Además, el Plan presentado se articula con el plan de elaboración y ajuste de Estándares de Aprendizaje del Ministerio de Educación, de manera de incluir evaluaciones censales para los cursos y áreas en las que aún no se han introducido Estándares pero se tiene programado hacerlo.

A continuación se presenta una tabla con el plan de elaboración de Estándares del Mineduc.

¹⁸ Los Estándares de Aprendizaje de 2° medio deberán ser modificados cuando entren en vigencia las Bases Curriculares de Educación Media para dicho curso.

Tabla N° 7: Plan de elaboración de Estándares de Aprendizaje

		Año de elaboración de evidencia	Año de aplicación de la prueba con la que se establecerán los puntajes de corte	Año en el que se llevarán a cabo los ejercicios de puntajes de corte	Año presentación al CNED	Año de promulgación del Decreto y publicación de documentos de difusión	Prueba a la que se le aplicarán los Estándares
4° básico	Lectura Matemática	2016	2017	2018	2018	2019	Prueba 2020 . Resultados referidos a estos Estándares.
6° básico	Lectura Matemática Escritura	2015	2015	2016	2016	2017	Prueba 2016 . Resultados referidos a estos Estándares (retroactivo a 2015).
	Cs. Naturales	2016	2016	2017	2017	2018	Prueba 2018 . Resultados referidos a estos Estándares (retroactivo a 2016).
8° básico	Lectura Matemática Historia, Geografía y Cs. Sociales	2016	2017	2018	2018	2019	Prueba 2019 . Resultados referidos a estos Estándares (retroactivo a 2017).
2° medio	Lectura Matemática Cs. Naturales	2017	2018	2019	2019	2020	Prueba 2020 . Resultados referidos a estos Estándares (retroactivo a 2018).
	Historia, Geografía y Cs. Sociales	2018	2019	2020	2020	2021	Prueba 2021 . Resultados referidos a estos Estándares (retroactivo a 2019).

De la tabla anterior se desprende que el proceso de elaboración de los Estándares mencionados contempla varios años. Debido al enfoque basado en evidencia adoptado, el proceso se inicia con la elaboración de evidencia (proceso que se lleva a cabo en conjunto con la Agencia de Calidad de la Educación) y termina dos o tres años después, una vez que los Estándares han sido aprobados por el CNED, su decreto está totalmente tramitado y se han publicado los documentos de difusión respectivos.

Tal como lo muestra la tabla, se tiene contemplado introducir este tipo de Estándares para 6° básico en Lectura, Matemática y Escritura para que entren en vigencia en 2017 y se puedan reportar resultados referidos a ellos para la prueba rendida en 2016. En el caso de Ciencias Naturales en 6° básico, se tiene planificado introducir estos Estándares para que entren en vigencia en 2018 y se puedan reportar resultados referidos a ellos para la prueba que se aplique ese mismo año.

En lo que respecta a 8° básico, la entrada en vigencia de las Bases Curriculares de 7° y 8° mandata la revisión y ajuste de los actuales Estándares de Aprendizaje. Como se observa en la tabla, se tiene contemplado actualizar los Estándares de Aprendizaje para Lectura, Matemática e Historia, Geografía y Ciencias Sociales para que entren en vigencia en 2019 y se puedan reportar resultados referidos a ellos para la prueba que se aplique ese mismo año.

En el caso de 2° medio, se tiene programado ajustar los Estándares de Aprendizaje de Lectura y Matemática una vez que entren en vigencia las Bases Curriculares para ese curso. Además, se tiene planificado elaborar este tipo de Estándares para Ciencias Naturales y para Historia, Geografía y Ciencias Sociales en los plazos que se muestran en la tabla.

En relación con el Plan de Evaluaciones, este se articula con el plan de elaboración de Estándares, ya que incorpora la aplicación de pruebas censales para 6° básico en todas las asignaturas o áreas en las que se elaborarán Estándares. Asimismo, incluye evaluaciones censales para Ciencias Naturales y para Historia, Geografía y Ciencias Sociales en 2° medio. Además, es importante notar que la periodicidad del Plan de Evaluaciones coincide con los tiempos requeridos para la elaboración de Estándares y aporta a dicho proceso facilitando la entrega de la evidencia necesaria para llevarlo a cabo.

Con respecto al **cuarto criterio** (permitir una clasificación u ordenación confiable de los establecimientos de educación básica y media), el Plan posibilita alcanzar el número mínimo de evaluaciones censales necesarias para cumplir con los requerimientos de la Ordenación de establecimientos requerida por ley, sin afectar la calidad métrica del índice de ordenación y clasificación de escuelas. En este sentido, el Plan no implica una pérdida de información o calidad psicométrica respecto del plan actualmente en ejecución (ver Anexo 1 para

mayor información). En efecto, desde un punto de vista psicométrico, el Plan de Evaluaciones y las metodologías utilizadas por la Agencia son capaces de entregar información válida y confiable sobre el desempeño de los y las estudiantes y los establecimientos. Al mismo tiempo, el análisis de los resultados SIMCE obtenidos por los establecimientos educativos resultan altamente consistentes entre pruebas de distintos subsectores y niveles evaluados ($r \geq 0.80$), lo que indica que la eliminación de algunos subsectores (por ejemplo, Ciencias en 4° básico) o niveles (por ejemplo, 2° básico) no tendrá efectos significativos en la puntuación que un establecimiento obtenga en el índice de ordenación de escuelas.

Respecto del **quinto criterio** (permitir una continuidad con el plan histórico, de manera de no perder la capacidad de monitoreo), el Plan cautela apropiadamente la continuidad de la serie de datos históricos de 4° básico, 8° básico y 2° medio. Históricamente, para el SIMCE se ha utilizado para fines de monitoreo un modelo de *status* (es decir, mediciones consecutivas, espaciadas por uno o dos años, en que se evalúan a distintos grupos de estudiantes con pruebas distintas) aplicando equiparación entre años evaluados, vía un diseño de anclaje. Dado que la Agencia seguirá utilizando las mismas metodologías de calibración y equiparación basadas en la Teoría de Respuesta al Ítem (IRT, por sus siglas en inglés), no habrá problemas técnicos con el monitoreo. Así, aunque el Plan supone un menor número de niveles y ocasiones evaluadas respecto del plan usado por la Agencia en el periodo 2011–2015, cabe señalar que se alinea sin problema al modelo de monitoreo utilizado hasta el año 2010.

En relación con el **sexto criterio** (incorporar la evaluación de estudiantes con necesidades educativas especiales), el Plan incluye este tipo de evaluaciones. En el contexto de la Ley N° 20.422, que establece normas sobre igualdad de oportunidades e inclusión social de personas con discapacidad y estipula que el Ministerio de Educación debe realizar las adecuaciones necesarias para que los y las estudiantes con necesidades educativas especiales puedan participar en las mediciones de la calidad de la educación, constituye una prioridad avanzar en dar la oportunidad a todos y todas de participar en estas evaluaciones. Este Plan lleva a la práctica el enfoque de inclusión en la evaluación, pues mantiene las acomodaciones para estudiantes con necesidades educativas especiales en 4° y 6° básico, y las incorpora para 8° básico.

Con respecto al **séptimo criterio** (permitir la incorporación de estudios de indicadores de Progreso), este Plan permite la estimación de algunos indicadores de progreso o valor agregado, dado que posibilita que algunas cohortes de estudiantes sean evaluadas dos veces durante su vida escolar. Por ejemplo, en algunos casos, se podrá hacer seguimiento entre 4° y 6° básico y, en otros, entre 8° básico y 2° medio. No obstante, es importante señalar que, por un lado, por la alta movilidad de estudiantes entre distintos cursos en el sistema educacional chileno y la gran proporción de escuelas pequeñas que existen en 4° básico, no es posible generar medidas de progreso o valor agregado para todos los establecimientos educacionales chilenos, lo que limita su aplicabilidad para efectos de *accountability* a nivel establecimiento. Por otro lado, aquellos establecimientos para los cuales sí es plausible estimar indicadores de progreso y valor agregado se

caracterizan por condiciones particulares (por ejemplo, alta matrícula, alta permanencia de estudiantes, predominancia de establecimientos particulares pagados o subvencionados), por lo que representan una muestra sesgada de la población de establecimientos, lo que limita la validez externa de las estimaciones para efectos de monitoreo del sistema educacional. A estas razones se suma la evidencia y el debate que existe en la comunidad académica respecto de la baja consistencia de los indicadores de valor agregado estimados con distintas metodologías y variables de control. Debido a ello, los resultados son muy sensibles a decisiones técnicas en el proceso de estimación.

Adicionalmente, los indicadores de progreso y valor agregado se focalizan en la capacidad del establecimiento para conseguir que sus estudiantes progresen en relación con sus capacidades y conocimientos de entrada. Por ello, los establecimientos de alto desempeño suelen exhibir bajos niveles de valor agregado y, por el contrario, los establecimientos con excelentes indicadores de valor agregado podrían tener, como resultado final, a la totalidad de sus estudiantes en niveles de desempeño insuficientes. En suma, por sus limitaciones prácticas y dificultades técnicas, se propone dar un desarrollo cauteloso a este tipo de indicadores, comenzando por realizar estudios con los pocos establecimientos que mantienen la mayoría de su matrícula. En esta línea, la Agencia tiene planificado realizar estudios de Progreso y/o Valor Agregado explorando diversas metodologías y sus posibles alcances y limitantes, dentro del quinquenio de vigencia del Plan de Evaluaciones, dado que contará con datos para ello. Es relevante anticipar que existen diversas aproximaciones metodológicas a los Estudios de Progreso y/o Valor Agregado y que estas metodologías enfrentan complejidades de orden práctico, especialmente en sistemas de evaluación con propósitos de responsabilización de establecimientos.

Finalmente, en lo que se refiere al **octavo y último criterio** (articulación con la participación de Chile en estudios internacionales), el Plan contempla la realización de pruebas nacionales e internacionales dentro de los mismos años, de manera de poder obtener información que se complementa, tal como se muestra en la tabla N°4.

IV. Elementos que complementan la evaluación externa de aprendizajes

Las evaluaciones censales y muestrales mencionadas anteriormente son evaluaciones estandarizadas y externas al establecimiento, consideradas muy importantes para la política pública y el proceso educativo, en cuanto entregan información sobre el logro de aprendizajes de las y los estudiantes, tanto a nivel de sistema como a nivel de establecimientos. Por su parte, las evaluaciones que realizan los y las docentes al interior de los establecimientos generan información detallada muy valiosa. Así, ambos tipos de evaluación son fuentes de información complementaria para la toma de decisiones por parte de los equipos docentes. En este sentido, desde el Ministerio de Educación y la Agencia de Calidad, se propone avanzar hacia un sistema de evaluaciones más equilibrado entre lo externo y lo interno, y potenciar un uso de los resultados para la mejora en todos sus niveles: las evaluaciones externas, las evaluaciones externas aplicadas internamente y las evaluaciones internas.

Es por ello que ambas instituciones están trabajando en dispositivos cuyo propósito es fortalecer las capacidades de las escuelas para promover y desarrollar procesos permanentes de evaluación con propósitos formativos. Con estos dispositivos se espera que los establecimientos, en primer lugar, integren la evaluación como parte integral de la enseñanza, cuyo fin último es mejorar el aprendizaje de los y las estudiantes. En segundo lugar, que visibilicen la evaluación como una oportunidad para mejorar las prácticas de las y los docentes y equipos directivos. Y en tercer lugar, que usen de manera adecuada las herramientas e información que se genere, es decir, con un propósito formativo, en términos de identificar

En esta sección se presentan dos elementos que complementan el sistema de evaluación externa de aprendizajes. Primero, se describe un tipo de evaluación externa, que se desarrollará por parte de la Agencia para uso interno y pedagógico por parte de los establecimientos educacionales. Luego, se rescata el rol de los Otros Indicadores de Calidad (OIC), que entregan información complementaria al logro de los Estándares de Aprendizaje sobre el desarrollo personal y social de los y las estudiantes, en el contexto de una formación integral.

Propuesta y fundamentación de evaluaciones externas para uso interno

Con el fin de fomentar un sistema de evaluación más equilibrado entre sus propósitos sumativos y formativos y con un mejor balance entre lo externo y lo interno, se propone incorporar evaluaciones externas para ser usadas internamente, con el propósito de que los equipos docentes cuenten con una herramienta elaborada externamente, y que utilicen internamente para retroalimentar las prácticas pedagógicas y la gestión curricular. Estas evaluaciones permitirán obtener información de resultados referidos al currículum en

distintas oportunidades a lo largo del año escolar, lo que posibilita el monitoreo de las trayectorias de aprendizaje de cada estudiante y una intervención pedagógica mucho más oportuna y pertinente.

Fundamentación de evaluaciones externas usadas internamente para Lectura en 2° básico

El desarrollo temprano de la habilidad lectora es fundamental. Son numerosos los autores que sostienen que la lectura es un proceso que incluye más que la mera decodificación y que se relaciona fuertemente con procesos cognitivos y de pensamiento, el desarrollo de conocimientos y la comprensión del mundo que alcanza la persona (Cunningham y Stanovich, 1997 y 2007; Bravo, Villalón y Orellana, 2006; Lerner, 2001; Roe y Smith, 2012; Holloway, 1999; Olson, 1994; entre otros).

En lo relativo a la educación escolar, el logro de una alfabetización a tiempo permite beneficiarse de esta sin perder las oportunidades de aprendizaje que provee la escuela, ya que la habilidad lectora es transversal a todas las asignaturas. Una adecuada alfabetización lectora permite a las y los estudiantes tener acceso independiente a la información, sin requerir de la mediación constante de docentes, padres u otros; por ello, no desarrollar la habilidad lectora tempranamente dificulta tanto el aprendizaje como la relación que se establece con el mundo (Roe y Smith, 2012).

Un análisis comparado de currículos de distintos países constata la compartida preocupación por desarrollar la lectura tempranamente. El currículum chileno no es una excepción: destaca la importancia de que los y las estudiantes desarrollen desde los primeros años habilidades comunicativas como la lectura, la escritura y la comunicación oral.

Basándose en lo anterior, resulta fundamental asegurarse de que las y los estudiantes en los primeros años escolares estén progresando sólidamente en sus habilidades lectoras. En esta lógica, en 2012, el Ministerio de Educación introdujo al sistema nacional de evaluación la aplicación de una prueba SIMCE censal de Lectura en 2° básico cuyo principal propósito es monitorear estos aprendizajes y proveer a los establecimientos con información oportuna respecto de qué tan adecuada es la comprensión lectora de las y los estudiantes, de manera de que puedan contar con un diagnóstico temprano que permita tomar medidas y orientar las estrategias para mejorar los aprendizajes de sus estudiantes (Mineduc, 2013). En esa misma línea, el Ministerio elaboró Estándares de Aprendizaje para esta área de aprendizaje y nivel, los cuales fueron promulgados a principios de 2014 y se han utilizado para reportar los resultados de las pruebas SIMCE a partir de ese mismo año.

Después de transcurridas tres evaluaciones SIMCE de Lectura en 2° básico y sus respectivas entregas de resultados a los establecimientos, el Ministerio de Educación ha podido concluir que, si bien contar con esta evaluación en formato censal ha permitido generar información valiosa para el sistema escolar, el propósito de obtener información oportuna y lo suficientemente específica como para tomar decisiones pedagógicas

pertinentes no es posible de abordar de manera adecuada en este tipo de evaluaciones. Además, de acuerdo con las consecuencias que, conforme a la ley, se le atribuyen a las pruebas con estas características (las que se relacionan con su uso en la Ordenación de los establecimientos), se ha considerado que dicho uso puede introducir distorsiones al propósito original, lo que puede desembocar en la aparición de efectos no deseados en el trabajo con estudiantes de edades tempranas, tales como estrechamiento del currículum y entrenamiento para la prueba.

Además, esta decisión es consistente con la recomendación del Equipo de Tarea –convocado en 2014 para la revisión del SIMCE– de reducir significativamente el número de mediciones censales y fortalecer las políticas de desarrollo de capacidades, para alcanzar así un mayor equilibrio entre la presión derivada de las mediciones externas y las condiciones con que cuentan los establecimientos escolares para mejorar (Equipo de Tarea, 2014). El informe de este equipo propone eliminar la prueba censal de Lectura en 2º básico y transferir esta evaluación a los establecimientos como una herramienta que permita fortalecer el conocimiento tanto de los objetivos definidos en el currículum como de la forma en que estos se pueden evidenciar (evaluar) de modo periódico con sus estudiantes, potenciando las competencias para desarrollar y gestionar información al interior del establecimiento para la mejora. De modo complementario, se espera que la transferencia de esta prueba al interior de los establecimientos educacionales fomente una lógica de responsabilización interna por los resultados de aprendizajes de los y las estudiantes, por sobre una externa centrada en consecuencias.

En lo que respecta a los aspectos que el formato censal de la prueba no permite cubrir, se encuentra el nivel de detalle de la información que el instrumento existente entrega a los establecimientos. Esta prueba, debido a sus propósitos centrales (de monitoreo de los aprendizajes en el sistema escolar, ordenación de escuelas y entrega de información externa para la gestión curricular a nivel escuela), a sus características y a sus restricciones (legales y psicométricas), entrega información a nivel de escuela en términos del porcentaje de estudiantes que alcanza cada Nivel de Aprendizaje establecido. Las características de la evaluación (tiempo de evaluación, formato de la evaluación, etc., en función de un diseño que permita abarcar una cobertura curricular amplia) dificultan la posibilidad de entregar información muy detallada. Junto con lo anterior, los resultados obtenidos en esta prueba se reportan al año siguiente de su aplicación. Ambos factores –los tipos de resultados que se pueden entregar y la oportunidad temporal de su entrega– dificultan su uso para la gestión pedagógica en aula.

Para poder contar con una herramienta realmente útil para un diagnóstico que oriente la toma de decisiones pedagógicas, se requiere que ésta provea de información oportuna, que contenga una precisión tal que permita conocer el nivel de aprendizaje alcanzado por cada estudiante y con un grado de desagregación que genere información respecto de las dimensiones al interior de cada nivel de aprendizaje que están más o menos logradas. Esto es crucial en el área de la lectura durante los primeros años de escolaridad, debido a que

las y los estudiantes ingresan a la escuela con una diversidad de experiencias previas y distintos niveles de desarrollo (Reynolds, Taylor *et al.*, 1981) y van iniciándose en la lectura a diferentes ritmos. Una herramienta que entregue información individual y precisa le facilitaría a la o el docente complementar la información evaluativa que tiene para identificar y hacerse cargo de la diversidad de aprendizajes dentro del aula al diagnosticar en qué lugar del proceso de desarrollo de las habilidades lectoras se encuentra, y así poder generar estrategias pedagógicas pertinentes a los diversos grupos instruccionales que siempre existen en un aula.

De esta manera, reconociendo la importancia del desarrollo temprano de la habilidad lectora así como la necesidad de que los equipos docentes puedan ir fortaleciendo su capacidad para generar y gestionar información diagnóstica detallada y temprana respecto del nivel de comprensión lectora que está alcanzando cada estudiante, --en el marco de los Estándares de Aprendizaje--, el Ministerio de Educación y la Agencia de Calidad de la Educación han elaborado la siguiente propuesta para transferir la prueba SIMCE de Lectura de 2º básico a los establecimientos.

La propuesta del presente Plan implica remplazar la prueba SIMCE censal de Lectura para 2º básico por una evaluación muestral periódica bianual (descrita en la sección anterior sobre evaluaciones muestrales), que permita al sistema contar con información para el monitoreo de los aprendizajes en este nivel educativo (estudiantes entre 6 y 8 años), y una nueva herramienta de evaluación que pasará a ser administrada, analizada y utilizada autónomamente por los equipos profesionales de los propios establecimientos educacionales, con orientaciones por parte de la Agencia de Calidad. De esta forma, las escuelas no solo conservarán el acceso a información temprana de los aprendizajes -complementaria a la que recogen en su práctica pedagógica-, sino que podrán obtener información con un nivel de riqueza y detalle mayor y en un lapso de tiempo menor, lo que les permitirá tomar decisiones de gestión pedagógica más atingentes, sin perder la referencia a las expectativas curriculares nacionales que brindan los Estándares de Aprendizaje. Esto posibilitará empezar a tender un necesario puente entre lo que se evalúa en la evaluación externa -y la forma en que los y las docentes comprenden y evidencian esos aprendizajes en su práctica pedagógica cotidiana al interior del aula. Así, esta estrategia puede modelar prácticas de elaboración de este tipo de preguntas y construcción de este tipo de instrumentos, y los resultados pueden servir tanto a los y las estudiantes y a sus apoderados -para reconocer sus logros y aspectos por desarrollar- como a los y las docentes, informando de manera más precisa las acciones pedagógicas a tomar. A nivel de escuela, puede servir a los equipos técnicos como información muy valiosa para complementar la que ya generan internamente y así tomar decisiones de gestión curricular cada vez más sustentadas.

i. Propósitos de esta herramienta de evaluación

En síntesis, los propósitos que guían el diseño de esta herramienta y sus resultados son los siguientes:

- Aportar información rica, precisa y oportuna a las escuelas respecto de los logros de aprendizaje de sus estudiantes en comprensión de lectura, de modo que puedan complementar la información que generan al interior de la escuela, apoyando la toma de decisiones pedagógicas y de gestión por parte de los y las docentes y los directivos.
- Fomentar la articulación de las expectativas curriculares nacionales con las que mantienen los y las docentes y en las que basan sus prácticas de planificación, enseñanza y evaluación, es decir, tender un puente entre la evaluación externa y lo que se realiza en los establecimientos. Contar con evidencia evaluativa referida a Estándares facilita esto. Esto se da en dos niveles:
 - o Nivel de gestión curricular de sostenedores (varias escuelas) o equipo directivo o técnico de la escuela: se cuenta con información referida a Estándares para tomar decisiones de gestión curricular pertinentes y oportunas (por ejemplo, generación de estrategias de apoyo pertinentes, focalización de determinadas estrategias en ciertos cursos o grupos de estudiantes de distintos cursos que presentan niveles de aprendizaje similares, definición conjunta de desarrollo profesional docente en ciertas áreas, etc.).
 - o Nivel de gestión pedagógica en aula: brinda la oportunidad para que el o la docente de aula contraste los aprendizajes que busca que sus estudiantes desarrollen con respecto a las expectativas de aprendizaje nacionales; permite modelar formas de evidenciar, en concreto, el logro de esas expectativas, en distintos niveles de aprendizaje; promueve una visión del aprendizaje como un proceso progresivo que se desarrolla con diferentes ritmos entre los y las estudiantes; impulsa una visión de la enseñanza como algo que tiene que ser diferenciado según el desarrollo de habilidades que tenga cada estudiante; promueve una preocupación por hacerse cargo de los resultados de cada uno de sus estudiantes (responsabilización interna).
- Fortalecer competencias docentes pedagógicas-evaluativas para evidenciar los niveles de aprendizaje de lectura que tienen sus propios y propias estudiantes, para distinguir distintos tipos de habilidades y puntos de desarrollo de la lectura y así visibilizar la consecuente necesidad de diferenciación de estrategias o prácticas pedagógicas, y para fomentar procesos de reflexión

pedagógica acerca de la validez de las evaluaciones y formas de evidenciar el logro de los aprendizajes en esta área.

- Favorecer la promoción de una lógica de responsabilización interna por los resultados de aprendizaje de los y las estudiantes, lo que permite actuar de modo pertinente y oportuno sobre la base de estos.

ii. Descripción de la herramienta

En breve, se propone poner a disposición de los y las docentes una página web que incluya:

- a. Un conjunto de pruebas equivalentes referidas a Estándares de Aprendizaje, en formato de lápiz y papel o en versión digital, que permitan reportar resultados individuales y agregados, según niveles de aprendizaje de los Estándares.
- b. Un tutorial que describa el funcionamiento de la página, para que los y las docentes puedan utilizar adecuadamente sus contenidos.
- c. Sugerencias de estrategias, acciones o actividades pedagógicas que pueden implementarse para mejorar la comprensión de lectura de las y los estudiantes, según su nivel de aprendizaje en el Estándar de 2º básico.

La página web y sus contenidos se desarrollarán en 2016, de modo que esté disponible para los establecimientos a partir del segundo semestre de ese año. Las pruebas elaboradas podrán aplicarse en distintos momentos del año (a elección de los equipos docentes), para monitorear el progreso individual de las y los estudiantes, y darán origen a reportes para docentes y directivos.

Durante su primer año de implementación, se pondrán a disposición de los establecimientos educacionales dos pruebas para evaluar las habilidades de comprensión lectora de alumnos de segundo año básico en dos periodos diferentes (junio y octubre). La aplicación de estas pruebas es voluntaria y podrán ser aplicadas en formato de lápiz y papel o en versión digital. Tanto la aplicación de las pruebas como el formato a utilizar, serán decisiones que tomarán directamente los establecimientos con base en sus necesidades y características particulares. Los resultados de estas evaluaciones se entregarán de manera rápida y oportuna y proporcionarán información tanto a nivel grupal como individual. Además, proporcionarán información detallada, entregando evidencia sobre los logros de aprendizaje alcanzados en relación a los estándares de aprendizaje y a las diferentes habilidades involucradas en la comprensión lectora.

Los directivos y docentes podrán tener acceso a la prueba, a sus resultados y a orientaciones pedagógicas, información que les permitirá reflexionar y definir estrategias de mejora considerando las necesidades específicas de sus estudiantes. Luego de la segunda aplicación, se entregará un reporte de progreso, lo que

permitirá al profesor identificar las diferentes trayectorias de aprendizaje, y redefinir las estrategias de enseñanza, en base a la nueva evidencia recogida.

iii. Características técnicas

Para asegurar la validez de las pruebas, todas sus preguntas serán sometidas a un exhaustivo proceso de revisiones de expertos y expertas y a un pilotaje. Estas revisiones permitirán verificar su alineamiento curricular y el cumplimiento de criterios de accesibilidad, entre otros. El pilotaje será útil para conocer la dificultad de las preguntas y su adecuación al grupo etario evaluado.

Como apoyo al análisis e interpretación de los resultados, se pondrá a disposición de los y las docentes un tutorial explicativo.

Junto con lo anterior, para fomentar un uso formativo, los resultados se encontrarán vinculados a ejemplos de estrategias y acciones que se pueden generar para avanzar en la lectura, según el nivel de desarrollo de la lectura de cada estudiante o grupo de estudiantes.

Otros Indicadores de Calidad Educativa

Con la promulgación de las leyes N° 20.370 (Ley General de Educación) y N° 20.529 (Sistema Nacional de Aseguramiento de la Calidad de la Educación Parvularia, Básica y Media y su Fiscalización), se ha avanzado hacia una concepción más amplia de la calidad de la educación, la cual incorpora indicadores respecto del desarrollo personal y social de las y los estudiantes.

En esta línea, se incluye la evaluación de Otros Indicadores de Calidad Educativa (OIC), los cuales son un conjunto de índices que entregan información acerca de los resultados obtenidos por un establecimiento educacional, en relación con el logro de los Objetivos Generales de la Educación Escolar y de los Estándares de Aprendizaje. Estos indicadores se instauraron por medio del Decreto Supremo de Educación N°381/2013, luego de ser aprobados por el Consejo Nacional de Educación.

Los Otros Indicadores de Calidad Educativa constituyen uno de los insumos de la Ordenación. La importancia de evaluar estos indicadores radica en que proporcionan a los establecimientos información relevante respecto de distintas áreas de desarrollo de los y las estudiantes, y entregan una señal sobre la relevancia de implementar acciones sistemáticas para desarrollar aspectos no académicos que se consideran fundamentales para la formación integral.

Actualmente, los Otros Indicadores de Calidad Educativa están conformados por un total de ocho indicadores:

1. Autoestima académica y motivación escolar.
2. Clima de convivencia escolar.
3. Participación y formación ciudadana.
4. Hábitos de vida saludable.
5. Asistencia escolar.
6. Retención escolar.
7. Equidad de género.
8. Titulación técnico-profesional.

Estos indicadores son evaluados mediante dos medios principales: los cuestionarios SIMCE y los registros del Ministerio de Educación y la Agencia de Calidad. Específicamente, los primeros cuatro indicadores son evaluados por medio de cuestionarios para estudiantes, docentes, padres y apoderados, los que se aplican junto con las pruebas SIMCE de carácter censal. Los indicadores de Asistencia escolar, Retención escolar y Titulación técnico-profesional son evaluados mediante los registros del Mineduc y la Agencia, mientras que el de Equidad de género se evalúa tomando en cuenta las diferencias, según género, detectadas en los resultados de las pruebas SIMCE de cada año.

Los resultados obtenidos anualmente en la evaluación censal de los OIC en aquellos cursos en los que se aplican pruebas censales SIMCE complementan la información que entregan las evaluaciones propuestas.

V. Comunicación y uso de resultados

El presente Plan de Evaluaciones genera un conjunto de información importante para orientar el mejoramiento del sistema educativo en sus distintos niveles: nacional, regional, comunal y local. En este contexto, la Agencia de Calidad de la Educación está ampliando y mejorando los dispositivos de reporte de información para retroalimentar la toma de decisiones del sistema escolar en sus distintos niveles.

Para el desarrollo de esta función, la Agencia ha elaborado un Plan de Comunicación y Difusión que tiene los siguientes objetivos:

- Responder a los propósitos que tiene el sistema nacional de evaluación de aprendizajes.
- Orientar a los sostenedores y los establecimientos en el buen uso de los resultados para la toma de decisiones de gestión.

Respecto del primer objetivo, los resultados son puestos a disposición de los **establecimientos y los sostenedores** a nivel de grado, curso y escuela con el siguiente detalle:

- Puntajes promedio de las asignaturas evaluadas, variaciones y diferencias con la evaluación anterior.
- Tendencias de puntajes en las tres últimas mediciones (si las hubiere).
- Distribución de estudiantes según Estándares de Aprendizaje, variaciones y diferencias respecto a la medición anterior.
- Comparación de resultados según GSE.
- Resultados, a nivel de escuela, de los Indicadores de Desarrollo Personal y Social¹⁹.

Los **padres y apoderados** reciben un informe de los resultados del establecimiento al que asisten sus hijos o hijas en el que se muestran, de manera gráfica, la distribución de estudiantes en los niveles de los Estándares de Aprendizaje (si los hubiere), el puntaje promedio obtenido y si es similar, más bajo o más alto que el de establecimientos de similares características, y orientaciones para su involucramiento en el proceso de aprendizaje.

Estos informes contienen los resultados de los y las estudiantes del establecimiento en Simce y en cuatro de los Indicadores de desarrollo personal y social, según grado y área evaluada. Esta información es presentada en un formato y lenguaje sencillo para facilitar su comprensión y uso. Los documentos incluyen además orientaciones para que madres, padres y apoderados acompañen activamente el proceso de aprendizaje de su hijo o pupilo. La distribución de estos informes se realiza primero de forma digital en la página web de la

¹⁹ Definidos en el contexto de la ley 20.529: Convivencia escolar, Autoestima académica y motivación escolar, Participación y formación ciudadana y Hábitos de vida saludable.

Agencia (de acceso público) y, posteriormente, de forma física a cada establecimiento en cajas dirigidas al director, para ser distribuidos entre los padres y apoderados de cada grado evaluado y además, a los del grado superior al evaluado, de modo que tengan una panorámica del nivel alcanzado. Por último, cada caja contiene una carta que explicita su contenido y la necesidad de que los informes sean entregados a sus destinatarios.

En el caso de las **autoridades nacionales, regionales y provinciales** del área de educación, se entrega:

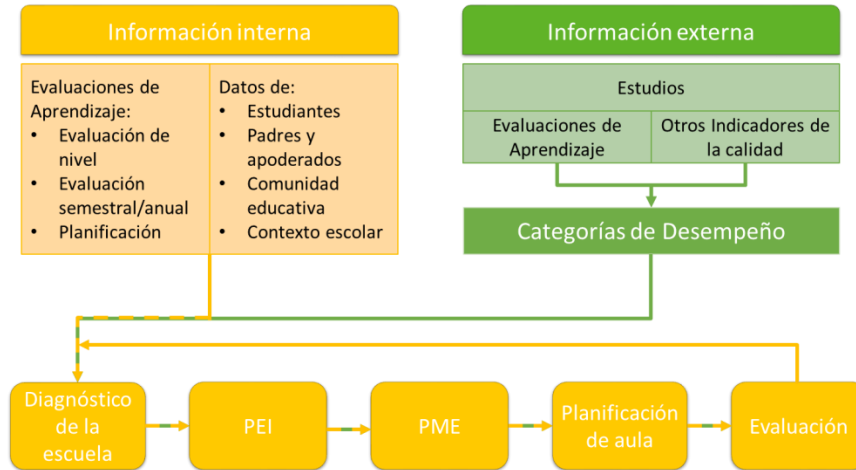
- Resultados nacionales en los Indicadores de Desarrollo Personal y Social.
- Resultados nacionales por grado y asignatura. Incluye tendencia de las últimas 10 mediciones (si las hubiere).
- Comparación de resultados según GSE a nivel nacional.
- Resultados de los factores asociados externos e internos a la escuela que evidencian relación con los resultados de aprendizaje.
- Tendencias de puntaje regional en las tres últimas mediciones (si las hubiere).
- Distribución regional en los niveles de los Estándares de Aprendizaje para los grados y asignaturas evaluadas.
- Orientaciones de política pública según el análisis de los factores asociados a los resultados.

Los **investigadores e investigadoras** pueden tener acceso a las bases de datos nacionales de cada medición 90 días después de la última entrega de resultados, previa firma de un compromiso de confidencialidad y buen uso.

El **público en general**, incluidos los medios de comunicación, tienen acceso a la presentación nacional de resultados, que incluye los mismos datos referidos anteriormente a nivel del país, así como a los de cada establecimiento por medio del sitio web de la Agencia, ingresando el nombre o RBD de este. No existe acceso a los resultados de más de un establecimiento a la vez, lo cual busca dificultar el acceso a la información para realizar *rankings* u otro tipo de comparaciones para las cuales estas evaluaciones no han sido diseñadas.

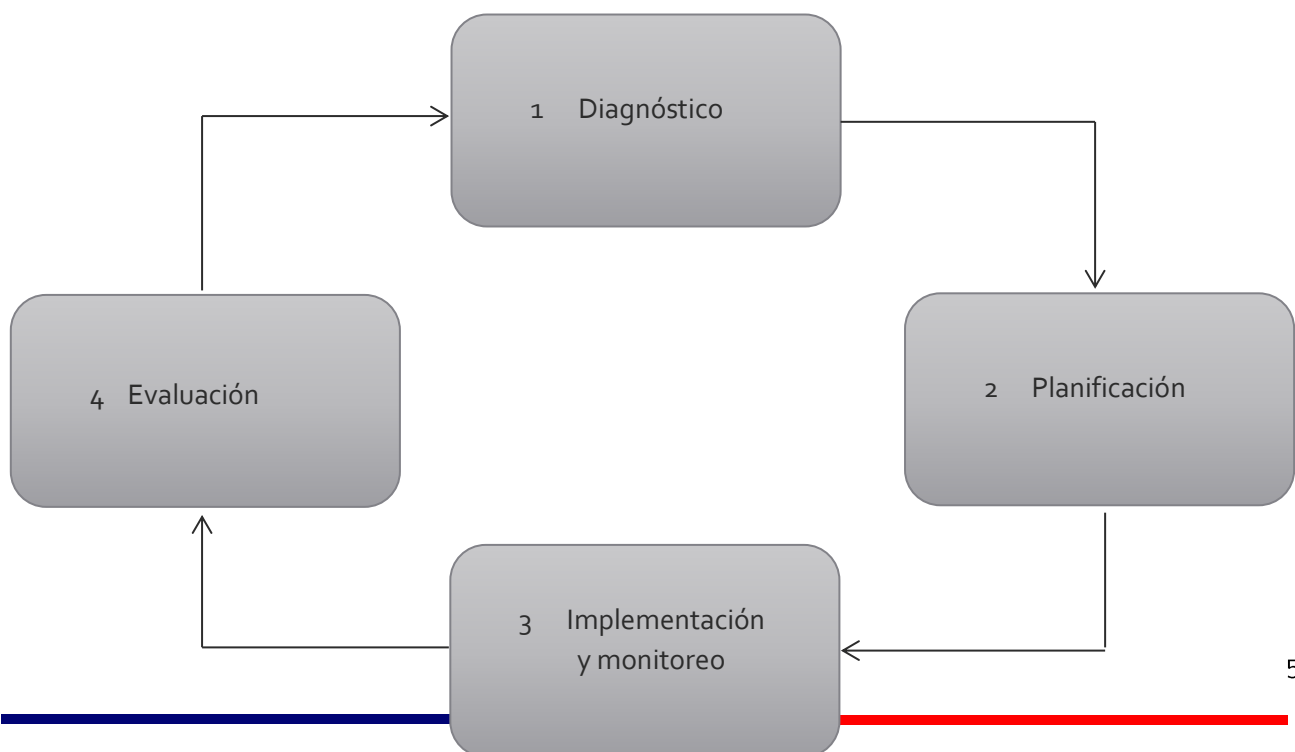
En relación con orientar a los sostenedores y los establecimientos en el buen uso de los resultados para la toma de decisiones de gestión, el foco está puesto en cómo generar capacidades en los directivos y docentes de los establecimientos para utilizar la información que les proveen las evaluaciones nacionales, integrándola con su información interna y herramientas de gestión para tomar decisiones que les aporten en el mejoramiento de sus procesos y resultados (ver figura 2).

Figura N° 2: Modelo de integración de la información externa e interna para el diagnóstico institucional.



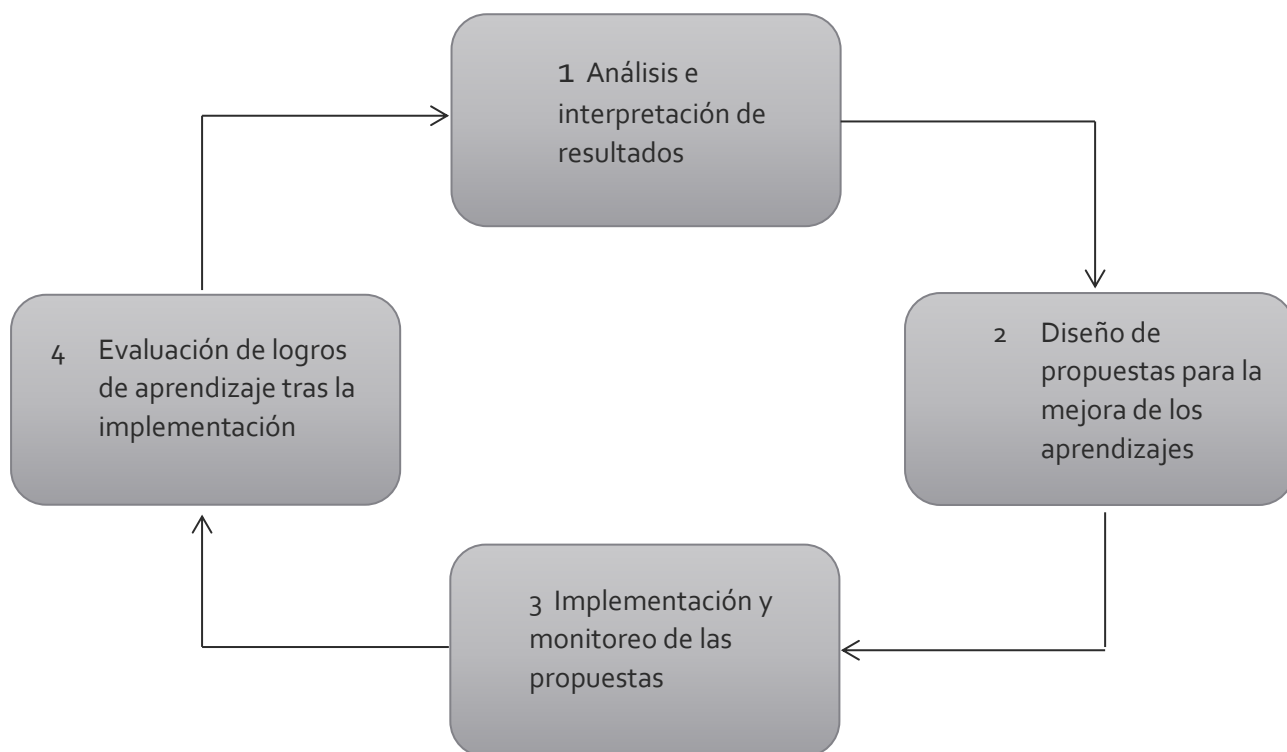
En este contexto, el diseño que se impulsa a partir de los dispositivos con los cuales la Agencia interactúa con las escuelas en el contexto de la entrega de resultados (informes, talleres, material digital en web), tiene a la base un modelo esperado de uso por parte de estas, tanto en el caso de los directivos como docentes. En el caso de los directivos, el modelo de uso impulsa que estos integren los datos de sus evaluaciones internas con los que proveen las evaluaciones externas para enriquecer su diagnóstico institucional y planificar trayectorias de mejoramiento según el ciclo descrito en la figura 3.

**Figura N° 3:
Ciclo para transformar la información que reportan las evaluaciones en acciones de mejora**



En relación con los y las docentes, el modelo de uso enfatiza adaptar las cuatro etapas del ciclo enfocadas desde la perspectiva del diseño instruccional (ver figura 4).

Figura N° 4:
Ciclo para transformar la información que reportan las evaluaciones en acciones de mejora de la gestión pedagógica, por parte de los y las docentes



No obstante los avances y desarrollos que en esta línea se están realizando, existe el desafío de seguir orientando en el uso adecuado de las evaluaciones y sus resultados por parte de las comunidades escolares y distintos actores del sistema, de tal manera que la información entregada sea útil para la toma de decisiones y los procesos de mejora educativa.

Como sistema escolar tenemos el desafío de que nuestro sistema nacional de evaluaciones provea información y orientaciones útiles, tanto para monitorear el sistema escolar como para el trabajo pedagógico y de gestión escolar en general, enfocadas en impactar en el aprendizaje de los y las estudiantes en la sala de clases.

A continuación se incluye una tabla de síntesis con los actores y productos con los cuales se reporta la información generada de las pruebas de evaluación de aprendizaje y de los indicadores de desarrollo personal y social de los estudiantes.

Tabla N° 8:
Síntesis con los actores y productos de reporte de información

Nivel	Actor	Informe/ producto
Nacional	Autoridades gubernamentales (parlamentarios, autoridades SAC y Ministros, entre otros)	- Presentación nacional de resultados - Síntesis de resultados nacionales - Informe nacional de resultados
	Ministerio de Educación	- Presentación nacional de resultados - Síntesis de resultados nacionales - Base de datos nacional - Informe nacional de resultados
	Expertos e investigadores	- Presentación nacional de resultados - Síntesis de resultados nacionales - Informe nacional de resultados - Base de datos nacional
	Medios de comunicación	- Presentación nacional de resultados - Comunicado de prensa - Síntesis de resultados nacionales
	Comunidad en general	- Información disponible en página web
Regional	Intendente	- Presentación nacional de resultados y regional - Base de datos establecimientos de la región
	Secretaría Regional Ministerial (SEREMI)	- Presentación nacional de resultados y regional - Informe de resultados agregados de la región y por provincia - Comunicado de prensa - Base de datos establecimientos de la región
	Departamento de Educación Provincial (DEPROV)	- Presentación nacional de resultados y regional - Informe de resultados agregados de la región y por provincia - Comunicado de prensa
	Medios de comunicación	- Presentación nacional de resultados - Comunicado de prensa - Síntesis de resultados nacionales
Comunal	Alcalde	- Base de datos establecimientos de la comuna - Informe agregado y por establecimiento ²⁰
	Sostenedores escolares	- Base de datos establecimientos del sostenedor
Establecimientos	Directivos	- Informe de Resultados Educativo (IRE) e Informe de los Indicadores de Desarrollo Personal y Social - Presentación power point de resultados para trabajo con la comunidad escolar
	Docentes	- Informe de Resultados Educativo (IRE) e Informe de los

²⁰ Informe a ser reportado durante la vigencia del presente Plan.

		Indicadores de Desarrollo Personal y Social
	Padres y apoderados	- Informe de Resultados Educativos para Padres y Apoderados

VI. Aseguramiento de la calidad técnica de las pruebas

La validación de los instrumentos es un proceso de suma importancia para los sistemas de evaluación. Variados organismos internacionales, tales como el consorcio de AERA, APA y NCME y la International Association for the Evaluation of Educational Achievement (IEA), han diseñado y publicado estándares de calidad internacionales con el objetivo de entregar criterios para perfeccionar y asegurar la calidad técnica de los instrumentos.

Por un largo tiempo, la validación de los instrumentos ha tenido una mirada mayoritariamente psicométrica, en que los criterios de calidad han sido principalmente la validez y confiabilidad de los instrumentos. Sin embargo, Messick (1995) ha relevado la importancia de considerar los resultados de las mediciones y sus posibles usos e interpretaciones dentro de los criterios de calidad para juzgar la calidad técnica de los instrumentos. Dado lo anterior, los estándares de la AERA, APA y NCME (2014) han enfatizado tres aspectos fundamentales para evaluar y asegurar la calidad técnica de las pruebas: validez, confiabilidad y equidad de los instrumentos.

La Agencia de Calidad de la Educación, en su rol de asegurar la calidad técnica de los instrumentos utilizados para la evaluación de los aprendizajes de las y los estudiantes, ha ido implementado variados procedimientos para asegurar y perfeccionar continuamente la calidad técnica de las pruebas, considerando los criterios de calidad internacionales mencionados. La transparencia de sus procesos ha sido un factor clave de validación, pues pone a disposición de la población la información sobre cómo y con qué objetivo han sido diseñados y construidos los instrumentos, cómo se ha determinado la muestra o población para cada test, cómo se realizan los procesos de aplicación, procesamiento y análisis de estos datos, y la entrega de sus resultados. Estos procedimientos han sido presentados al público en los informes técnicos que la Agencia ha publicado desde el año 2012 y que pretende seguir publicando, facilitando así, a quien lo requiera, toda la información referida al desarrollo de las evaluaciones que realiza.

La validez, considerada como el grado en que la evidencia y la teoría apoyan las interpretaciones de los resultados para los propósitos con que las pruebas fueron diseñadas, es un aspecto fundamental en el desarrollo y análisis de los instrumentos. La Agencia desarrolla estudios internos para recoger la evidencia de distintas fuentes (en cuanto al contenido medido, los procesos de respuesta de las y los estudiantes, las relaciones de constructos, etc.), de manera de asegurar la calidad de los instrumentos. Además, actualmente existe un protocolo de trabajo conjunto entre la Agencia y la Unidad de Currículum y Evaluación del Mineduc, orientado a velar por el alineamiento curricular de los instrumentos de evaluación. Este protocolo se

encuentra en proceso de actualización para introducir ajustes que, después de algunos ciclos de aplicación, se consideran convenientes de incorporar.

En materia de confiabilidad –que se refiere a la capacidad del sistema para minimizar el error de medición–, la Agencia ha puesto especial énfasis en someter las evaluaciones a exhaustivas revisiones antes de ser aplicadas de manera censal y definitiva. Por ello se realizan pruebas experimentales que tienen como objetivo probar las preguntas elaboradas en poblaciones similares para asegurar que el error de medición sea lo menor posible y así verificar y elegir los ítems con los mejores datos psicométricos. Ya que los resultados de las pruebas referidas a los Estándares de Aprendizaje tienen consecuencias para los establecimientos, se verifica que los índices de confiabilidad sean lo más alto posible cumpliendo con los estándares internacionales al respecto.

La accesibilidad e imparcialidad de las evaluaciones realizadas por la Agencia son dos condiciones requeridas para minimizar las posibilidades de que los resultados de algún grupo o estudiante se expliquen por algún motivo distinto al nivel de desarrollo de la habilidad evaluada (AERA, APA y NCME, 2014). Esto apunta a que los ítems en todas las evaluaciones puedan ser respondidos por toda la población y a que no exista varianza irrelevante de constructo, lo que impediría a ciertos grupos responder por motivos ajenos a la habilidad (por ejemplo, por religión, grupo racial, territorio, etc.). Para abordar este tema, la Agencia lleva a cabo la revisión de los ítems por medio de un comité de imparcialidad, de manera de asegurar que estos no presenten ningún tipo de sesgo. Asimismo, luego de que cada ítem ha sido experimentado, se verifica que los datos métricos no muestren indicios de sesgo.

Además de los estudios internos para asegurar la calidad técnica de los instrumentos, se considera importante tener una mirada externa de instituciones de prestigio internacional en el área. Por esto, la Agencia promoverá la realización de un plan de validación de evaluaciones²¹, con el objetivo de verificar que los estándares de calidad de los instrumentos se estén cumpliendo.

VII. Monitoreo de la evaluación

Con el propósito de orientar a la política pública a partir de la información generada mediante la evaluación, se desarrollará y promoverá investigación asociada a la comprensión y uso de los datos generados por el sistema de evaluación, tanto en su evaluación nacional como internacional. Para ello, se recogerá información de directores, jefes de UTP, docentes, apoderados y estudiantes, lo que permitirá identificar posibles dinámicas recurrentes que derivan en usos no deseados de la información.

Como parte del plan de monitoreo, se coordinará la realización de un estudio longitudinal de validez y comparabilidad de las mediciones censales y muestrales, para verificar que los cambios propuestos al Plan de

²¹ En la medida que el presupuesto lo permita.

Evaluaciones cumplen adecuadamente con sus propósitos. Además, en el caso de 2° básico, la Agencia realizará un estudio específico que permita determinar el grado de apropiación y los usos de las evaluaciones externas para uso interno, por parte de los equipos docentes y directivos escolares. Con el propósito de disponer de los recursos necesarios para la realización de estos estudios, tanto el Mineduc como la Agencia harán las gestiones a su alcance para obtenerlos mediante la Ley de Presupuesto.

En cuanto a la evaluación de los indicadores de desarrollo personal y social (u otros indicadores de calidad), se realizarán estudios con levantamiento de información para estudiantes, docentes, apoderados y directores para refinar el desarrollo de las preguntas de los instrumentos. En paralelo, mediante un trabajo colaborativo -que ya se encuentra en desarrollo- entre la Agencia y el Ministerio, se elaborará una propuesta que permitirá enriquecer la entrega de información relacionada con dichos indicadores.

Considerando el desarrollo de una evaluación que analiza el desarrollo de competencias genéricas durante la Enseñanza Media Técnico-Profesional y para potenciar el uso de la información generada por esta, se recogerá evidencia de distintos actores para analizar el concepto de calidad en las definiciones de las competencias genéricas del currículum de educación técnico-profesional.

El Mineduc, por su parte, contará con líneas de fomento de estudios de percepciones, interpretaciones y usos del sistema de evaluación externo por medio de FONIDE y otras herramientas.

VIII. Análisis de la factibilidad técnica y económica del Plan

La implementación del presente Plan de Evaluaciones constituye un gran desafío para la Agencia, más allá de los recursos presupuestarios requeridos, puesto que supone la consolidación de los equipos de trabajo y la puesta en marcha de una agenda de desarrollo de capacidades en evaluación y medición educativa. En este sentido, se harán los esfuerzos para asegurar que los y las profesionales responsables de la elaboración y análisis de las evaluaciones cuenten con las condiciones laborales adecuadas para cumplir su labor y, en forma paralela, adoptar medidas para fomentar el desarrollo de capital humano avanzado en evaluación.

En cuanto a capacidades operativas, la Agencia cuenta con una cartera de proveedores externos y contrapartes internas debidamente calificados, para una adecuada implementación del Plan. En este aspecto, cabe destacar que la expansión del sistema de evaluación durante el periodo 2011-2015 fue un factor determinante para el desarrollo de dichas capacidades internas y externas.

Por último, es importante señalar que la implementación de este Plan cuenta con el debido respaldo presupuestario, bajo el entendido de que la Agencia elaborará la solicitud presupuestaria correspondiente. No obstante, cabe precisar que la asignación real de presupuesto se sanciona anualmente, en el contexto de la aprobación de la Ley de Presupuestos.

IX. Referencias

AERA, APA & NCME (2014). *Standards for Educational and Psychological Testing*.

Agencia de Calidad de la Educación de Chile (2015). *Propuesta Sistema Nacional de Evaluación de la Calidad de la Educación Escolar*. Planes de Evaluación de Aprendizajes 2016–2020. Santiago: Autor.

Barber, M. & Mourshed, M. (2007). *How the World's Best-Performing School Systems Come Out On Top*. S.S.O.M.C. McKinsey & Company (Ed.) Recuperado de:
http://www.mckinsey.com/client-service/social-sector/resources/pdf/Worlds_School_Systems_Final.pdf

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>.

Bransford, J.D., Brown, A.L. & Cocking, R.R. (2000) *How people learn: brain mind, experience and school*. Washington D.C.: National Academy Press, citado en Margaret Forster, op.cit.

Bravo, V.L., Villalón, M. & Orellana, E. (2006). *Predictibilidad del rendimiento en la lectura: una investigación de seguimiento entre primer y tercer año*. Revista latinoamericana de psicología, 38(1), 9–20.

Consejo Asesor Presidencial para la Calidad de la Educación (2006). *Informe Final*. 14–17 y 24–25. Gobierno de Chile.

Cunningham, A. E. & Stanovich, K. E. (2007). *Los efectos de la lectura en la mente*. Santiago de Chile: Centro de Estudios Públicos (CEP).

Cunningham, A. E. & Stanovich, K. E. (1997). *Early Reading Acquisition and Its Relation to Reading Experience and Ability 10 Years Later*. Developmental Psychology, 33(6), 934–945.

Darling Hammond, L. & Bransford, J. (Eds.). (2005). *Preparing Teachers for a changing world, What teachers should learn and be able to do*. (First ed.): Jossey-Bass.

Decreto N°614/2013. *Establece bases curriculares de 7° año básico a 2° año medio en asignaturas que indica*. Publicado el 4 de marzo de 2014. Santiago, Chile. Disponible en: www.leychile.cl

Dinham, S., Ingvarson, L. & Kleinhenz, E. (2008). *Teaching talent: the best teachers for Australia's classrooms*. B. C. O. Australia (Ed.) Recuperado de:

<http://www.bca.com.au/DisplayFile.aspx?FileID=436>

Equipo de Tarea para la Revisión del SIMCE (2014). *Hacia un sistema completo y equilibrado de evaluación de los aprendizajes en Chile: Informe Equipo de Tarea para la Revisión del SIMCE*. No publicado.

Forster, M. (2009). *Informative Assessment: Understanding and guiding learning*. Paper presented at the Research Conference "Assessment and Student Learning: Collecting, interpreting and using data to inform teaching", Perth.

Griffin, P., Care, E., Francis, M., Hutchinson, D., Arratia, A. & McCabe, C. (2013). *Assessment and Learning Partnerships: The influences of teaching practices on students achievement: Assessment Research Centre*. Australia: University of Melbourne.

Hattie, J. (2003). *Teachers make a difference: what is the research evidence*. *Interpretations*, 36(2), 27–38.

Hattie, J. (2005). *What is the nature of evidence that makes a difference to learning? Using data to support learning*.

Hattie, J. & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), pp. 81–112.

Hill, R. & C. DePascale (2003): *Adequate Yearly Progress under NCLB: Reliability Considerations*. Paper presented at the 2003 Annual Meeting of the National Council on Measurement in Education.

Holloway, J. H. (1999). *PISA: Evaluación de las competencias lectoras para el siglo XXI*. Santiago de Chile: Ministerio de Educación.

Horndike, R. & Thorndike–Christ, T. (2010). *Measurement and evaluation in psychology and education* (Eighth Edition ed.). Boston: Pearson Merrill.

International Reading Association and the National Association for the Education of Young Children. (1998). *Learning to read and write: Developmentally appropriate practices for young children*. Newark, Del.: International Reading Association.

Kim, K. H. (2011). The Creativity Crisis: The Decrease in Creative Thinking Scores on the Torrance Tests of Creative Thinking. *Creativity Research Journal*, 23(4), pp.285–295.

Lerner, D. (2001). *Leer y escribir en la escuela, lo real, lo posible y lo necesario*. México D.F.: SEP/Fondo de Cultura Económica.

Ley N°20.370. *Establece la Ley General de Educación*. Publicación 12 de septiembre de 2009. Santiago, Chile. Disponible en: www.leychile.cl

Ley N°20.529. *Sistema Nacional de Aseguramiento de la Calidad de la Educación Parvularia, Básica y Media y su Fiscalización*. Publicación 27 de agosto de 2011. Santiago, Chile. Disponible en: www.leychile.cl

Ley N°19.712. *Ley del Deporte*. Publicación 9 de febrero de 2001. Santiago, Chile. Disponible en: www.leychile.cl

Looney, J. W. (2011). *Alignment in Complex Education Systems: Achieving Balance and Coherence*. OECD Education Working Papers, No. 64, OECD Publishing. Paris, France.

Louden, W., Rohl, M., Barrat Pugh, C., Brown, C., Cairney, T., Elderfield, J. (2005). *In Teachers' Hands: Effective Literacy Teaching Practices in the Early Years of Schooling*, 197. Department of Education, S. a. T., Australian Government.

Manzi, J., González, R. & Sun, Y. (2011). *La Evaluación Docente en Chile*. Santiago de Chile: Centro de Medición MIDE UC, Pontificia Universidad Católica de Chile.

Messick, S. (1995). *Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning*. *American Psychologist*, 50(9), 741-749. Recuperado en: <http://dx.doi.org/10.1037/0003-066X.50.9.741>

Ministerio de Educación de Chile. (2002). *Objetivos Fundamentales y Contenidos Mínimos Obligatorios de la Educación Básica y Media, Actualización 2009*. Santiago: Autor.

Ministerio de Educación de Chile (2008). *Por qué debemos evaluar el trabajo de los niños*. Organización del Bachillerato Internacional, enero 2002. En *Evaluación para el Aprendizaje: Educación Básica Primer Ciclo*.

Ministerio de Educación de Chile (2012). *Bases Curriculares 2012: Educación Básica*. Santiago: Autor.

Ministerio de Educación de Chile. (2013). *Bases Curriculares 2013: Educación Básica*. Santiago: Autor.

Ministerio de Educación de Chile. (2013). *Fundamentos Estándares de Aprendizaje Lenguaje y Comunicación: Lectura 2° Básico*. Santiago de Chile: Autor.

Ministerio de Educación de Chile (2014). *Otros Indicadores de Calidad Educativa*. Santiago: Autor.

Mullis, I. V. S., Kennedy, A. M., Martin, M. O. et al. (2006). *Estudio Internacional de Progreso en Comprensión Lectora 2006: Marcos Teóricos y Especificaciones de Evaluación*. Amsterdam/Madrid: Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA) e INECSE, Ministerio de Educación y Ciencia.

Näsström G. & Henriksson, W. (2008). *La alineación de estándares y evaluación. Un estudio teórico y empírico de métodos para la alineación*. Revista Electrónica de Investigación Psicoeducativa. ISSN. 1696-2095. Nº 16, Vol 6(3). 2008, 667-690.

OECD. (2012). *Preparing Teachers and Developing School Leaders for the 21st Century: Lessons from around the world*. In A. Schleicher (Ed.). Paris.

OECD. (2013). *Synergies for Better Learning: An International Perspective on Evaluation and Assessment*, OECD Reviews of Evaluation and Assessment in Education, OECD Publishing, Paris.

Olson, D. R. (1994). *PISA: Evaluación de las competencias lectoras para el siglo XXI*. Santiago de Chile: Ministerio de Educación.

Reynolds, R. E., Taylor, M. A., Steffensen, M. S., Shirey, L. L. & Anderson, R. C. (1981). *Cultural schemata and reading comprehension*. Technical report N°201. Illinois: University of Illinois, Center for the Study of Reading.

Roe, B.D. & Smith, S.H. (2012). *Teaching reading in today's elementary schools*. Belmont: Wadsworth.

Rowe, K. (2007). *ASSESS and ASSIST: Capacity building for ALL teachers of students -with and without learning difficulties*, Australia 2007. En Evaluación para el aprendizaje: Educación Básica primer Ciclo, Ministerio de Educación de Chile, 2008.

Stanovich, K. E. (1986). *Matthew Effects in Reading: Some Consequences of Individual Differences in the Acquisition of Literacy*. Reading Research Quarterly, 21(4), 360-407.

Wolfe, E. & Smith, E. (2007a). *Instrument Development Tools and Activities for Measure Validation Using Rasch Models: Part I Instrument Development Tools*. In E. Smith & R. Smith (Eds.), Rasch measurement: Advanced and Specialized Applications, 202-242. Maple Grove, Minnesota, USA: JAM Press.

Wolfe, E. & Smith, E. (2007b). *Instrument Development Tools and Activities for Measure Validation Using Rasch Models: Part II Validation Activities*. In E. Smith & R. Smith (Eds.), *Rasch measurement: Advanced and Specialized Applications*, 243-290. Maple Grove, Minnesota, USA: JAM Press.

X. Anexos

Anexo N° 1

Implicancias del plan para la ordenación de escuelas

Cabe señalar que el nuevo plan de evaluaciones nacionales censales implica ciertas modificaciones al calendario de ordenación o clasificación en categorías de desempeño que se realiza en el marco de la Ley 20.529. En la siguiente tabla se presentan las evaluaciones que, de acuerdo a lo estipulado por la ley, se utilizaron para la ordenación 2014, y las que, en el marco del plan de evaluación propuesto, se debieran utilizar para cada clasificación desde 2015. Este ejercicio se construyó haciendo las siguientes interpretaciones de la ley: a) cuando se elimina una evaluación del plan de evaluaciones, ésta pasa a ser automáticamente eliminada de las Categorías de Desempeño; b) cuando una evaluación pasa de ser anual a bianual, ésta se mantiene en su calidad de anual (y por lo tanto considera las tres últimas mediciones) hasta que efectivamente se genere una brecha de dos años en las últimas dos evaluaciones.²² Por otro lado, la tabla asume que en cada uno de los años en que se genera la clasificación, se cuenta con Estándares de Aprendizaje para las evaluaciones consideradas, y que éstos pueden ser aplicados retroactivamente. Esto aplica para el caso de 6° básico, donde se asume que los estándares de aprendizaje estarán disponibles en 2017.

²² Estos criterios o interpretaciones no son vinculantes y podrán ser modificadas en el futuro.

Tabla 2-2. Año de prueba considerada para cálculo de categorías de desempeño para educación Básica, de acuerdo a plan de evaluaciones propuesto (últimos dos dígitos de cada año considerado)

Grado/Prueba		Plan Vigente						Plan Propuesto														
		2014/a			2015/b			2016			2017			2018			2019			2020		
2o	Lectura	-	-	-	1	13	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
					2		4															
4o	Matemática	1	1	1	1	13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		1	2	3	2		4	3	4	5	4	5	6	5	6	7	6	7	8	7	8	9
	Lectura	1	1	1	1	13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		1	2	3	2		4	3	4	5	4	5	6	5	6	7	6	7	8	7	8	9
	Sociales	1		1	1		1	-		-	-		-	-		-	-		-	-		-
		0		2	2		4															
	Naturales	1		1	1		1	-		-	-		-	-		-	-		-	-		-
		1		3	1		3															
6o	Matemática	-	-	-	-	-	-	-	-	-	1	1	1	1	1	1		1	1		1	1
											4	5	6	4	5	6	6	8	6		8	8
	Lectura	-	-	-	-	-	-	-	-	-	1	1	1	1	1	1		1	1		1	1
											4	5	6	4	5	6	6	8	6		8	8
	Escritura	-	-	-	-	-	-	-	-	-	1	1	1	1	1	1		1	1		1	1
											4	5	6	4	5	6	6	8	6		8	8
	Sociales	-		-	-		-	-		-	-		-	-		-	-		-	-		-
	Naturales	-		-	-		-	-		-	1		1	1		1		1	1		1	1
											4		6	4		6	6	8	6		8	8
8o	Matemática	1		1	1	13	1	1	1	1	1	1	1	1	1	1		1	1		1	1
		1		3	1		4	3	4	5	3	4	5	5	7	5		7	7		7	9
	Lectura	1		1	1	13	1	1	1	1	1	1	1	1	1	1		1	1		1	1
		1		3	1		4	3	4	5	3	4	5	5	7	5		7	7		7	9
	Sociales	0		1	1		1	1		1			1	1		1		1	1		1	1
		9		1	1		4	1		4	1		4	4		7	4	7	7		7	9
	Naturales	1		1	1		1	-		-	-		-	-		-	-		-	-		-
		1		3	1		3															

Según el plan y los supuestos realizados, las categorías de desempeño se realizarán, hasta 2016, en base a los tres grados en los cuales se cuenta con estándares de aprendizaje (4° y 8° Básico). Luego, a partir de 2017, se incorporará la información de 6° básico, lo cual agregará información para un tercer grado en Educación Básica. Esto es relevante, pues, uno de los aspectos más importantes para asegurar la validez de las categorías de desempeño es la cantidad de observaciones independientes realizadas en cada establecimiento, lo que a su vez depende de la cantidad de grados considerados (Hill & DePascale, 2003).²³ En este sentido es importante destacar que mientras el traspaso de 8° básico a la Educación Media no se haga efectivo, la robustez de la ordenación de Educación Media (que hasta entonces se realizará exclusivamente en base a información de II Medio) será considerablemente menor que la de la Educación Básica.

²³ Dado que los resultados de las distintas pruebas están altamente correlacionados dentro de cada grado, las pruebas no constituyen observaciones independientes del establecimiento.

Tabla 2-3. Número de Pruebas consideradas para la ordenación en educación básica, según plan vigente y plan propuesto.

	Plan Vigente							Plan Propuesto				
	2014	2015	2016	2017	2018	2019	2020	2016	2017	2018	2019	2020
2o	0	3	3	3	3	3	3	0	0	0	0	0
4o	10	10	10	10	10	10	10	6	6	6	6	6
6o	0	0	0	11	13	13	13	0	11	11	8	8
8o	8	10	10	10	10	10	10	8	8	8	8	8
Total	18	23	23	34	36	36	36	14	25	25	22	22

A modo de resumen, en la tabla 2-3 se presenta la cantidad de grados y pruebas considerados para las categorías de desempeño de Educación Básica. A partir de esta tabla es posible observar desde 2017, tanto para el plan vigente como para el plan propuesto, una importante alza en el número de pruebas consideradas para la clasificación de desempeño, debido a la disponibilidad de estándares de aprendizaje en 6º básico, que posibilitan la incorporación de este grado a la clasificación de establecimientos.

No obstante, la disminución de pruebas que contempla el nuevo plan de evaluación implica una disminución de la cantidad de información respecto a lo que habría sido incorporado en caso que el plan vigente se mantenga. Esta disminución de información se genera porque el plan propuesto implica que las pruebas aplicadas a 6º y 8º básico pasarían a ser bi-anales (en comparación con la anualidad actual de dichas pruebas) y no con la eliminación de grados. La bi-anualidad de un mayor número de pruebas implica que, cada dos años, la información se repetiría respecto a la del año anterior para un número relativo mayor de pruebas. En efecto (y considerando las ponderaciones a cada prueba/año), a partir de la tabla 2-4, que muestra que la información referente a pruebas que se repite entre uno y otro, es posible observar que la información repetida llega a ser de un 100% en 6º básico (años par) y en 8º básico (años impar).

Tabla 2-4. Porcentaje de pruebas repetidas respecto al año anterior en Categorías de Desempeño de Básica, 2014- 2020 para plan vigente (Vig.) y plan propuesto (Prop.)

Grado	2 Básico		4 Básico		6 Básico		8 Básico		Promedio	
	Vig.	Prop.	Vig.	Prop.	Vig.	Prop.	Vig.	Prop.	Vig.	Prop.
2014	0%		42%				42%		42%	
2015	50%		58%				58%		47%	
2016	50%		58%	50%			58%	67%	58%	58%
2017	50%		58%	50%	0%	0%	58%	100%	39%	50%
2018	50%		58%	50%	50%	100%	58%	33%	56%	61%
2019	50%		25%	50%	57%	33%	58%	100%	47%	61%
2020	50%		17%	33%	57%	100%	58%	33%	44%	56%
Promedio	42%		45%	47%	41%	58%	56%	67%	47%	57%

Nota: porcentajes consideran pesos

Es posible pensar que mientras mayor sea la información repetida respecto al año anterior, menor es la potencial movilidad entre categorías de desempeño en el año corriente. De hecho, nuestras simulaciones demuestran que la relación entre el porcentaje de pruebas repetidas y el porcentaje de pruebas y el porcentaje de establecimientos que se mantiene en la misma categoría es considerablemente alta (0,97). No obstante y como se demuestra a continuación, este problema afecta a un número relativamente pequeño de establecimientos.

Como muestra la tabla 2-4, si bien es esperable encontrar un porcentaje de pruebas repetidas respecto al año anterior, este porcentaje es levemente mayor en el escenario del plan propuesto que en el del plan vigente. En promedio, el plan vigente presenta un porcentaje de pruebas repetidas que llega al 58% en 2016, mientras que en el plan propuesto este porcentaje llega a 61% en 2018 y 2019. De hecho, un año crítico es 2019, donde el porcentaje de pruebas repetidas aumenta de 47% en el plan vigente a 61% en el plan propuesto. En base a las simulaciones realizadas y descritas en el Anexo 3, con un 47% de pruebas repetidas, alrededor de 75% de establecimientos que se mantienen de categoría en el año siguiente, porcentaje que aumenta a 80% con un 61% de pruebas repetidas. Esta diferencia se traduce en que, en el peor de los años, alrededor de 280 establecimientos más en el plan propuesto que en el vigente (considerando un universo de 5.500 establecimientos con datos suficientes para ser clasificados) se mantendrían en la misma categoría de desempeño. Para evaluar qué tipo de establecimientos se ven afectados por este artefacto, es ilustrativo considerar a los establecimientos que se clasifican con un solo grado evaluado bi-anualmente como por ejemplo los que tienen solo 8º en lo que respecta a educación básica. Hasta que 8º sea parte de la educación básica, estos establecimientos repetirán su categoría de desempeño cada 2 años pues efectivamente serán evaluados (tanto en pruebas SIMCE como en OIC) cada dos años. Estos colegios representan a un 1,7% del total de establecimientos con datos suficientes para ser clasificados en 2013 y en 2014 (n=86).

Quizás más relevante es el caso de los establecimientos que vería limitada su transición desde la categoría de desempeño insuficiente. Las simulaciones realizadas y descritas en el Anexo 3, indican que para 2019, un 69% de los establecimientos insuficientes se mantendría en la misma categoría en el escenario del plan vigente. Este porcentaje aumenta a 72% en el escenario del plan propuesto, lo que implica que alrededor de 17 establecimientos más en el plan propuesto que en el vigente (considerando un universo de 580 establecimientos en categoría insuficiente) se mantendrían en la misma categoría de desempeño.

Por otro lado, es importante mencionar que el nuevo plan de evaluación no tiene mayor efecto en la potencial cobertura de la ordenación. En efecto, un establecimiento requiere un mínimo de datos para poder ser clasificado. La tabla 2-5 a continuación provee los porcentajes de establecimientos y estudiantes ordenados en los escenarios de plan de evaluación vigente y propuesto, considerando que 8 pasa a media en 2018 y que no lo hace. Observamos que en ambos escenarios existe un alza importante del número de establecimientos ordenados en 2017 (fruto de la incorporación de 6º básico al ejercicio de clasificación), y que la disminución en cobertura es entre 1 y 3 puntos de porciento en lo que respecta a establecimientos. Sin embargo, la proporción de alumnos sujetos al ejercicio de clasificación es alto y se mantiene virtualmente inalterado.

Tabla 2-5. Proyección de número de establecimientos y estudiantes clasificados en Categorías de Desempeño de Básica, 2016- 2020 para plan vigente y plan propuesto (%)

	Establecimientos		Estudiantes	
	Vigente	Propuesta	Vigente	Propuesta
2016	68,0%	69,0%	98,0%	98,0%
2017	74,0%	73,0%	98,0%	98,0%
2018	76,0%	75,0%	98,0%	98,0%
2019	77,0%	76,0%	98,0%	98,0%
2020	79,0%	76,0%	99,0%	98,0%

Los análisis de confiabilidad realizados en base a la información y metodología disponible, indican que la disminución de la consistencia dada la disminución de pruebas asociada al nuevo plan de evaluación, no sería de alta magnitud. En efecto, nuestros datos tienden a corroborar lo planteado por los expertos en medición educacional respecto a que lo más importante para mantener la confiabilidad de la ordenación es aumentar la cantidad de observaciones independientes de un establecimiento (o grados), y no necesariamente aumentar las observaciones dentro de cada grado (Hill y DePascale, op cit). La tabla 2-6 muestra un ejercicio preliminar donde se calcula el indicador según los métodos de Split Half y simulaciones de Montecarlo en tres escenarios: uno donde se incorporan 4 asignaturas en 4º básico y 3 asignaturas en 8º básico (similar al plan actual), uno donde se elimina completamente 8º básico y un tercero donde se mantienen ambos grados pero se eliminan asignaturas (similar al plan propuesto). El ejercicio realizado, muestra que, en comparación con el escenario inicial y para ambos indicadores, la consistencia de la clasificación disminuye en 3 puntos al

eliminar un grado. Sin embargo, al eliminar asignaturas la consistencia disminuye menos: 2 puntos en el caso del método Split-half y 1 punto en el caso del método de Montecarlo.

Tabla 2-6. Consistencia para clasificación 2014 en tres escenarios: plan vigente, eliminación de grado, eliminación de asignaturas manteniendo grado (Método de Split- Half y Repeticiones de Monte Carlo)²⁴

	Alternativas de Plan de Evaluación		
	4° Básico 4 asignaturas 8 Básico 3 asignaturas	4° Básico 4 asignaturas	4° Básico 2 asignaturas 8 Básico 2 asignaturas
Método Split-Half	75%	72%	73%
Método Repliaciones Tipo Monte Carlo	85%	82%	84%

En resumen, si bien el plan propuesto disminuye la cantidad de información disponible para la ordenación en educación básica, esto no genera mayores problemas de confiabilidad ni de cobertura en la clasificación. Por otro lado, la mayor cantidad de pruebas bi- anuales que se traduce en un mayor porcentaje de pruebas repetidas entre años contiguos, genera una disminución en las probabilidades de transitar de una categoría a otra. Sin embargo, nuestras simulaciones indican que este problema afecta a un número reducido de establecimientos y, en el peor de los casos, a solo 17 establecimientos de la categoría insuficiente.

²⁴ Los resultados de Split-Half fueron generados aplicando 500 combinaciones distintas de formas en que los establecimientos pueden haber sido partidos en dos. Los ejercicios tipo Monte Carlo fueron generados realizando 500 muestreos aleatorios con reposición. Todos los resultados fueron generados en la Agencia y además en el National Laboratory for High Performance Computing del Centro de Modelamiento (CMM) Matemático de la Facultad de Ciencias Físicas y Matemáticas de la Universidad de Chile.