

## PROGRAMA DE EDUCACIÓN

# Hacia una política nacional de evaluación educativa a gran escala al servicio de la mejora de los aprendizajes

## Resumen

Esteban Torre

Juan Cruz Perusia

Juan Xanthopoulos

El mundo enfrenta una “crisis del aprendizaje” (UNESCO, 2013; Banco Mundial, 2018; UNICEF, 2021). Una proporción elevada de niños, niñas y adolescentes que están dentro del sistema educativo no alcanza resultados de aprendizajes satisfactorios.

En Argentina, tanto las evaluaciones internacionales como las nacionales evidencian un déficit marcado de habilidades fundamentales en Lengua y Matemática. En secundaria, el último dato disponible surge de las pruebas Aprender 2022, que muestran que 4 de cada 10 y 8 de cada 10 adolescentes escolarizados en el último año del nivel secundario no alcanzan un nivel satisfactorio en Lengua y Matemática, respectivamente.

En este contexto, existe un acuerdo amplio entre las autoridades educativas nacionales y organismos especializados en educación sobre la importancia que tienen las evaluaciones de aprendizajes a gran escala como un insumo esencial para la mejora de la calidad educativa. Este tipo de evaluaciones se implementan a nivel del sistema educativo sobre la totalidad o una muestra de estudiantes de un grado específico, y son estandarizadas en cuanto a dominios (contenidos y habilidades) y grados evaluados, período de evaluación, metodología de corrección y resultados.

Ahora bien, para que las evaluaciones de aprendizajes a gran escala se transformen en una verdadera herramienta para la mejora, es necesario tomar un conjunto de decisiones estrechamente vinculadas entre sí. La primera de ellas es definir su propósito general: no contar con un “para qué” claramente especificado puede generar inconsistencias que van desde el diseño de las pruebas hasta cómo se difunden y utilizan sus resultados. El segundo paso es precisar quiénes serán los usuarios de la evaluación, los productos que se derivarán de sus resultados y el uso esperado para cada uno de ellos. Estas decisiones son fundamentales para determinar tres características clave del diseño de la evaluación: su alcance –es decir, si será censal o muestral–, los grados o años y el contenido y las habilidades a evaluar y, por último, la frecuencia con la que se llevará a cabo.

A 30 años del inicio de las evaluaciones de aprendizajes a gran escala en Argentina, su sostenimiento por parte de los distintos gobiernos, más los recientes planes nacionales de evaluación 2020-2021 y 2023-2024, promovidos desde el Ministerio de Educación y aprobados por el Consejo Federal de Educación, reflejan un consenso sobre la importancia de su implementación. Esto también se evidencia en la cantidad de jurisdicciones que han llevado a cabo evaluaciones propias: desde 2019, un total de 12 jurisdicciones lo han hecho, mientras que otras 3 tienen previsto hacerlo en 2023.

Sin embargo, una de las debilidades principales de las evaluaciones a gran escala en Argentina durante todos estos años ha radicado en la falta de definiciones acerca de cómo se espera que ocurran las mejoras en el sistema educativo a partir de sus resultados. Esta situación de indefinición sobre los propósitos de la evaluación luego se refleja principalmente en dos áreas: primero, en la falta de previsibilidad respecto de los productos que se derivarán de sus resultados y el uso esperado para cada uno de ellos; y segundo, en el diseño de



Este documento fue elaborado en el marco del Laboratorio de Innovación y Justicia Educativa del Programa de Educación de CIPPEC (EduLab), que tiene por objetivo aportar estrategias de transformación integral de la educación en Argentina para lograr una mayor justicia educativa que garantice trayectorias completas y aprendizajes de calidad para todos los niños, niñas y jóvenes.

las pruebas, donde la cobertura censal o muestral, los dominios o grados evaluados y la frecuencia de evaluación han ido cambiando sin que se tenga muy claro el porqué.

En este marco, este documento presenta un conjunto de recomendaciones para el diseño de una política nacional de evaluación de aprendizajes a gran escala que sirva como una herramienta eficiente al servicio de la mejora de la calidad educativa. Propone la necesidad de planificar las evaluaciones a gran escala en un horizonte temporal de 10 años, período que posibilitaría implementar distintos tipos de dispositivos evaluativos, examinar diferentes dominios sin sobrecargar a las escuelas y estudiantes con una gran cantidad de pruebas en un mismo año, y realizar una definición más precisa de usuarios y usos esperados de los resultados de las pruebas.

En concreto, esta política nacional de evaluación se basaría en: 1) implementar evaluaciones censales quinquenales para elaborar un mapa nacional de aprendizajes y trayectorias educativas a nivel de escuela; 2) continuar las evaluaciones de diagnóstico sobre las áreas fundamentales de Matemática y Lengua, con operativos muestrales bienales al final de la primaria y la secundaria; 3) iniciar la evaluación de habilidades socioemocionales también en los niveles primario y secundario; 4) fortalecer los esquemas de coordinación entre Nación y las jurisdicciones para evitar superposiciones de evaluaciones y fomentar su complementariedad; y 5) promover procesos de evaluación formativa en el aula de manera de reforzar el uso y calidad de este dispositivo pedagógico por parte de los y las docentes.



El uso de un lenguaje que no discrimine, que no reproduzca estereotipos sexistas y que permita visibilizar todos los géneros es una preocupación de quienes trabajaron en esta publicación. Dado que no hay acuerdo sobre la manera de hacerlo en castellano, se consideraron tres criterios a fines de hacer un uso más justo y preciso del lenguaje: 1) evitar expresiones discriminatorias, 2) visibilizar el género cuando la situación comunicativa y el mensaje lo requieren para una comprensión correcta y, 3) no visibilizarlo cuando no resulta necesario.

## Introducción

En la actualidad existe un acuerdo en muchos países y entre los principales organismos internacionales que trabajan en el mejoramiento de las políticas públicas en educación sobre la importancia que tienen las “evaluaciones de aprendizaje a gran escala” como una herramienta esencial para generar insumos para la mejora de la calidad educativa. Este tipo de evaluaciones educativas son aquellas implementadas a nivel de sistema, a diferencia del otro gran grupo de evaluaciones que se utilizan en el aula como parte del proceso de enseñanza.

Las evaluaciones a gran escala brindan una foto de los logros de aprendizaje de un grupo determinado de estudiantes en un año determinado y en un número limitado de dominios. A menudo se clasifican como evaluaciones subnacionales, nacionales o internacionales (según el o los sistemas educativos sobre los que se implementen), son uniformes y estandarizadas en sus contenidos, procesos y tiempos de administración, metodología de corrección y generación de resultados. Por lo general, se basan en muestras, aunque, en las últimas décadas, un número creciente de países han adoptado un enfoque basado en censos (Verger et al., 2018; UNESCO, 2019).

En Argentina, la Ley Federal de Educación de 1993 dispuso que el Ministerio de Educación de la Nación, en coordinación con las provincias, evaluara de manera permanente el sistema educativo en todas las jurisdicciones del país. Esta disposición dio origen a los Operativos Nacionales de Evaluación (ONE) que, ese mismo año, comenzaron a implementarse, evaluando de manera muestral a estudiantes de 6° grado del nivel primario y los años 2° y 3° del secundario.

Hoy, a 30 años del inicio del sistema de evaluación de calidad educativa a gran escala implementado por el Ministerio de Educación de la Nación, lo que se ha llamado la “crisis del aprendizaje” representa un desafío prioritario en la agenda internacional de desarrollo educativo. En este contexto, es oportuno reflexionar sobre qué elementos tienen que considerarse para que estas evaluaciones puedan realmente ser una política pública efectiva a fin de alcanzar el objetivo principal del sistema educativo: que la totalidad de niños, niñas y adolescentes pueda finalizar la escolarización primaria y secundaria, garantizando “resultados de aprendizaje pertinentes y efectivos”, de acuerdo con lo estipulado por la meta 4.1 de los Objetivos de Desarrollo Sostenible.

Este documento tiene como objetivo principal brindar un conjunto de lineamientos para la consolidación de una política nacional de evaluación de aprendizajes a gran escala en Argentina que sirva como herramienta eficiente para el mejoramiento de la calidad educativa. Para esto, en la primera sección se presentan algunos datos que caracterizan la situación actual de aprendizajes en el nivel primario y secundario. También, se aborda el rol que se asigna actualmente a las evaluaciones educativas para superar esta crisis. Por su parte, la segunda sección profundiza en los propósitos teóricos de las evaluaciones a gran escala, mostrando que una definición precisa del “para qué evaluamos” es el primer paso para luego determinar las características técnicas de los dispositivos de evaluación y que estos alcancen los objetivos que se proponen. Luego, la tercera sección repasa la evolución y estructura actual de las evaluaciones nacionales, jurisdiccionales e internacionales en Argentina a fin de entender cuál es la situación de nuestro país respecto de la temática. En la cuarta sección se analizan los dispositivos nacionales de evaluación de aprendizajes a gran escala en nuestro país con relación a: propósito, usuarios y tipos de usos previstos, cobertura, años y contenidos a evaluar y frecuencia. La quinta sección elabora un conjunto de recomendaciones para consolidar una política nacional de evaluación a gran escala que sirva al mejoramiento de los aprendizajes de las y los estudiantes en Argentina. Por último, se proponen a algunas reflexiones finales de temas a considerar para avanzar en la implementación de las recomendaciones brindadas.

## La “crisis del aprendizaje” y el rol actual de las evaluaciones de aprendizaje a gran escala

En el marco de la Naciones Unidas, más de 190 países aprobaron en 2015 los Objetivos de Desarrollo Sostenible (ODS), donde se reserva un lugar especial para la educación, ya que reconoce la centralidad del proceso educativo para poder alcanzar ese desarrollo. El objetivo número 4 pone el foco en la educación, en el cual resalta que no basta con que las personas lleguen y “pasen” por la escuela, sino que es crucial que adquieran conocimientos y competencias esenciales para su desarrollo en el futuro. En la terminología de los ODS, los sistemas educativos deben brindar a las personas la posibilidad de adquirir aprendizajes “efectivos y relevantes”.

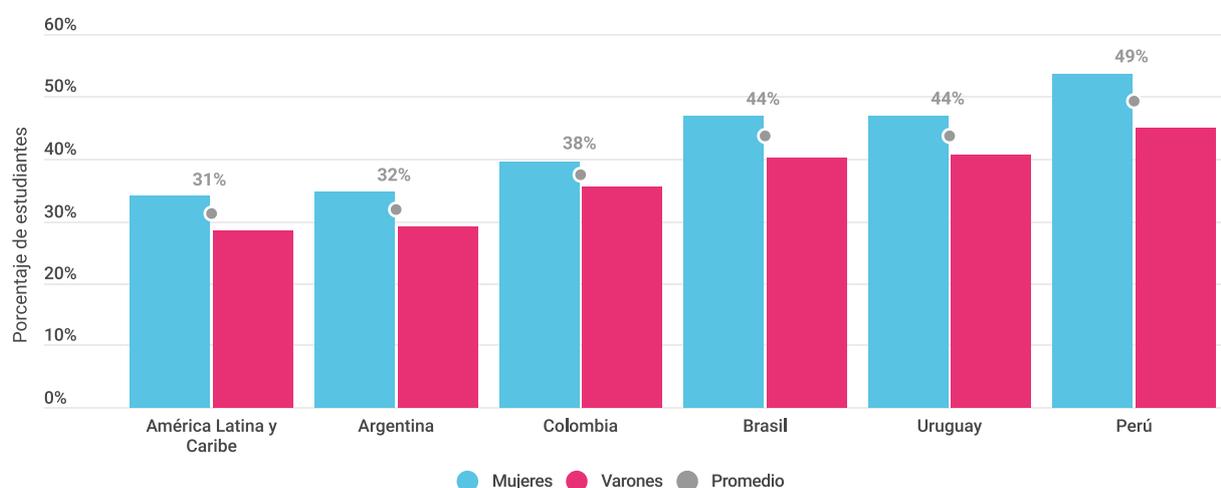
Esta meta, cuya relevancia fue acordada ocho años atrás, manifestaba la preocupación mundial por lo que se ha llamado la “crisis del aprendizaje”, situación que se caracteriza por el hecho de que una elevada proporción de niños, niñas y adolescentes que están incluidos en el sistema educativo no alcanzan resultados de aprendizajes satisfactorios (UNESCO, 2013; Banco Mundial, 2018; UNICEF, 2021).

Se trata de un fenómeno transversal a la educación primaria y secundaria. En Argentina, en promedio, sólo 3 de cada 10 estudiantes que están en 6° grado alcanzaron un desempeño por encima del nivel mínimo de competencia en Lectura, según el Estudio Regional Comparativo y Explicativo (ERCE) de la UNESCO en 2019. Este desempeño es muy similar al alcanzado, en promedio, por los países de América Latina que participaron en el estudio y está por debajo de varios países del Cono Sur, como son Colombia, Brasil, Uruguay y Perú (**Gráfico 1**).

En esta misma línea, el panorama que se presenta en la educación secundaria también es preocupante. Las evaluaciones PISA administradas en adolescentes de 15 años en 2018 mostraron que, en Argentina, apenas 5 de cada 10 estudiantes tuvieron un desempeño por encima del estándar mínimo esperado para la edad. Una vez más, este valor es prácticamente idéntico al alcanzado, en promedio, por los y las estudiantes de los países de América Latina y el Caribe que participaron en la prueba. Adicionalmente, con estas pruebas se puede constatar la brecha de aprendizajes existente con estudiantes de algunos de los países más desarrollados en América del Norte y Europa, donde casi 8 de cada 10 alcanzan este desempeño mínimo esperado (**Gráfico 2**).

GRÁFICO 1

Porcentaje de estudiantes en 6° grado de primaria que alcanzan los aprendizajes mínimos esperados en Lectura, en América Latina, Argentina y otros países seleccionados (Pruebas ERCE 2019)

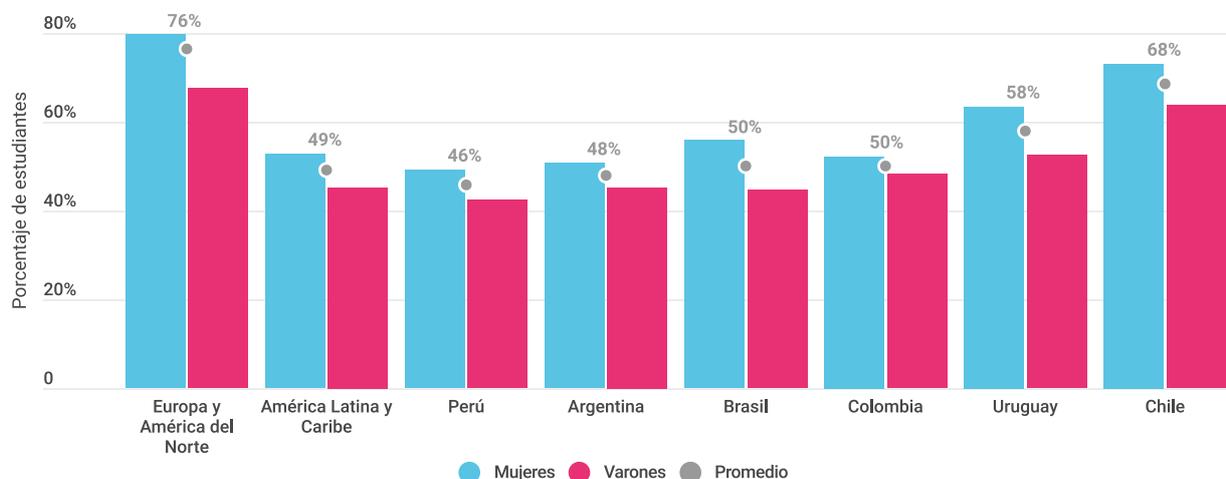


Nota: los “aprendizajes mínimos” corresponden a estudiantes cuyos resultados se ubicaron en los niveles III y IV de la prueba de referencia.

Fuente: elaboración propia en base a datos obtenidos del explorador del ODS 4 por el Instituto de Estadística de la UNESCO. Los datos de América Latina y Caribe corresponden a promedios simples de los 16 países que participaron de la prueba ERCE en 2019.

GRÁFICO 2

Porcentaje de estudiantes de 15 años que alcanzan los aprendizajes mínimos esperados en Lectura, en Europa y América del Norte, América Latina y el Caribe, Argentina y países seleccionados (Pruebas PISA 2018)



Nota: los "aprendizajes mínimos" corresponden a estudiantes cuyos resultados se ubicaron por encima del nivel II inclusive de la prueba de referencia.

Fuente: elaboración propia en base a los datos obtenidos de WIDE Education Inequalities, 2021 (Ver <https://www.education-inequalities.org/>). Los datos de América Latina y Caribe corresponden a promedios simples de los 10 países que participaron de la prueba PISA en 2018, mientras que los datos de Europa y América del Norte corresponden a promedios simples de los 41 países que participaron de la prueba en ese mismo año.

La crisis del aprendizaje, que para el inicio de 2020 estaba bien documentada en cuanto a su alcance, a partir de la pandemia se agravó en muchos contextos (Perusia, 2021). La razón principal es el cierre abrupto de escuelas que se inició en febrero de 2020 en casi todos los países y que, para abril, ya afectaba a alrededor de 1.500 millones de estudiantes, lo que equivalía al 90% de la población estudiantil mundial (Srivastava et al, 2020).

A propósito de esto, estudios recientes realizados en países como Brasil, Países Bajos y Reino Unido confirman que el cierre de escuelas tuvo un efecto de deterioro en los aprendizajes. En otras palabras, durante ese año de pandemia las y los estudiantes aprendieron mucho menos que sus pares de años anteriores (Perusia y Cardini, 2021; Linchand et al, 2021; Engzell et al, 2021; Renaissance Learning, Education Policy Institute, 2021). En Estados Unidos, el Centro Nacional de Estadísticas Educativas (NCES, por sus siglas en inglés), confirmó, a partir de un estudio realizado en 2022, que los puntajes promedio de los y las estudiantes de 9 años disminuyeron 5 puntos en Lectura y 7 puntos en Matemática en comparación con 2020: es la mayor disminución de puntaje promedio en Lectura desde 1990 y la primera disminución de puntaje en Matemática de la serie histórica que se inicia en los '70 (Nation's Report Card, 2022).

En el caso de Argentina, una vez atravesada la pandemia, las evaluaciones a gran escala realizadas dan cuenta de un escenario donde persisten grandes dificultades para que una fracción considerable de estudiantes alcance niveles satisfactorios en Lengua y Matemática, tal como sucedía antes de la pandemia. En algunos casos, estas dificultades se profundizaron, de manera consistente con lo que sucedió en otros países. Esto es lo que mostraron las evaluaciones nacionales que implementa el Ministerio de Educación de la Nación, denominadas Aprender, donde en 2022 se verificó un incremento del porcentaje de estudiantes del nivel secundario que se ubican en los dos niveles más bajos de desempeño con respecto a 2019, tanto en Lengua como en Matemática (Ministerio de Educación, 2023b). En Lengua, se pasó del 38% al 43%; en Matemática, del 71% al 82%.

## Las evaluaciones educativas como instrumento para la superación de la “crisis del aprendizaje”

Recientemente y con posterioridad a la pandemia, varios organismos especializados en educación acordaron que las políticas necesarias para desarrollar y acelerar la incorporación de aprendizajes deben basarse en al menos cinco pilares, desarrollados a continuación en la **Figura 1** (Banco Mundial et al, 2022).

El marco “RÁPIDO” recomienda, en su segundo pilar, realizar regularmente dos tipos de evaluaciones educativas a distintos niveles:

1. Evaluaciones diagnósticas y formativas a nivel de aula, implementadas por los y las docentes como parte del proceso de enseñanza y que tienen como propósito introducir los ajustes necesarios para mejorar los resultados de los y las estudiantes, al dar una retroalimentación sobre el progreso del aprendizaje individual.
2. Evaluaciones a gran escala a nivel del sistema educativo, implementadas usualmente por un ministerio u organismo evaluador con competencia para realizar esa tarea a nivel nacional o jurisdiccional, que consisten en evaluaciones estandarizadas aplicadas a una muestra o a toda la población objetivo de estudiantes de un determinado grado o nivel educativo. La estandarización implica que se aplican las mismas pruebas, en idénticas condiciones (i.e., día, horario, cantidad de tiempo para realizar la prueba) para luego ser puntuadas con criterios comunes. Es decir, se busca garantizar la homogeneidad del proceso evaluativo y de los resultados obtenidos.

Usualmente, ambos tipos de evaluaciones coexisten en la mayoría de los sistemas educativos del mundo y tienen un objetivo común: generar información para que distintos actores del sistema educativo puedan actuar de una manera más eficiente para facilitar la incorporación de aprendizajes por parte de las y los estudiantes.

Este documento se enfoca particularmente en las evaluaciones a gran escala, cuya relevancia como herramientas para monitorear y mejorar la calidad de los sistemas educativos ha ganado peso en las últimas dos décadas (Clarke, M., y Luna-Bazaldúa, D., 2021). Su importancia radica en la capacidad que tienen para responder, de acuerdo con las características que adopten, preguntas que son clave para el diseño de la política educativa (Greaney, V. y Kellaghan, T., 2016):

- ¿Qué tan bien están aprendiendo los y las estudiantes en el sistema educativo evaluado?  
¿Están cumpliendo con los estándares de aprendizaje específicos?

FIGURA 1  
Marco RÁPIDO para la acelerar de la incorporación de los aprendizajes



- ¿Qué fortalezas y debilidades particulares existen en el conocimiento y las capacidades cognitivas de los y las estudiantes?
- ¿Hay subgrupos particulares que se desempeñan peor que otros? ¿Existen disparidades, por ejemplo, entre el desempeño de niños y niñas, estudiantes de localidades urbanas y rurales, o estudiantes de diferentes grupos lingüísticos?
- ¿Qué factores están asociados con el rendimiento de las y los estudiantes?
- ¿En qué medida varía el rendimiento según las características del entorno de aprendizaje (por ejemplo, los recursos escolares o la preparación del docente) o según las circunstancias del hogar de los y las estudiantes?
- ¿Cambia el rendimiento de los y las estudiantes con el tiempo? ¿Qué factores están vinculados a los cambios en su rendimiento a lo largo del tiempo?

Responder estos interrogantes con datos precisos y confiables es esencial para enfocar los sistemas educativos en el logro de los aprendizajes. En otras palabras, la evaluación de aprendizajes a gran escala es necesaria para hacer de estos aprendizajes un objetivo preciso de política educativa, a partir de la generación de datos que guíen la acción (Banco Mundial, 2018).

## ¿Para qué hacer evaluaciones de aprendizajes a gran escala?

La definición del propósito principal de una evaluación a gran escala es el primer paso para diseñar un instrumento evaluativo que realmente tenga la potencialidad de ser efectivo. Como señalamos anteriormente, no contar con un “para qué” definido con claridad puede llevar a múltiples errores o inconsistencias que pueden perjudicar al diseño de las pruebas, así como también a la posterior utilización y difusión de sus resultados. A continuación, la **Tabla 1** detalla tres propósitos bajo los cuales la investigación sobre evaluación educativa a gran escala suele categorizar los distintos usos previstos de los resultados.

En la literatura educativa, las evaluaciones diseñadas con los propósitos 2 y 3 son conocidas como evaluaciones de “implicancias altas” (*high-stake assessments*), como una manera de poner de relieve que los resultados tienen una consecuencia directa y tangible sobre los sujetos evaluados o sobre las instituciones educativas involucradas y/o su personal. En contraposición, a las evaluaciones diseñadas de acuerdo con el propósito 1 se las conoce como evaluaciones de “implicancias bajas” (*low-stake assessments*), dado que los sujetos evaluados, sus docentes o instituciones no observan una consecuencia directa basada en los resultados (Ravela, 2006).

	Evaluaciones de implicancias bajas	Evaluaciones de implicancias altas	
¿Cuál es su propósito general?	1. Evaluaciones de carácter diagnóstico.	2. Evaluaciones para establecer mecanismos de competencia.	3. Evaluaciones de certificación de los logros educativos de los y las estudiantes (exámenes nacionales).
¿Qué las caracteriza?	Permiten establecer niveles de aprendizajes para utilizar como insumo en el diseño intervenciones para la mejora a distintos niveles (nacional, subnacional, escuela, estudiantes).	Permiten establecer mecanismos de competencia entre escuelas y/o docentes.	Permiten establecer quiénes han alcanzado los conocimientos y desempeños necesarios para aprobar un curso o nivel educativo.

<p><b>¿Qué usos previstos específicos posibilitan?</b></p>	<ul style="list-style-type: none"> <li>• Desarrollar políticas de apoyo a las escuelas, generales y focalizadas, para la mejora de los resultados.</li> <li>• Producir conocimiento sobre logros educativos y factores asociados a estos.</li> <li>• Orientar la definición del currículo y la elaboración de materiales pedagógicos.</li> <li>• Fortalecer programas de formación docente.</li> <li>• Rendir cuentas respecto de los resultados de la gestión educativa.</li> </ul>	<ul style="list-style-type: none"> <li>• Otorgar incentivos económicos, a escuelas y/o docentes, según el desempeño de sus estudiantes.</li> <li>• Propiciar un mercado competitivo utilizando los resultados para construir rankings de escuelas según su desempeño y hacerlos públicos. Esto puede fomentar la responsabilidad de las escuelas por sus resultados; entregar a las familias información para la toma de decisiones sobre la escuela a la cual enviar a sus hijos/as, y promover la competencia entre los centros educativos en pos de obtener mejores resultados.</li> </ul>	<ul style="list-style-type: none"> <li>• Acreditar la finalización exitosa de un nivel educativo.</li> <li>• Establecer un mecanismo de selección para acceder al nivel superior (usualmente terciario o universitario).</li> </ul>
--	--	---	---

Fuente: elaboración propia en base a Ravela (2006) e IIPE-UNESCO (2021b).

Si bien en América Latina hay algunas experiencias de usos de resultados de implicancias altas, la cultura de la evaluación a gran escala en Argentina –al igual que en la mayoría de los países de la región– se ha desarrollado y construido desde la visión de las evaluaciones de carácter diagnóstico (**Recuadro 1**). Por esta razón, sumado al hecho de que no hay evidencia de que los usos de implicancias altas tengan más consecuencias positivas que negativas para los sistemas educativos, este documento decide enfocarse en los cambios que requiere un sistema de evaluación a gran escala concebido desde esa perspectiva de usos de implicancias bajas, para que pueda actuar como catalizador de la mejora de los aprendizajes.

### RECUADRO 1

#### Usos de los resultados de las evaluaciones en América Latina entre 2010 y 2021

Un estudio reciente sobre los usos de los resultados de las evaluaciones a gran escala en Argentina, Brasil, Chile, Ecuador, México y Uruguay (IIPE-UNESCO, 2021b), que abarcó el período 2010-2021, ha mostrado que estos se usan, más comúnmente, como insumos para el diseño de distintos tipos de programas de mejoras de logros sobre áreas de conocimiento específicas (Lengua y Matemática, por lo general).

Con menor frecuencia, en algunas circunstancias los resultados se utilizan en distintas temáticas enfocadas a la formación docente: programas de formación docente continua, como insumos para la preparación de recursos pedagógicos, o para diseñar convocatorias a universidades o centros educativos que prestan servicios de profesionalización docente.

Un tercer tipo de uso es el que se vincula con el propósito de seleccionar establecimientos educativos específicos para ser incluidos en programas de apoyo para el mejoramiento de sus resultados. Usualmente, el criterio aquí es seleccionar aquellas escuelas con un rendimiento promedio por debajo de un umbral determinado, observando también otros indicadores de eficiencia interna (por ejemplo, repitencia, abandono) para luego poder brindar distintos tipos de recursos que permitan mejorar esa situación.

Los ejemplos reseñados anteriormente constituyen tipos de usos conectados con sistemas de evaluación diseñados bajo el paradigma de implicancias bajas. Sin embargo, también existen algunos casos de usos con implicancias altas. En este sentido, una primera serie de ejemplos responde al objetivo de establecer incentivos a los y las docentes. En México, los resultados de las evaluaciones se utilizaron, entre varios usos, como parte de un esquema organizador de la carrera magisterial y de distribución de incentivos hacia el cuerpo docente. En Brasil, en los estados de Ceará y de San Pablo y el municipio de Sobral (Ceará), también se consideran esquemas de incentivos a docentes sobre la base de los resultados de sus propios sistemas de evaluación. Un tercer antecedente de este tipo de uso se verifica en Chile, donde los establecimientos mejor evaluados, a partir de un índice en que los resultados de la evaluación nacional (SIMCE) tienen un gran peso, se hacen acreedores por dos años de la Subvención por Desempeño de Excelencia destinada a otorgar incentivos económicos a docentes y asistentes escolares.

Por otra parte, otras experiencias refieren a la existencia de mecanismos que contienen los resultados de la evaluación para asignar recursos financieros a las instituciones educativas. Este tipo de mecanismos existen en Brasil, donde hay programas de transferencias de recursos del gobierno federal hacia los estados y municipios asociados con el Índice de Desempeño de la Educación Básica (IDEB, por sus siglas en portugués).

Al mismo tiempo, otro tipo de uso lo constituye la publicación de resultados por establecimientos en formato de rankings. En Brasil, si bien ha perdido presencia en los últimos años, esta práctica ha sido utilizada. En Chile, a su vez, la publicación de los ordenamientos de escuelas según su desempeño en el SIMCE es un mecanismo instalado como parte de los procesos de difusión de la información.

Por último, también existieron y/o existen evaluaciones a gran escala cuyos resultados se utilizan para certificar conocimientos de los y las estudiantes y así acreditar la finalización de un nivel educativo, generalmente el secundario, o como mecanismo de acceso y admisión a la educación superior. Los países donde se presentan este tipo de evaluaciones son Ecuador, con las pruebas Ser Bachiller; Brasil, con el Examen Nacional de la Educación Secundaria (ENEM); México cuenta con el Instrumento para el Diagnóstico de Alumnos de Nuevo Ingreso a la Educación Secundaria (IDANIS); y Chile, con distintas variantes de pruebas estandarizadas que son parte del proceso de selección para el ingreso a la educación universitaria, siendo la Prueba de Selección Universitaria (PSU) un instrumento que se aplicó entre 2003 y 2020.

## El diseño de las evaluaciones de aprendizajes a gran escala

Una vez determinado el propósito principal de la evaluación educativa, se debería definir con la mayor precisión posible los usos previstos y usuarios de los resultados para, de esa manera, poder determinar las características de su diseño.

Las evaluaciones de implicancias bajas establecen como propósito el diagnóstico para la mejora. Sin embargo, muchas veces estos procesos evaluativos realizan definiciones difusas sobre quiénes serán los usuarios de la información y los objetivos más específicos que persiguen. Por esta razón, el desarrollo de una política nacional de evaluación de aprendizajes a gran escala efectiva debería considerar una trayectoria de toma de decisiones precisas sobre al menos el conjunto de elementos que se presentan a continuación en la **Tabla 2**.

TABLA 2 Trayecto de decisiones para definir las características de una evaluación de aprendizajes a gran escala	
Decisión sobre características clave para el diseño de una evaluación	Aspectos a considerar en la toma de decisiones
1. Definir el propósito general de la evaluación	Los propósitos posibles se definieron en la <b>Tabla 1</b> . En las evaluaciones de carácter diagnóstico, las consecuencias no serán directas ni tangibles en el corto plazo, aunque sí deberían afectar y modificar conductas y procesos de distintos actores dentro del sistema educativo en el mediano plazo.
2. Identificar usuarios de los resultados y tipos de usos previstos por usuarios	Una evaluación de carácter diagnóstico puede tener múltiples usuarios potenciales (responsables de política educativa a distintos niveles, supervisores, directores, docentes, etc.). La definición de estos usuarios y cómo se utilizarán los resultados es clave en esta instancia para determinar el alcance requerido del diagnóstico a realizar y los productos esperados (por ejemplo, informes generales de resultados para autoridades educativas, informes por escuelas para directores y supervisores, informes de desempeño por materias para docentes).
3. Determinar la cobertura de la población objetivo (censal o muestral)	Obtener un diagnóstico sobre el logro educativo de las y los estudiantes a nivel de un sistema educativo determinado sólo requiere una evaluación con una muestra probabilística. Esto quiere decir que no hace falta evaluar a la totalidad de estudiantes del nivel, sino que alcanza con un subgrupo que, por sus características, sea representativo del conjunto. Por el contrario, el diseño de una política universal de apoyo a las instituciones educativas con mayores dificultades puede requerir la realización de una evaluación censal para identificarlas. En consecuencia, esta decisión está ligada a los usos previstos de la información y la definición de los usuarios de esta.
4. Determinar grados/años a evaluar y qué contenidos y habilidades	Definir los grados o años a evaluar y qué contenidos dependerá estrechamente de las preguntas que se quieren responder con la evaluación. Por ejemplo, conocer si el proceso de alfabetización inicial que tiene como objetivo la educación primaria se cumple en tiempo y forma requerirá evaluar Lengua en 3° grado. Sin embargo, si el propósito fuera conocer con qué conocimientos están accediendo los y las estudiantes a la educación secundaria, tendrá más sentido realizar una evaluación al final del último año de la primaria.
5. Definir la frecuencia/ periodicidad bajo la cual es necesario y adecuado evaluar	Para que un sistema de evaluación tenga impacto, debe tener una periodicidad que posibilite que los datos puedan ser analizados, discutidos, comprendidos y utilizados. De esta manera, por ejemplo, realizar evaluaciones una vez por año puede no ser lo adecuado si el sistema no tiene la capacidad de generar los resultados rápidamente y plasmar estos en acciones concretas de mejora. Incorporar la información nueva y traducirla en decisiones y acciones lleva tiempo. Los cambios en el sistema educativo requieren más tiempo aún. Estos conceptos deben ser tenidos en cuenta al definir la frecuencia de las evaluaciones.

Fuente: elaboración propia sobre la base de Ravela et al, (2008) y Clarke y Luna-Bazaldua (2021).

## La evaluación de aprendizajes a gran escala en Argentina: dispositivos nacionales, internacionales y jurisdiccionales

En los países suelen implementarse dos tipos de evaluaciones de aprendizajes a gran escala: evaluaciones nacionales y evaluaciones internacionales. En las últimas dos décadas, estas se han implementado cada vez más y han ido ganando espacio entre países de distinto nivel de desarrollo: entre 2013 y 2017, estos dispositivos fueron aplicados en al menos 177 países en la educación primaria y en 150 países en el nivel educativo secundario (UNESCO, 2019).

La necesidad de contar con estos dos tipos de evaluaciones radica en el objetivo de caracterizar de la mejor manera los aprendizajes de los y las estudiantes. Por un lado, las evaluaciones nacionales permiten conocer cuánto saben con relación a los contenidos o estándares nacionales y cómo evolucionan los saberes en el tiempo. Por el otro, las evaluaciones internacionales permiten contar con información de los aprendizajes con respecto a ciertos estándares internacionales, posibilitando adicionalmente la comparación con sistemas educativos de otros países. De esta manera, el diagnóstico nacional se ve enriquecido a partir de la incorporación de la dimensión internacional de comparación.

Luego, en el contexto federal argentino, como en otros países donde la gestión educativa es responsabilidad primordial de las jurisdicciones, se suma un tercer tipo de evaluación: las evaluaciones jurisdiccionales. Las evaluaciones elaboradas a este nivel pueden ser útiles para complementar las que se realizan a nivel nacional, a partir de información más específica para los sistemas educativos subnacionales.

A continuación, se presenta un estado de situación correspondiente a cada uno de los tres tipos de evaluación, para luego concluir en el análisis de los planes nacionales de evaluación aprobados en 2021 y 2022 en Argentina como mecanismos que buscan coordinar la implementación de estos dispositivos.

### Las evaluaciones nacionales de aprendizajes en Argentina

La Ley Federal de Educación (Ley N° 24.195), sancionada en 1993, dispuso que el Ministerio de Educación de la Nación, en coordinación con las provincias, sea el responsable de garantizar la calidad del sistema educativo mediante su evaluación permanente. Para eso, le otorgó la responsabilidad de diseñar un sistema de evaluación y control periódico de la calidad, a concertar en el ámbito del Consejo Federal de Educación. Esta disposición dio origen al Operativo Nacional de Evaluación (ONE), que comenzó a implementarse ese mismo año, evaluando de manera muestral a los y las estudiantes de 6° grado del nivel primario y de 2° y 3° año del nivel secundario en las asignaturas de Lengua y Matemática.

Como se observa en la **Tabla 3**, el ONE se implementó anualmente de forma ininterrumpida hasta el 2000, aunque con algunos cambios en sus características. En primer lugar, a partir de 1994, además de los y las estudiantes anteriormente señalados, comenzó a evaluarse a los y las que estuvieran cursando el último año del nivel secundario. A su vez, se incorporaron las asignaturas de Ciencias Sociales y Ciencias Naturales. Un año después, se comenzó a evaluar a estudiantes de 3° grado del nivel primario y, a partir de 1997, el operativo adquirió carácter censal para ciertos grados y asignaturas.

La periodicidad anual que sostuvo el dispositivo a lo largo del período 1993-2000 se interrumpió en 2001. Luego de un año sin implementarse, volvió a realizarse en 2002 con las mismas características que había adoptado en su última edición, sin embargo, perdió su carácter anual, ya que volvió a desarrollarse recién en 2005 y, posteriormente, en 2007.

TABLA 3

Dispositivos nacionales de evaluación de aprendizajes a gran escala en Argentina, por nivel educativo y grados/años de estudio. Evaluaciones implementadas entre 1993 y 2022 y planeadas para 2023 y 2024

Año	Nombre del operativo	Nivel primario								Nivel secundario							
		3° grado				6° grado				2°/3° año				5°/6° año			
1993	ONE					M	L			M	L						
1994	ONE					M	L	CS	CN	M	L	CS	CN	M	L		
1995	ONE	M	L			M	L	CS	CN	M	L			M	L		
1996	ONE	M	L			M	L	CS	CN	M	L			M	L		
1997	ONE	M	L			M	L	CS	CN	M	L			M	L		
1998	ONE	M	L			M	L	CS	CN	M	L			M	L	CS	CN
1999	ONE	M	L			M	L	CS	CN	M	L			M	L	CS	CN
2000	ONE	M	L	CS	CN	M	L	CS	CN	M	L			M	L	CS	CN
2001	-																
2002	ONE	M	L	CS	CN	M	L	CS	CN	M	L			M	L	CS	CN
2003	-																
2004	-																
2005	ONE	M	L	CS	CN	M	L	CS	CN	M	L			M	L	CS	CN
2006	-																
2007	ONE	M	L	CS	CN	M	L	CS	CN	M	L			M	L	CS	CN
2008	-																
2009	-																
2010	ONE	M	L	CS	CN	M	L	CS	CN	M	L	CS	CN	M	L	CS	CN
2011	-																
2012	-																
2013	ONE	M	L	CS	CN	M	L	CS	CN	M	L	CS	CN	M	L	CS	CN
2014	-																
2015	-																
2016	Aprender	M	L			M	L			M	L	CS	CN	M	L	CS	CN
2017	Aprender							CS	CN					M	L		
2018	Aprender					M	L										
2019	Aprender													M	L	CS	CN
2020	-																
2021	Aprender					M	L										
2022	Aprender					M	L							M	L		
2023	Aprender					M	L										
2024	Aprender													M	L		

Nota: el sombreado en la asignatura indica aplicación censal sobre la población objetivo. Las letras en la tabla identifican a la asignatura: M=Matemática; L=Lengua; CS=Ciencias Sociales; CN=Ciencias Naturales.  
Fuente: elaboración propia sobre la base del Ministerio de Educación de la Nación.

En términos normativos, en el 2010 se aprobó el documento “Lineamientos para la evaluación de componentes de la calidad educativa 2010-2020” (Res. CFE N° 116/10), mediante el cual se estableció que el ONE debía aplicarse cada tres años y se incorporaron instrumentos de recolección de información complementarios con el objetivo de dar cuenta de una visión más integral de la calidad educativa. Bajo esta nueva normativa, el ONE se aplicó en dos oportunidades más: 2010 y 2013.

En 2016 el ONE fue reemplazado por la evaluación Aprender, que retomó la frecuencia anual que había tenido inicialmente su antecesor (Res. CFE N° 280/16). En su primera aplicación, adquirió características similares a la última evaluación implementada: evaluó a estudiantes de los niveles primario y secundario en las áreas de Lengua, Matemática, Ciencias Sociales y Ciencias Naturales, con un alcance censal para algunos grados y asignaturas, y muestral para otros.

No obstante, a partir del 2017, algunas características del dispositivo experimentaron modificaciones. En primer lugar, desde ese año se evaluó el último grado o año de cada nivel educativo, aunque no siempre se evaluaron ambos niveles. En segundo lugar, se restringieron las asignaturas evaluadas. Así, en 2018, 2021 y 2022 el dispositivo se limitó a Lengua y Matemática. En tercer lugar, también varió su alcance: en algunos años tuvo carácter censal (2017, 2018 y 2021), mientras que en otros combinó algunos grados y asignaturas con alcance censal y otros con carácter muestral.

## Participación de Argentina en las evaluaciones internacionales de aprendizajes

Además de las evaluaciones nacionales de aprendizajes, la Argentina participó a lo largo del período analizado en varias evaluaciones internacionales. Tal como se mencionó anteriormente, estas proporcionan información valiosa, dado que permiten realizar comparaciones con los resultados de estudiantes de otros países, para así identificar áreas de mejora, respaldar la formulación de políticas basadas en evidencia y estimular la mejora continua. Adicionalmente, aumenta la visibilidad internacional del sistema educativo y permite a especialistas de todo el mundo analizar el desempeño del sistema educativo nacional, dado que la estandarización de las pruebas garantiza la comparabilidad de los resultados.

A continuación, la **Tabla 4** resume la información de los dispositivos en los que participó Argentina entre 1993 y 2022, que en su totalidad fueron de carácter muestral.

Año	Nombre del operativo	Nivel educativo evaluado	Estudiantes evaluados/as	Asignatura evaluada	Consortio responsable de su implementación
1997	PERCE	Nivel primario	3° y 4° grado	Lectura y Matemática	UNESCO
2000	PISA	Nivel secundario	Estudiantes de 15 años	Lectura, Matemática y Ciencias	OCDE
2001	PIRLS	Nivel primario	4° grado	Lectura	IEA
2006	SERCE	Nivel primario	3° y 6° grado	Lectura y escritura, Matemática y Ciencias Naturales	UNESCO
2006	PISA	Nivel secundario	Estudiantes de 15 años	Lectura, Matemática y Ciencias	OCDE
2009	PISA	Nivel secundario	Estudiantes de 15 años	Lectura, Matemática y Ciencias	OCDE
2012	PISA	Nivel secundario	Estudiantes de 15 años	Lectura, Matemática y Ciencias	OCDE
2013	TERCE	Nivel primario	3° y 6° grado	Lectura y escritura, Matemática y Ciencias Naturales	UNESCO
2015*	PISA	Nivel secundario	Estudiantes de 15 años	Lectura, Matemática y Ciencias	OCDE
2018	PISA	Nivel secundario	Estudiantes de 15 años	Lectura, Matemática y Ciencias	OCDE

2019	ERCE	Nivel primario	3° y 6° grado	Lectura y escritura, Matemática, Ciencias Naturales y Habilidades Socioemocionales	UNESCO
2022	PISA	Nivel secundario	Estudiantes de 15 años	Lectura, Matemática y Ciencias	OCDE
<p>*Nota: si bien la evaluación se implementó en el 2015, los resultados del país no fueron publicados por problemas con la muestra. Fuente: elaboración propia sobre la base del Ministerio de Educación de la Nación.</p>					

La primera de ellas fue el Primer Estudio Regional Comparativo y Explicativo (PERCE), desarrollada por la UNESCO en 1997. Allí se evaluaron a estudiantes de 3° y 4° grado de nivel primario en Lectura y Matemática. Posteriormente, Argentina también participó en las tres evaluaciones siguientes organizadas por el organismo (SERCE, TERCE y ERCE). A partir de la segunda edición, los grados evaluados fueron 3° y 6°, y se incorporó el área de Ciencias Naturales. Además, en la última edición se evaluaron las habilidades socioemocionales de las y los estudiantes.

Por otra parte, desde el 2000, Argentina también participa de las evaluaciones PISA, el Programa para la Evaluación Internacional de Estudiantes –por su sigla en inglés– que lleva adelante la Organización para la Cooperación y el Desarrollo Económicos (OCDE) cada tres años. Allí se evalúa a estudiantes de 15 años en Lectura, Matemática y Ciencias. Nuestro país formó parte del estudio en todas las aplicaciones, a excepción del año 2003.

En el año 2001, y por única vez, se aplicó PIRLS, una evaluación desarrollada por la Asociación Internacional para la Evaluación de los Logros Educativos (IEA). Su objetivo es proveer información sobre el desempeño en Lectura de los y las estudiantes de 4° grado del nivel primario y comprender el contexto escolar y familiar en el que se desarrolla su proceso de aprendizaje. Si bien desde su primera implementación la evaluación se realizó cada cinco años, la Argentina no volvió a participar.

## Las evaluaciones jurisdiccionales de aprendizajes en Argentina

El panorama de las evaluaciones de aprendizajes a gran escala en la Argentina se completa con las evaluaciones jurisdiccionales, las cuales comenzaron a desarrollarse a partir de la segunda mitad de la década del 90. Al igual que en el caso de las evaluaciones nacionales, fueron impulsadas por la Ley Federal de Educación, dado que allí se establecía que la atribución de evaluar el sistema educativo pertenecía al ministerio nacional, pero también a las provincias y a la Ciudad Autónoma de Buenos Aires.

Además de ese marco normativo, según Larripa (2010), otras variables contribuyen a explicar el surgimiento de evaluaciones jurisdiccionales. En primer lugar, la instalación de una creciente “cultura de la evaluación”, es decir, la generación de una posición más favorable hacia la evaluación. Al mismo tiempo, como un segundo elemento, las críticas al sistema nacional de evaluación también tuvieron injerencia. En tercer lugar, vale mencionar que existen diferencias en la estructura de ciclos y en el currículum en muchas jurisdicciones con respecto al marco acordado federalmente y considerado en las evaluaciones nacionales. Por último, en cuarto lugar, la posibilidad de obtener financiamiento externo para el desarrollo de sistemas de evaluación acordes a las necesidades de las jurisdicciones también operó como una variable en este sentido.

A treinta años de este primer impulso, se observa que surgieron nuevos dispositivos de evaluación de aprendizajes a gran escala en varias jurisdicciones (**Tabla 5**). En la gran mayoría de los casos, pareciera que estos dispositivos fueron creados con el propósito de evaluar el impacto del COVID 19 sobre el aprendizaje de los y las estudiantes, de forma tal de contar con información actualizada para favorecer el proceso de toma de decisiones de política educativa a nivel jurisdiccional.

TABLA 5 Evaluaciones jurisdiccionales de aprendizajes implementadas en Argentina, por jurisdicción y características seleccionadas entre 2019 y 2022 y planeadas para 2023							
Jurisdicción	Nombre del dispositivo	Año/s de implementación	Frecuencia	Nivel educativo evaluado	Estudiantes evaluados/as	Asignaturas evaluadas	Cobertura
<b>Buenos Aires</b>	-	<b>2022</b>	<b>No especificada</b>	<b>Primario</b>	<b>3° y 6° grado</b>	<b>Prácticas del Lenguaje Matemática</b>	<b>Censal</b>
CABA	FEPBA	2019 y 2021	Bianual	Primario	7° grado	Prácticas del Lenguaje Matemática	Censal
	TESBA	2019 y 2021	Bianual	Secundario	3° año	Lengua y Literatura Matemática	Censal
	<b>Fluidez y comprensión lectora</b>	<b>2022</b>	<b>Anual</b>	<b>Primario</b>	<b>4° grado</b>	<b>Fluidez y comprensión lectora</b>	<b>Muestral</b>
Catamarca	Pensar en Aprender – Aprender a Pensar	2019	No especificada	Secundario	6° año	Lengua Matemática	Censal
Chaco	Dispositivo de Evaluación Chaqueño (DECh)	2019	Anual	Primario Secundario	7° grado 2° año	Lengua y Literatura Matemática Ciencias Sociales Ciencias Naturales Ejes transversales	Censal
Córdoba	PRISMA	2022	Anual	Primario Secundario	3° grado 3° y 5° año	Lengua Matemática	Muestral
Corrientes	<b>Corrientes Operativo Provincial de Evaluación (COPE)</b>	<b>2022</b>	<b>Anual</b>	<b>Primario Secundario</b>	<b>3° y 6° grado 3° y 6° año</b>	<b>Lengua y Literatura Matemática</b>	<b>Muestral</b>
	<b>Estudio comparativo de aprendizaje</b>	<b>2022</b>	<b>No especificada</b>	<b>Secundario</b>	<b>Ciclo Básico</b>	<b>Lengua y Literatura Matemática</b>	<b>Muestral</b>
Entre Ríos	<b>ALEER</b>	<b>2022</b>	<b>No especificada</b>	<b>Primario</b>	<b>2° grado</b>	<b>Lectura</b>	<b>Muestral</b>
Formosa	Evaluaciones procesuales	2019, 2020, 2021 y 2022	Anual	Primario	3° y 6° grado	Lengua Matemática	Muestral
La Pampa	<b>Monitoreo de Aprendizajes Pampeanos (MAP)</b>	<b>2022</b>	<b>No especificada</b>	<b>Primario</b>	<b>3° y 5° grado</b>	<b>Lengua Matemática</b>	<b>Muestral</b>
Jujuy	Evaluación Saber Más	2021	No especificada	Secundario	3° año	Lengua Matemática	Censal

Jurisdicción	Nombre del dispositivo	Año/s de implementación	Frecuencia	Nivel educativo evaluado	Estudiantes evaluados/as	Asignaturas evaluadas	Cobertura
La Rioja	<b>Evaluación Formativa para la Etapa de Intensificación de la Enseñanza</b>	<b>2022</b>	<b>Anual</b>	<b>Primario</b>	<b>3°, 5° y 7° grado</b>	<b>Lengua Matemática Ciencias Sociales Ciencias Naturales</b>	-
	<b>Evaluación Formativa de los Aprendizajes en el Ciclo Básico</b>	<b>Prevista para 2023</b>	<b>No especificada</b>	<b>Secundario</b>	<b>1° y 3° año</b>	<b>Lengua Matemática Biología, Físico-Química Ciudadanía Ciencias Sociales</b>	-
Mendoza	Relevamiento Provincial de Aprendizaje (RPA)	2021 y 2022	Anual	Primario Secundario	4° grado 2° año	Lengua Matemática	Muestral
	Censo de Fluidez Lectora	2021 y 2022	Anual	Primario Secundario	2° a 7° grado 1° a 6° año	Comprensión lectora	Censo
Misiones	<b>Evaluación Provincial de Logros de Aprendizaje</b>	<b>2021</b>	<b>No especificada</b>	<b>Primario</b>	<b>4° grado 3° año</b>	<b>Lengua Matemática</b>	<b>Muestral</b>
Santa Cruz	-	<b>Prevista para 2023</b>	<b>No especificada</b>	-	-	-	-
Santa Fe	-	<b>Prevista para 2023</b>	<b>No especificada</b>	<b>Primario</b>	<b>6° grado</b>	<b>Lengua</b>	<b>Muestral</b>

Nota: el uso de la negrita indica que el dispositivo se implementó o se implementará por primera vez ese año.  
Fuente: elaboración propia sobre la base de Ministerio de Educación de la Nación (2023a).

Como se observa en la **Tabla 5**, un total de 12 jurisdicciones han llevado a cabo evaluaciones de aprendizaje propias a lo largo de ese período, y otras 3 tienen previsto hacerlo en 2023. El impulso de la pandemia sobre el desarrollo de nuevos dispositivos se hace evidente si se considera que en el 2019 sólo 5 jurisdicciones (CABA, Catamarca, Córdoba, Entre Ríos y Formosa) contaban con evaluaciones de aprendizajes propias.

Sin embargo, en muchos casos, estas evaluaciones adquirieron características similares a evaluaciones nacionales. En este sentido, la gran mayoría de ellas evalúa las asignaturas de Lengua y Matemática, al igual que Aprender. En algunos casos, además, evalúan a estudiantes del mismo grado o año, lo que implica una superposición de las evaluaciones que redundan en una duplicación de esfuerzos y de información generada. Incluso, es habitual que coincida el año en que se implementan los dispositivos nacionales y provinciales de un mismo nivel, lo que puede suponer una alteración excesiva del calendario escolar y una carga extra de trabajo para las escuelas. Finalmente, en muchos casos se ha verificado que son rápidamente discontinuadas, sin una sostenibilidad en el tiempo.

### El Plan Nacional de Evaluación Educativa como mecanismo de coordinación de los distintos tipos de evaluaciones

En los años 2021 y 2022 se acordaron los planes nacionales de evaluación educativa 2020-2021 y 2023-2024 (Res. CFE N° 396/21 y 435/22, respectivamente). Ambos acuerdos, alcanzados en el marco del Consejo Federal de Educación, reflejan un intento por integrar los distintos dispositivos de evaluación y los múltiples instrumentos de recolección de información existentes.

El objetivo de estos planes nacionales es “componer un estado de situación integral de la educación” (en la terminología del Plan 2023-2024) en nuestro país, haciendo foco en cinco dimensiones del proceso educativo: la enseñanza, el aprendizaje, las trayectorias escolares, la organización escolar y el financiamiento educativo. Como se puede ver en la **Tabla 6**, estas dimensiones se abordan a partir de la articulación de una variedad de metodologías de análisis y estrategias de generación y recolección de información.

Dentro de los componentes enumerados se incluyen evaluaciones de aprendizajes a gran escala –tanto nacionales como internacionales, aunque excluye las jurisdiccionales– y otro tipo de instrumentos de monitoreo, acopio de datos e informes de investigación, que aportan información complementaria al desempeño estudiantil.

Según la resolución que le da origen al plan, de esta manera se busca “adoptar un enfoque integral sobre los diversos componentes de la educación escolar –enseñanza, aprendizajes, trayectorias, organización, recursos materiales e institucionales– y desplegar una perspectiva que incluya y articule una variedad de metodologías de análisis, estrategias de generación de datos y recolección de información, instrumentos de medición, y lecturas analíticas que superen la reducción de la calidad educativa al rendimiento en pruebas estandarizadas” (Res. CFE N° 435/22).

TABLA 6 Componentes del Plan Nacional de Evaluación Educativa 2023-2024	
Componente	Descripción
Aprender	Dispositivo nacional de evaluación de aprendizajes a gran escala y de sistematización de información acerca de algunas condiciones en las que ellos se desarrollan.
Evaluación formativa en línea	Evaluación para estudiantes de mitad de ciclo de la educación primaria y de la educación secundaria, que permite que los y las docentes diseñen trayectos específicos para que sus estudiantes sean evaluados/as en cada área, a partir de los insumos disponibles en una plataforma digital.
La voz de las y los estudiantes sobre sus aprendizajes	Encuesta muestral para estudiantes del último año del nivel primario. Indaga sobre las dimensiones relativas a la percepción de los aprendizajes realizados, al desarrollo de su autonomía, las características del acompañamiento recibido en el hogar, sus usos de las tecnologías digitales, sus vínculos con compañeros/as y equipos directivos y docentes de la escuela.
ERCE	Dispositivo regional latinoamericano de evaluación de los aprendizajes a gran escala a cargo de la UNESCO, y de sistematización de información acerca de factores que pueden asociarse a los logros de aprendizaje.
PISA	Dispositivo internacional de evaluación de los aprendizajes de los estudiantes a gran escala a cargo de la OCDE, y de sistematización de información acerca de algunas condiciones en las que ellos se desarrollan.
Encuesta Juventudes Argentinas	Cuestionario para jóvenes de 15 a 25 años de todas las provincias del país, que tiene por objetivo generar información novedosa y sistemática sobre las percepciones, experiencias personales y familiares, intereses, preferencias y hábitos de los y las jóvenes, con foco en el pasaje de la adolescencia a la adultez.
Indicadores educativos	Análisis descriptivo de los principales indicadores de los niveles educativos del ciclo común y superior no universitario.
Matriz de seguimiento de la inversión educativa	Matriz de seguimiento de la evolución del financiamiento educativo nacional y jurisdiccional. Tiene como objetivo contribuir al acceso de información sistematizada que permita analizar la evolución de los recursos del sistema educativo y conocer los principales componentes que lo constituyen.
Investigación educativa	Producción de conocimiento sobre el sistema educativo a través de la elaboración de informes nacionales sobre temáticas de interés.

Fuente: elaboración propia sobre la base de la Resolución del Consejo Federal de Educación N° 435/22.

## Evaluaciones nacionales de aprendizajes a gran escala de la Argentina: rasgos, desafíos y oportunidades

Con la intención de brindar un marco organizativo común, los planes nacionales de evaluación educativa 2020-2021 y 2022-2023 constituyeron iniciativas que agruparon distintos instrumentos de evaluación y monitoreo que se fueron formulando a lo largo del tiempo junto con algunos dispositivos novedosos. Al mismo tiempo, dichos planes pueden entenderse como el resultado de la continuación de una inercia del desarrollo de los procesos evaluativos y de monitoreo iniciados 30 años atrás, en los que se han ido acumulando iniciativas y generando algunas nuevas. No obstante, este desarrollo se dio sin definir propósitos claros, así como tampoco una articulación clara entre ellos a fin de alcanzar una evaluación integral, para la cual no hay un modelo explícito definido.

Un plan o política nacional de evaluación de aprendizajes a gran escala orientada a mejorar la calidad educativa debería:

- Definir claramente el propósito de las evaluaciones nacionales y los usuarios y usos previstos de los resultados a distintos niveles.
- Reflexionar sobre el sentido y prioridades respecto de la participación en evaluaciones internacionales.
- Fomentar la complementariedad de las evaluaciones jurisdiccionales, reconociendo la autonomía de las provincias y considerando que la utilidad de estos dispositivos radica en la posibilidad de ajustarlos a las particularidades y necesidades de cada una de ellas.

En suma, contar con definiciones precisas sobre el propósito de cada dispositivo –yendo más allá de la definición general de “diagnóstico para la mejora”–, identificar a los usuarios de los resultados y definir los tipos de usos o consecuencias de los resultados y para quiénes es esencial en el objetivo de que la evaluación a gran escala sea una verdadera herramienta para la mejora. Estas precisiones son las que permiten luego avanzar de manera consistente en la determinación de: 1) la cobertura deseada de cada dispositivo, 2) los grados o años a evaluar y los dominios de conocimiento, y 3) la frecuencia bajo la cual es más conveniente implementar los distintos dispositivos.

A continuación, se realiza un análisis de los dispositivos nacionales de evaluación de aprendizajes a gran escala que se implementaron en Argentina, con relación a las definiciones propuestas en el párrafo precedente (que se presentaron con más detalle en la **Tabla 3**), con el objetivo de analizar si esas definiciones son claras, consistentes entre ellas y en el tiempo, de manera que configuren diagnósticos con la potencialidad de impulsar las mejoras que se proponen. Para esto, se analizan las prácticas evaluativas seguidas al respecto en los últimos 30 años junto con la información contenida en el Plan Nacional de Evaluación Educativa 2023-2024.

### Propósitos de las evaluaciones de aprendizajes a gran escala

El marco legal dentro del cual se encuadran las evaluaciones de aprendizajes a gran escala en Argentina (LEN 26.206 de 2006) define a las evaluaciones como un instrumento “para la toma de decisiones tendiente al mejoramiento de la calidad de la educación”, pero no aclara quiénes serían los que toman estas decisiones ni el mecanismo por el cual se espera que se utilicen los resultados.

Luego de 30 años de funcionamiento de los distintos dispositivos de evaluación nacionales y de acuerdo con el plan nacional de evaluación vigente, en primer lugar, se puede afirmar que estas evaluaciones a gran escala nunca fueron concebidas –ni está previsto hacerlo– con un propósito principal de tener implicancias altas, como por ejemplo acreditar conocimientos de un nivel educativo o establecer mecanismos de competencia de mercado dentro del sistema educativo.

Una segunda conclusión es que estas evaluaciones en Argentina son concebidas con un propósito general de carácter diagnóstico, para mejorar los procesos educativos y sus resultados, aunque no se explicita cómo se espera o proyecta que sucedan estas mejoras.

En tercer lugar, la sucesión de evaluaciones nacionales implementadas en Argentina desde 1993 da cuenta de una trayectoria de consolidación en cuanto a su permanencia. Sin embargo, de manera consistente con la falta de definiciones acerca de cómo se espera que ocurran las mejoras en el sistema educativo, se registran vaivenes en lo que hace a las características de su diseño, que parecieran estar ligados a la discusión acerca de las funciones de la evaluación y la noción de calidad de la educación.

Las controversias en Argentina pasan por la construcción de un modelo de evaluación que asuma una noción “amplia” o una noción “estrecha” de la calidad de la educación, según los períodos. Como se indica en una investigación reciente (IIPE-UNESCO, 2021b), “la no resolución de esta controversia a nivel institucional posterga la definición del empleo de los resultados en la macropolítica”, situación que, de cualquier modo, no impide ciertos usos específicos en distintos niveles del sistema y en las provincias, aunque estos sean con bastante falta de previsibilidad y sostenibilidad en el tiempo.

## Usuarios de los resultados y tipos de usos previstos

La definición de usuarios específicos de los resultados y tipos de usos previstos durante estos 30 años ha sido un aspecto difuso en la experiencia de Argentina. Probablemente esta sea la razón que explica que, en la actualidad, varios actores reconocen “la necesidad de revisar y precisar qué evaluar y para qué” en el sistema de evaluación argentino (IIPE-UNESCO, 2021a). Esto implica determinar si los resultados de la evaluación deben ser utilizados, por ejemplo, para el diseño de políticas a nivel de sistema y/o a nivel de instituciones educativas individuales; y si puede y debe darse prioridad a algunos de estos niveles en particular. Avanzar en estas definiciones permitiría una mejor organización de las intervenciones y usos factibles de los resultados de las evaluaciones.

En este contexto el Plan Nacional de Evaluación 2023-2024, a diferencia de su predecesor correspondiente a 2020-2021, avanzó en definir un conjunto de productos a generar a partir de las evaluaciones Aprender, que se detallan a continuación:

1. Informe nacional de resultados: análisis sobre los logros de aprendizaje y sus condiciones.
2. Informes jurisdiccionales de resultados.
3. Informes temáticos nacionales.
4. Recomendaciones metodológicas para la enseñanza.
5. Reportes por escuela: mirar la escuela desde los datos y mirar la escuela desde los aprendizajes.

La definición de este conjunto de productos es un aspecto positivo porque se busca alcanzar distintas audiencias como así también anticipar de qué manera los resultados serán difundidos. Sin embargo, estos productos no son necesariamente novedosos: durante los 30 años de implementación de las evaluaciones a gran escala, ha existido una multiplicidad de productos que fueron variando en el tiempo. En consecuencia, habrá que ver en qué medida se logra la consolidación y continuidad de esta producción conjunta.

Por otra parte, si bien desde esos productos se podría inferir quiénes son algunos de los usuarios potenciales con cierto nivel de precisión, estos no se explicitan y tampoco hay especificaciones sobre los tipos de usos previstos para cada uno de ellos. Por ejemplo, ¿se espera que las recomendaciones pedagógicas para la enseñanza sean tomadas por las y los docentes voluntariamente en su quehacer cotidiano, o hay un plan para incorporarlas en los procesos de formación inicial y continua de estos?

Por último, debería examinarse si las características de los productos generados permiten una apropiación adecuada de los resultados por parte de los usuarios previstos en cada caso. Este es un análisis difícil de realizar en el marco de la ausencia de las definiciones

requeridas respecto de los usuarios. Testimonio de ello es el último informe del Ministerio de Educación de la Nación que fue enviado a las escuelas, “Mirar a la escuela desde los datos. Reporte de datos e indicadores educativos”, generado a partir de Aprender 2019. Allí se observa que la información que les brinda sobre el desempeño en las pruebas de sus estudiantes es muy escasa: únicamente la distribución de las y los estudiantes por niveles de desempeño<sup>2</sup> e información referida a algunas variables contextuales y actitudinales. De esta manera, las posibilidades de interpretación de esta información resultan muy limitadas en ausencia, por ejemplo, de los resultados de otras escuelas que puedan considerarse como grupos de referencia o comparación (por ejemplo, resultados de escuelas de la provincia, públicas y privadas, escuelas en el mismo distrito, escuelas con familias con de grupo socioeconómico similar, etc.)<sup>3</sup>.

## Cobertura de la población objetivo: muestra o censo

Una definición importante cuando se planifica el diseño de una evaluación de aprendizajes a gran escala es determinar si tendrá una cobertura censal o muestral de la población objetivo. Es decir, si se evaluará a la totalidad de estudiantes de la población definida o a un subconjunto que sea representativo.

Si se examina la serie de las operaciones evaluativas desde 1993 en términos de cobertura (**Tabla 3**), se observa un esquema irregular, donde las pruebas censales se desarrollaron desde 1997 hasta 1999 en el final del secundario, con la incorporación del 6° grado de primaria con esta cobertura universal en 2000, para luego ser irregulares entre niveles y años. Desde el inicio de las evaluaciones a gran escala en Argentina, se realizaron cuatro evaluaciones censales de 6° grado de primaria y nueve en 5°/6° de secundaria, sobre Matemática y Lengua.

A su vez, los planes de evaluación educativa 2020-2021 y 2023-2024, aprobados en el Consejo Federal de Educación en julio de 2021 y octubre de 2022, respectivamente, ya definieron qué coberturas tendrán las evaluaciones a gran escala implementadas en los años cubiertos: el esquema propuesto consiste en evaluaciones censales bienales para el final de la primaria y el final de la secundaria, donde estos operativos censales no se superponen por nivel. Así, las evaluaciones censales del 6° grado de primaria quedan para 2021 y 2023, mientras que las del último año de la secundaria (5°/6°) se realizan en 2022 y 2024.

Más allá de los esfuerzos de los planes nacionales en la definición de los tipos de cobertura de las evaluaciones por venir, no se encuentra una justificación en esta toma de decisiones. Este problema está asociado a una situación ya mencionada, que consiste en la falta de acuerdos sobre qué usos se quieren priorizar para las evaluaciones en Argentina y en base a qué propósitos. Más concretamente, cuando se observan los objetivos del Plan Nacional de Evaluación Educativa 2023-2024, de ninguno de ellos se infiere la necesidad de contar con evaluaciones a gran escala censales, dado que no hay una mención específica en ninguno de ellos al uso de la información a nivel de la institución educativa.

## Grados o años a evaluar

La administración de las evaluaciones nacionales a gran escala se concentró, hasta 2016 inclusive, en los grados 3° y 6° del nivel primario y en los años 2°/3° y 5°/6° del secundario (**Tabla 3**). Luego, a partir de 2017, sólo se han administrado aquellas evaluaciones más cercanas al final de cada uno de los niveles educativos.

La racionalidad de medir el logro educativo al final de un nivel no está necesariamente fundamentada en algún instrumento legal o plan de evaluación, pero se asume que se hace de ese modo para conocer la acumulación de conocimientos en un momento cercano

2 En Aprender se consideran 4 niveles de desempeño: por debajo del nivel básico, básico, satisfactorio y avanzado.

3 La falencia en los reportes por escuela indicada no se dio en los informes que se elaboraron a partir de Aprender 2017, los cuales sí brindaban información mucho más detallada. Esto muestra, una vez más, la falta de sostenibilidad de ciertas definiciones referidas a las características de los productos que los usuarios de la información requieren.

al final del año escolar de los y las estudiantes que transitan el último año de ese nivel. Sin embargo, esta decisión puede resultar en información que simplemente está llegando demasiado tarde a los hacedores de política si, por ejemplo, el objetivo fuera intervenir tempranamente para garantizar la finalización de un nivel educativo alcanzando aprendizajes satisfactorios.

Un ejemplo claro de esta situación lo constituyen las evaluaciones censales que se implementan en el último grado del nivel secundario. En la coyuntura actual, luego de la pandemia, han adquirido un mayor peso las políticas para fortalecer trayectorias educativas débiles, con el objetivo de evitar el abandono escolar y apuntalar la incorporación de aprendizajes relevantes. En ese contexto, resultaría de mucha mayor utilidad contar con evaluaciones diagnósticas que permitan identificar el rendimiento de los y las estudiantes y sus características al inicio de la educación secundaria, para así apuntalar a las instituciones que pudieran tener más dificultades, que analizar el rendimiento de estudiantes sólo al final de su nivel educativo, cuando ya una gran parte de los y las estudiantes con dificultades han sido excluidos/as del sistema.

### Dominios a evaluar

Al analizar los contenidos que se evaluaron, se observa que las pruebas a gran escala siempre hicieron foco en los conocimientos de las áreas conocidas como competencias fundamentales, es decir, Matemática y Lengua. El 6° grado de la primaria fue evaluado 18 veces desde 1993 en esos dominios, mientras que los y las estudiantes de 5°/6° año de secundario fueron evaluados/as 16 veces en el mismo período. Además, las evaluaciones también se realizaron sobre Ciencias Naturales y Ciencias Sociales, aunque con una menor frecuencia (13 y 10 veces en 6° de primaria y 5°/6° de secundaria, respectivamente (**Tabla 3**).

El marco de referencia de estas evaluaciones está dado por un conjunto de contenidos y capacidades cognitivas específicas a las áreas que se examinan, el cual se determina considerando lo que se conoce como Núcleos de Aprendizaje Prioritarios (NAP)<sup>4</sup>, los diseños curriculares jurisdiccionales y los consensos jurisdiccionales (Ministerio de Educación, 2016).

En la actualidad, la priorización de la enseñanza de los contenidos fundamentales es una de las intervenciones recomendadas para enfrentar la crisis de aprendizajes existente (Banco Mundial et al, 2022). Sin embargo, esto no necesariamente implica renunciar a evaluar otras áreas, que podrían ser estudiadas con menor frecuencia, y así proveer información muy relevante para mejorar los procesos de enseñanza y aprendizaje.

La incorporación de otros dominios a evaluar, como podrían ser otras áreas de conocimiento o incluso otras habilidades –por ejemplo, habilidades socioemocionales o habilidades digitales– requiere necesariamente contar con una planificación que contemple un horizonte temporal más extenso, por ejemplo de 10 años. Esto permitiría planificar en qué momento evaluar cada asignatura, dándole regularidad al proceso y evitando la superposición de múltiples evaluaciones en el mismo año que, de producirse, genera complicaciones logísticas en la administración de las pruebas, mayores costos y una sobrecarga evaluativa sobre las y los estudiantes.

### Frecuencia de las evaluaciones

La frecuencia con que se implementaron las evaluaciones nacionales en Argentina ha sido otra de las características variables a lo largo de los años y con poca previsibilidad en el largo plazo. Los planes de evaluación acordados en 2021 y 2022 sólo establecen acuerdos por dos años y no especifican la periodicidad esperada de las múltiples iniciativas detalladas.

Una primera advertencia es que no por mucho evaluar se producirán mejoras en el sistema educativo. Para que esto suceda la política de evaluación debe estar alineada, como venimos

4 Los Núcleos de Aprendizaje Prioritarios (NAP) son un conjunto de contenidos y habilidades considerados esenciales en el currículo educativo de Argentina. Fueron establecidos por el Ministerio de Educación como una herramienta para orientar y organizar la enseñanza en todos los niveles educativos.

sosteniendo, a un conjunto de elementos que posibiliten la mejora. Al mismo tiempo, una frecuencia elevada tiene costos presupuestarios considerables dado que gran parte de estos se explican por los recursos que se deben invertir para que el proceso de campo de evaluar cientos de instituciones educativas pueda realizarse.

Por lo tanto, cabe preguntarse sobre el objetivo de recoger todos los años la misma información, que muchas veces sólo lleva a replicar resultados del año anterior, y luego no puede ser analizada en profundidad ni tampoco utilizada intensivamente. Como se destaca en distintas investigaciones, “para que un sistema de evaluación tenga impacto, debe tener una periodicidad que garantice que los datos puedan ser analizados, discutidos, comprendidos y utilizados. Incorporar la nueva información y traducirla en decisiones y acciones lleva tiempo. Los cambios en el sistema educativo requieren más tiempo aún” (Ravela et al, 2008).

Desde esta perspectiva, la frecuencia de las evaluaciones debe organizarse de modo que permita ir evaluando distintos niveles, contenidos y competencias, dando tiempo para actuar sobre los hallazgos. Adicionalmente, al espaciar la frecuencia, también se posibilita mayor tiempo para el diseño metodológico de las evaluaciones de modo de robustecer la confiabilidad de sus resultados, tanto para el año específico en que se generan como también, muy especialmente, para las comparaciones de resultados entre años.

## Recomendaciones para una política nacional de evaluación de aprendizajes a gran escala

Como se puede observar, desde 1993 hasta 2022 se ha realizado una gran cantidad de evaluaciones nacionales de aprendizaje a gran escala, tanto en el nivel primario como en el nivel secundario. La mayoría se enfocó en las asignaturas de Lengua y Matemática, áreas que también fueron cubiertas tanto por las evaluaciones internacionales en las que Argentina participó regularmente como también por evaluaciones en algunas jurisdicciones que desarrollaron sus propios dispositivos.

De este escenario general se desprende un primer punto a resaltar, que es la permanencia de todas estas evaluaciones a lo largo del tiempo, lo que evidencia una trayectoria de consolidación y pareciera reflejar un consenso sobre la importancia de su implementación. También supone la generación de una gran cantidad de información sobre el sistema educativo, pero que, en general, ha sido subutilizada como herramienta para la mejora de su calidad (IIPE-UNESCO, 2021b).

En segundo lugar, las características de este tipo de evaluaciones nacionales, como son su frecuencia, asignaturas, años y grados evaluados y el tipo de cobertura, han ido variando a lo largo del tiempo. Estas modificaciones responden a las decisiones de las distintas gestiones de gobierno, acordadas en mayor o menor medida con las jurisdicciones. Sin embargo, tal como se mencionó, es poco lo que se sabe sobre los fundamentos de estos cambios, al menos públicamente<sup>5</sup>.

Las modificaciones recurrentes y la falta de previsibilidad sobre las características de las evaluaciones a gran escala en Argentina están relacionadas con la ausencia de una definición precisa sobre el “para qué” de las evaluaciones nacionales. Tener claridad sobre las funciones y usos de las evaluaciones es determinante a la hora de tomar decisiones relativas a la frecuencia, los contenidos y habilidades, años y grados a evaluar y la utilización de censos o muestras.

A partir de este diagnóstico general, se presentan a continuación cinco recomendaciones

5

A modo de ejemplo, cuando se determinó que el ONE asumiría una frecuencia trienal, la resolución que así lo determinaba no ofreció ninguna justificación para respaldar la decisión. Tampoco lo hizo la que dio origen a la evaluación Aprender a la hora de retomar primero la frecuencia anual, para luego pasar a una bienal. Lo mismo ocurre con los cambios en otras características, como las asignaturas, los años y grados evaluados, y el nivel de cobertura de las evaluaciones.

que podrían considerarse en el diseño de una política nacional de evaluación de aprendizajes a gran escala que pretenda ser una herramienta eficiente al servicio de la mejora de la calidad educativa. Estas recomendaciones se estructuran bajo el principio general de la implementación de evaluaciones de bajo impacto, diseñadas con el propósito principal de constituir diagnósticos para la mejora de los aprendizajes, que proveen información relevante a distintos niveles del sistema educativo.

### 1. Implementar evaluaciones censales quinquenales para elaborar un mapa nacional de aprendizajes y trayectorias educativas a nivel de escuela.

Existe acuerdo en la actualidad sobre la importancia de generar políticas para garantizar trayectorias educativas completas de los y las estudiantes de modo de evitar el abandono escolar. Un buen desempeño escolar es una variable clave para fortalecer el vínculo del estudiante con sus pares y la institución educativa.

En este contexto, se recomienda implementar evaluaciones educativas censales con un primer objetivo de elaborar un mapa nacional de caracterización de los aprendizajes y las trayectorias educativas, que integre los datos de las evaluaciones con los de progresión y graduación que provienen de otras fuentes de información<sup>6</sup>, que sirva como base para las políticas nacionales y jurisdiccionales de reforzamiento de trayectorias educativas. Los datos censales deberán servir para el desarrollo de productos a nivel de escuela, donde se informe sobre el desempeño de los y las estudiantes en comparación con el de otras instituciones educativas similares. Estos informes deberían ser la base para la elaboración de proyectos de mejoramiento institucional.

La frecuencia sugerida para estos censos es quinquenal, dado que obtener información una vez cada 5 años sobre el desempeño educativo de la totalidad de estudiantes de un curso determinado de todas las escuelas del país debería ser suficiente para caracterizar a las instituciones sobre los aspectos cubiertos por las pruebas.

Con respecto a la definición sobre en qué grados y materias realizarlos, se propone que el censo de evaluación educativa en el nivel primario se realice evaluando sólo competencias de Lectura en el 3° grado, momento en el que está previsto que los y las estudiantes hayan completado su proceso de alfabetización básica. A su vez, en el nivel secundario, este censo debería realizarse en el 1° año, sobre Matemática y Lengua, de manera de poder estudiar las características de la totalidad de estudiantes que transitaron de la primaria a la secundaria, evitando perder la información de aquellos/as que abandonan sus estudios antes de llegar al último año del nivel.

Finalmente, estos censos deberían considerar también el relevamiento de información del contexto de cada estudiante y de la escuela, que servirá para caracterizar la situación socioeconómica en donde se desenvuelve el proceso educativo.

### 2. Continuar las evaluaciones de diagnóstico sobre las áreas fundamentales de Matemática y Lengua, con operativos muestrales bienales, en los últimos años de los niveles primario y secundario.

El monitoreo de los logros de aprendizaje que posibilitaron las evaluaciones de diagnóstico sobre las áreas fundamentales de Matemática y Lengua, implementadas desde 1993, con una regularidad considerable y cobertura variable, debe persistir de modo de continuar con las series históricas que se generaron a partir de Aprender. Sin embargo, el mismo puede hacerse de una manera mucho más eficiente si se realiza con un esquema de evaluaciones muestrales y frecuencia bienal.

Los operativos muestrales simplifican mucho la logística de implementación en comparación a las evaluaciones censales, ya que posibilitan una administración de mayor calidad dada por diversos factores, como son un mayor resguardo de la confidencialidad de las pruebas, simplificación de los procesos de supervisión, carga de datos, oportunidad de la información, entre otros.

6

Los datos de progresión y graduación (por ejemplo, sobreedad, repitencia, egreso) se generan en Argentina a partir de la operación estadística que se conoce como Relevamiento Anual, que se implementa liderado por el Ministerio de Educación de la Nación en las 24 jurisdicciones del país.

En contraposición, un ejemplo claro de las dificultades que enfrentan las evaluaciones censales para sostener algunos parámetros de calidad de la administración se observa en la cobertura que estos operativos tienen a nivel de estudiante: en Aprender 2022 del nivel secundario, se registró que, en cuatro provincias, más del 40% de las y los estudiantes no participaron de las pruebas (Ministerio de Educación, 2023b).

De cualquier modo, una evaluación muestral tampoco significa evaluar “menos”. En la práctica, existen formas de diseñar pruebas donde se evalúa una gran cantidad de temas: se establecen distintos tipos de pruebas, donde no todos/as los y las estudiantes deben responder a las mismas consignas, lo que permite alcanzar un mayor espectro de los contenidos y competencias que se evalúan. Este tipo de mecanismos resulta mucho más sencillo de llevar a la práctica con evaluaciones muestrales que censales.

### 3. Incorporar el estudio de las habilidades socioemocionales en ambos niveles educativos.

Desde 1993, cuando se introdujeron en el país las evaluaciones gran escala, estas se concentraron mayormente en Lengua y Matemática en los grados o años finales de la primaria y secundaria, aunque en reiteradas oportunidades se evaluaron Ciencias Sociales y Ciencias Naturales. Sin embargo, existe dentro del sistema educativo un objetivo de avanzar en una evaluación que considere un concepto amplio de calidad educativa, tal como se manifiesta en los últimos planes nacionales de evaluación, de ir hacia una mayor integralidad del proceso educativo. En esta línea, se recomienda iniciar un proceso de evaluaciones de habilidades socioemocionales.

La ampliación de la evaluación al campo de las habilidades socioemocionales (también conocidas como transferibles, habilidades para la vida, o para el siglo XXI) se sustenta en la gran importancia que ellas adquieren en su función de apoyo para desarrollarse en la escuela, en la vida y en el trabajo. Las habilidades socioemocionales, en conexión con las habilidades fundamentales, proveen la base para la incorporación de otras competencias, las cuales “se relacionan con las destrezas necesarias para poder adaptarse a los diferentes contextos de la vida y que, potencialmente, los individuos pueden transferir a distintos ámbitos laborales o sociales” (UNICEF, 2020). Incluyen habilidades cognitivas, como por ejemplo resolución de problemas y concentración; emocionales, como la empatía y la resiliencia; y sociales, como comunicación y colaboración.

El estudio de las habilidades socioemocionales es un dominio relativamente nuevo en los sistemas educativos. En Argentina existió un antecedente de una evaluación de esas habilidades, implementado en la Ciudad Autónoma de Buenos Aires (UEICEE, 2017), mientras que, a nivel nacional, el único antecedente es el de la cuarta versión del Estudio Regional Comparativo y Explicativo (ERCE 2019) de la UNESCO. Esta evaluación incluyó un módulo administrado a una muestra representativa nacional de estudiantes del 6° grado de primaria, que se enfocó en tres habilidades o competencias específicas: empatía, apertura a la diversidad y la autorregulación escolar (UNESCO, 2021).

Operativamente, se debería considerar la posibilidad de alternar esta evaluación específica con los cuestionarios complementarios que tradicionalmente se administraron con las pruebas de conocimiento. Estos cuestionarios complementarios recogen información sobre las características de los y las estudiantes, las escuelas y la experiencia educativa que se da en ellas, con el objetivo principal de brindar un marco conceptual para interpretar los resultados e identificar aquellos factores que se encuentran asociados a un mejor rendimiento académico. Dada la gran cantidad de veces que estos se han administrado en los últimos años, existe mucha información disponible al respecto. Por lo tanto, un esquema rotatorio entre estos cuestionarios de contexto e instrumentos de evaluación para habilidades socioemocionales podría ser un esquema eficiente para aumentar el alcance de las pruebas sin sobrecargar a los y las estudiantes con más cantidad de dispositivos de recolección de información.

Considerando las tres recomendaciones formuladas anteriormente, la **Tabla 7** resume la estructura de una nueva política nacional de evaluación a gran escala que propone, en resumen:

- Evaluaciones educativas censales cada 5 años, evaluando Lengua en la mitad del ciclo de la primaria y Lengua y Matemática en el inicio del nivel secundario, que sirvan como insumo para elaborar mapas nacionales de aprendizajes y trayectorias educativas.
- Monitoreo bienal de los aprendizajes en el sistema educativo en los dominios fundamentales, Matemática y Lengua, al final de los ciclos de primaria y secundaria.
- Evaluaciones de habilidades socioemocionales cada cuatro años en primaria y secundaria, que se alternan con la recopilación de información de contexto de las y los estudiantes de los cuestionarios complementarios.

**TABLA 7**  
Propuesta de política nacional de evaluación 2025-2034

Año	Nivel primario						Nivel secundario					
	3° grado			6°/7° grado			1° año			5°/6° año		
2025	L	Co					M	L	Co			
2026				M	L	HSE				M	L	HSE
2027												
2028				M	L	Co				M	L	Co
2029												
2030	L	Co		M	L	HSE	M	L	Co	M	L	HSE
2031												
2032				M	L	Co				M	L	Co
2033												
2034				M	L	HSE				M	L	HSE

Nota: el sombreado indica aplicación censal sobre la población objetivo. Las letras en el cuadro identifican a la asignatura (M=Matemática; L=Lengua y HSE=Habilidades socioemocionales) y, en el caso de Co, la aplicación de cuestionarios complementarios que recogen información de contexto de los y las estudiantes y la escuela.

#### 4. Desarrollar esquemas para aumentar la coordinación de las evaluaciones entre Nación y las jurisdicciones y fortalecer así sus capacidades de evaluación.

En la organización del sistema educativo en Argentina, las jurisdicciones son las destinatarias principales de los resultados de evaluación pues son ellas las primeras responsables de la gestión del sistema. Una comparación entre las evaluaciones jurisdiccionales recientes (**Tabla 5**) y las evaluaciones nacionales implementadas en 2021 y 2022 (**Tabla 3**) permite identificar una superposición de grados y materias. Si bien esto podría responder a que algunas jurisdicciones estén buscando evaluar contenidos que les son propios y no tan bien reflejados en los que son cubiertos por las pruebas nacionales Aprender, es altamente probable que el solapamiento responda más a cuestiones políticas y falta de coordinación entre niveles de gobierno que a necesidades de información específicas de las gestiones educativas jurisdiccionales. Por esta razón, vincular la política nacional y las jurisdiccionales de evaluación resulta clave, a fin de favorecer la complementariedad y evitar superposiciones.

Un primer paso en esta dirección consiste en acordar una estrategia entre la Nación y las jurisdicciones donde se plasmen las especificidades que adquirirán los dispositivos de evaluación a gran escala nacionales. Partiendo de esta base, en primer lugar, las jurisdicciones conocerán con qué información contarán para apoyar la gestión educativa, y en segundo lugar, podrán determinar la necesidad de avanzar en el diseño de dispositivos de evaluación propios que realmente complementen lo realizado por el nivel nacional. De lo contrario, si no existe una planificación a largo plazo (por ejemplo, de 10 años) de las evaluaciones nacionales, resulta muy difícil pretender que las provincias avancen en evaluaciones que complementen a las nacionales.

La **Tabla 7** presenta una propuesta de política nacional de evaluación a 10 años, entre 2025 y 2034, y en ella se puede vislumbrar áreas donde la mirada jurisdiccional puede complementar a las evaluaciones nacionales. Por ejemplo, la evaluación de las Ciencias Sociales y Naturales sería una de ellas, como así también evaluaciones que se realicen al final del ciclo básico del secundario. Luego, la política nacional propuesta también ofrece años sin implementación de pruebas nacionales (por ejemplo, 2027, 2029, etc.), dando espacio para que las jurisdicciones implementen sus propias pruebas.

Al mismo tiempo, esta estrategia de coordinación debe contemplar recursos financieros adicionales, desde el Ministerio de Educación de la Nación, para apoyar el desarrollo de capacidades de evaluación y de uso de la información en las jurisdicciones que lo requieran, para que puedan efectivamente llevar a la práctica esta función de complementación.

#### 5. Promover procesos de evaluación formativa en el aula.

En el contexto de la salida gradual de la pandemia, en América Latina se dio un interés renovado por la evaluación formativa como proceso para mejorar la incorporación de aprendizajes, donde países como Colombia, Guatemala, Honduras, Panamá y Uruguay ya desarrollan nuevas acciones o profundizan líneas de acción existentes en esta área (Perusia, 2021).

La evaluación formativa refiere a las actividades formales e informales de evaluación que conducen los y las docentes durante los procesos de enseñanza y aprendizaje para modificar las estrategias pedagógicas con el objetivo de mejorar el logro de las y los estudiantes. A diferencia de la evaluación sumativa, que se concentra en el final de una etapa educativa, la evaluación formativa se realiza durante toda ella y se enfoca en el proceso mismo, con iteraciones continuas entre el/la estudiante y el/la docente (Bennett, 2011; Black y Wiliam, 2009; Allal, 1980; citado en UNESCO, 2021). De esta manera, pone a disposición información personalizada y en profundidad tanto a estudiantes como a docentes, a fin de adaptar sus estrategias de aprendizaje y enseñanza.

La incorporación de procesos de evaluación formativa como parte de una política nacional de evaluación de aprendizajes es recomendable y constituye un complemento apropiado a las pruebas de gran escala, ya que es una herramienta pedagógica pensada para el trabajo del docente a nivel individual con el o la estudiante. Dado que desde el Ministerio de Educación Nacional se impulsó este instrumento en el Plan 2023-2024, es esencial conocer cómo está siendo concebido en la actualidad, al igual que otras experiencias en la región, para así determinar la mejor manera de articularlo con las evaluaciones a gran escala.

## Reflexiones finales

A 30 años del inicio de las evaluaciones de aprendizajes a gran escala en la Argentina, su persistencia en el tiempo y su sostenimiento por parte de los distintos gobiernos dan cuenta de una política consolidada en cuanto a su presencia. Los planes nacionales de evaluación educativa diseñados para 2020-2021 y 2023-2024, impulsados desde el Ministerio de Educación de la Nación y acordados en el Consejo Federal de Educación, representaron un primer avance en términos de planificar y dar cierto marco de previsibilidad a la implementación de distintos instrumentos de evaluación.

Sin embargo, se debe avanzar mucho más para que estos planes se transformen en instrumentos eficientes para el diagnóstico y la mejora de los aprendizajes: es necesario romper con la inercia evaluativa que llevó a la acumulación de una gran cantidad de dispositivos donde no hay claridad respecto de sus objetivos ni de cómo se articulan entre ellos.

Esta presencia consolidada de las evaluaciones no se condice con una utilización intensiva de sus resultados. Uno de los requisitos para superar esta situación consiste en definir con mayor precisión los propósitos principales de las pruebas, identificando usuarios de la información y tipos de usos que se espera que estos den para apropiarse de los resultados. Una vez definidos estos propósitos, usuarios y usos, se deberían alinear las características técnicas de las pruebas en, al menos, lo que refiere a su nivel de cobertura, población objetivo, áreas de conocimiento a evaluar y periodicidad.

La implementación de las recomendaciones provistas en este documento requiere definir un horizonte temporal de planificación más extenso que los dos años propuestos por los planes nacionales de evaluación vigentes. La planificación a largo plazo, escenario de la propuesta 2025–2034 que este documento utiliza como ejemplo, es clave no sólo para diseñar un esquema donde vayan rotando los distintos instrumentos nacionales de evaluación a gran escala de acuerdo con los objetivos definidos, sino también para poder definir las características de las evaluaciones jurisdiccionales de modo tal que resulten complementarias y evitar la duplicación de esfuerzos y recursos. Dicho de otro modo, si no sabemos qué evaluaciones nacionales se implementarán de aquí a dos o tres años, ¿cómo podrían las jurisdicciones desarrollar una evaluación que complemente el accionar nacional?

Un segundo tema sobre el cual reflexionar es qué tipo de institucionalidad permitiría implementar de la mejor manera una política nacional de evaluación a gran escala como la que se propone en este documento. En la actualidad, el ámbito institucional de acuerdo es el Consejo Federal de Educación, donde la conducción de las evaluaciones nacionales está a cargo del Ministerio de Educación de la Nación, con el asesoramiento del Consejo Nacional de Calidad de la Educación<sup>7</sup>.

Existen varias experiencias en el mundo donde la responsabilidad de conducir el diseño e implementación de las evaluaciones a gran escala se asigna a una organización gubernamental que, si bien tiene una vinculación con el organismo que conduce la política educativa, también posee una mayor independencia dada por algunas características de su diseño institucional, por ejemplo, cómo se elige a la autoridad que la dirige, la asignación de autonomía presupuestaria, y/o personería jurídica propia. En América Latina, estos diseños institucionales se verifican en el INEP en Brasil, la Agencia de Calidad de la Educación en Chile, ICFES en Colombia, INEVAL en Ecuador, el IDEICE de República Dominicana, y el INEEd de Uruguay<sup>8</sup>.

Si bien cada una de las instituciones mencionadas en el párrafo anterior tiene sus particularidades, el punto en común que justifica su creación ha sido la búsqueda de una institucionalidad más independiente que permita generar las condiciones necesarias para

7 El Consejo Nacional de Calidad de la Educación es un organismo creado en el año 2006 por la Ley de Educación Nacional, definido como “un órgano de asesoramiento especializado, que estará integrado por miembros de la comunidad académica y científica de reconocida trayectoria en la materia, representantes del Ministerio de Educación de la Nación, del Consejo Federal de Educación, del Congreso Nacional, de las organizaciones del trabajo y la producción, y de las organizaciones gremiales docentes con personería nacional”. Sin embargo, su puesta en marcha fue recién reglamentada en 2020.

8 El INEP es el “Instituto Nacional de Estudios y Pesquisas Educativas”, el ICFES es el Instituto Colombiano para la Evaluación de la Educación, el INEVAL es el Instituto Nacional de Evaluación Educativa, el IDEICE es el Instituto Dominicano de Evaluación e Investigación de la Calidad Educativa, y el INEEd es el Instituto Nacional de Evaluación Educativa.

una planificación y continuidad de las políticas de evaluación a más largo plazo, intentando que las mismas se conviertan en una política de Estado.

En Argentina, a nivel nacional, existieron dos experiencias que buscaron avanzar en el diseño de una institucionalidad similar a la de los países reseñados: la primera fue la creación en 2001 del Instituto para el Desarrollo de la Calidad Educativa (IDECE, según Decreto Nacional 165/01, que fue disuelto en 2003), y la segunda fue un proyecto de ley –nunca aprobado– que se presentó en 2017 y que proponía la creación de un ente descentralizado del Ministerio de Educación, “con autonomía funcional y potestad para dictar su organización interna, autarquía económica financiera, personería jurídica propia” (Llach, 2020). En este contexto, vale la pena reflexionar sobre si la adopción en Argentina de un modelo institucional de estas características para la evaluación educativa a gran escala en Argentina debería ser parte de la solución que aborde los desafíos identificados para la actual política de evaluación.

Un tercer tema está dado por la importancia de continuar participando de evaluaciones internacionales, tales como los estudios de la UNESCO y de OCDE. Este tipo de estudios son los únicos capaces de comparar la evolución de los aprendizajes de los y las estudiantes de Argentina con los y las de otros países. A esto se suma el hecho de que la participación en ellos tiene el beneficio de conectar a los equipos técnicos del país con sus pares de otros países del mundo y así fortalecer las capacidades técnicas nacionales de evaluación.

Por último, considerando la extensión de la política propuesta, es deseable establecer un mecanismo de monitoreo y evaluación. Este mecanismo, luego de los primeros 4 años de implementación, cuando se hayan completado los censos en los dos niveles y las evaluaciones muestrales cubriendo todos los dominios propuestos, brindaría elementos para realizar los ajustes que permitan mejorar los resultados alcanzados.

## Bibliografía

- Banco Mundial. (2018). *World Development Report 2018: Learning to Realize Education's Promise*. Washington, DC: World Bank. Obtenido de: <https://www.worldbank.org/en/publication/wdr2018>
- Banco Mundial, UNESCO y UNICEF. (2021). *The State of the Global Education Crisis: A Path to Recovery*. Washington D.C., París, Nueva York: The World Bank, UNESCO, y UNICEF. Obtenido de: <https://www.unicef.org/reports/state-global-education-crisis>
- Banco Mundial. (2022). *La crisis de aprendizaje en América Latina y el Caribe y la pandemia del COVID-19: Resultados para reflexionar sobre una tendencia que se profundiza*. Obtenido de: <https://documents1.worldbank.org/curated/en/099443504042242867/pdf/IDU0e7871c270811504df009e620db34dfd8158b.pdf>
- Banco Mundial, Bill & Melinda Gates Foundation, FCDO, UNESCO, UNICEF, y USAID. (2022). *Guía para la recuperación y aceleración de los aprendizajes: Usando el marco RÁPIDo para abordar las pérdidas de aprendizajes por COVID-19 y construir mejor en el futuro*. Washington, DC: Banco Mundial. Obtenido de: <https://thedocs.worldbank.org/en/doc/e52f55322528903b27f1b7e61238e416-0200022022/related/Guide-for-Learning-Recovery-and-Acceleration-06-23.pdf>
- CFE. (septiembre de 2010). Resolución N° 116/2010. *Lineamientos para la evaluación de componentes de la calidad educativa 2010 – 2020*. Obtenido de <https://cfe.educacion.gob.ar/resoluciones/>
- CFE. (mayo de 2016). Resolución N° 280/2016. *Aprobación de la implementación del “Sistema de Evaluación Nacional de la Calidad y Equidad Educativa”*. Obtenido de <https://cfe.educacion.gob.ar/resoluciones/>
- CFE. (julio de 2021). Resolución N° 396/2021. *Aprobación del Plan de Evaluación Educativa 2020-2021*. Obtenido de [https://www.argentina.gob.ar/sites/default/files/rcfe\\_396-21\\_y\\_anexo.pdf](https://www.argentina.gob.ar/sites/default/files/rcfe_396-21_y_anexo.pdf)
- CFE. (octubre de 2022). Resolución N° 435/2022. *Plan Nacional de Evaluación Educativa 2023-2024*. Obtenido de [http://www.bnm.me.gov.ar/giga1/normas/RCFE\\_435-22.pdf](http://www.bnm.me.gov.ar/giga1/normas/RCFE_435-22.pdf)
- Clarke, M., y Luna-Bazaldua, D. (2021). *Primer on Large-Scale Assessments of Educational Achievement*. Washington, DC: Banco Mundial. Obtenido de: <https://openknowledge.worldbank.org/entities/publication/618b21d1-e54e-5a94-a088-b7ed8965a990>
- Engzell, P., Frey, A., & Verhagen, M. (27 de abril de 2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceeding of the National Academy of Sciences (PNAS)*.
- Greaney, V., y Kellaghan, T. (2016). *Evaluaciones nacionales del rendimiento académico. Volumen 1: Evaluación de los niveles nacionales de rendimiento académico*. Washington, DC: Banco Mundial. Obtenido de: <https://openknowledge.worldbank.org/server/api/core/bitstreams/cd5c01cb-5d53-5cbd-880a-4d17d8735561/content>
- IPEE-UNESCO, Oficina para América Latina. (2021a). *El uso de los resultados de las pruebas de evaluación de los aprendizajes en el planeamiento de las políticas educativas en Argentina: Informe nacional*. Buenos Aires. Obtenido de: <https://unesdoc.unesco.org/ark:/48223/pf0000379598>
- IPEE-UNESCO, Oficina para América Latina. (2021b). *El uso de los resultados de las pruebas de evaluación de los aprendizajes en el planeamiento de las políticas educativas en seis países de la región*. Informe regional. Buenos Aires. Obtenido de: <https://unesdoc.unesco.org/ark:/48223/pf0000379593>
- Larripa, S. (2010). *Sistemas jurisdiccionales de evaluación educativa en la Argentina: los casos de la Ciudad de Buenos Aires y la provincia de Buenos Aires (1997-2007)*. [Tesis doctoral, Universidad de San Andrés]. Repositorio institucional de la Universidad de San Andrés: <https://repositorio.udes.edu.ar/jspui/handle/10908/10986>

Ley Federal de Educación. (1993). Ley N° 24.195. Argentina.

Ley Educación Nacional. (2006). Ley N° 26.206. Argentina.

Linchand, G., Dória, C., Leal Neto, O., & Cossi, J. (2021). *Educational Impacts of Remote Learning in Secondary Education: Evidence from Brazil during pandemic*. Research Square. <https://doi.org/10.21203/rs.3.rs-568605/v1>

Llach, J. J. (2020). *¿Cuáles son los organismos responsables de la evaluación educativa en los países de la región? Ciudad Autónoma de Buenos Aires: Observatorio Argentinos por la Educación*. Obtenido en: <https://argentinosporlaeducacion.org/wp-content/uploads/2022/02/organismos-de-evaluacion-educativa.pdf>

Ministerio de Educación de la Nación. (2016). *Aprender 2016: notas metodológicas*. Ciudad Autónoma de Buenos Aires: Ministerio de Educación de la Nación. Obtenido de: [https://www.argentina.gob.ar/sites/default/files/notas\\_metodologicas\\_0.pdf](https://www.argentina.gob.ar/sites/default/files/notas_metodologicas_0.pdf)

Ministerio de Educación de la Nación (2023a). *Dispositivos de evaluación de aprendizajes de estudiantes: experiencias jurisdiccionales*. Ciudad Autónoma de Buenos Aires: Ministerio de Educación de la Nación.

Ministerio de Educación de la Nación (2023b). *Resultados Aprender 2022 en el nivel secundario*. Obtenido de: [https://www.argentina.gob.ar/sites/default/files/2023/06/\\_resultados\\_a\\_nivel\\_nacional\\_aprender\\_2022\\_en\\_nivel\\_secundario\\_-\\_secretaria\\_de\\_informacion\\_y\\_evaluacion\\_educativa.pptx\\_1.pdf](https://www.argentina.gob.ar/sites/default/files/2023/06/_resultados_a_nivel_nacional_aprender_2022_en_nivel_secundario_-_secretaria_de_informacion_y_evaluacion_educativa.pptx_1.pdf)

Nation's Report Card. (2022). *NAEP Long-Term Trend Assessment Results: Reading and Mathematics*. Nationsreportcard.gov. Obtenido de: <https://www.nationsreportcard.gov/highlights/ltr/2022/>

Perusia, J. (julio de 2021). *Evaluaciones educativas, una de las respuestas a la "segunda ola" de la crisis de los aprendizajes provocada por el COVID-19*. Buenos Aires: CIPPEC. Disponible en: [www.cippec.org/textual/evaluaciones-educativas-una-de-las-respuestas-a-la-segunda-ola-de-la-crisis/](http://www.cippec.org/textual/evaluaciones-educativas-una-de-las-respuestas-a-la-segunda-ola-de-la-crisis/)

Perusia, J. y Cardini, A. (septiembre de 2021). *Sistemas de alerta temprana en la educación secundaria: prevenir el abandono escolar en la era del COVID-19*. Documento de Política Pública N°233. Buenos Aires: CIPPEC. Disponible en: <https://www.cippec.org/publicacion/sistemas-de-alerta-temprana-en-el-nivel-secundario/>

Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Martínez, F., Aylwin, M. y Wolff, L. (2008). *Las evaluaciones educativas que América Latina necesita*. En Revista Iberoamericana de Evaluación Educativa. Vol 1. Número 1. Obtenido de: <http://www.rinace.net/riee/numeros/vol1-num1/art4.pdf>

Ravela, P. (2006). *Para comprender las evaluaciones educativas*. Fichas didácticas. Santiago de Chile: Ed. San Marino.

Renaissance Learning, Education Policy Institute. (2021). *Understanding progress in the 2020/21 academic year*. Department for Education.

Srivastava, P., Cardini, A., Matovich, I., Moussy, H., Gagnon, A. A., Jenkins, R., Reuge, N., Moriarty, K., Anderson, S. (2020). *Covid-19 and the global education emergency: Planning systems for recovery and resilience*. SAUDI ARABIA: T20.

UEICEE. (septiembre de 2017). *Las habilidades socioemocionales en el modelo de cinco factores. Un estudio con alumnos participantes de FEPBA 2014*. Ciudad de Buenos Aires. Obtenido de: <https://buenosaires.gob.ar/sites/default/files/media/document/2017/12/22/135ba7a49f409fdcf7450f783726d06f59343941.pdf>

UNESCO. (2013). *La crisis mundial del aprendizaje: Por qué todos los niños merecen una educación de buena calidad*. París. Obtenido de: <https://unesdoc.unesco.org/ark:/48223/pf0000223826>

UNESCO. (2019). *The promise of large-scale learning assessments: acknowledging limits to unlock opportunities*. Paris. Obtenido de: <https://unesdoc.unesco.org/ark:/48223/pf0000369697>

UNESCO (2021). *Habilidades socioemocionales en América Latina y el Caribe. Estudio Regional Comparativo y Explicativo (ERCE 2019)*. Obtenido de: <https://unesdoc.unesco.org/ark:/48223/pf0000380240>

UNESCO. (2021). *Evaluación formativa: Una oportunidad para transformar la educación en tiempos de pandemia; reflexión a partir de los resultados del estudio cualitativo sobre perspectivas docentes en torno a la evaluación formativa*. Santiago de Chile. Obtenido de: <https://unesdoc.unesco.org/ark:/48223/pf0000378045>

UNICEF. (2020). *Importancia del desarrollo de habilidades transferibles en América Latina y el Caribe*. Documento de discusión. Obtenido de: [https://www.unicef.org/lac/sites/unicef.org/lac/files/2020-07/Importancia-Desarrollo-Habilidades-Transferibles-ALC\\_0.pdf](https://www.unicef.org/lac/sites/unicef.org/lac/files/2020-07/Importancia-Desarrollo-Habilidades-Transferibles-ALC_0.pdf)

Verger, A., Parcerisa, L. and Fontdevila, C. (2018). *The growth and spread of large-scale assessments and test-based accountabilities: a political sociology of global education reforms*. *Educational Review*, Vol. 71, No. 1, pp. 5-30.

## Acerca de los autores



Las opiniones expresadas en este documento no reflejan necesariamente la posición institucional de CIPPEC en el tema analizado.

Esteban Torre

— Director de Educación en CIPPEC

Licenciado en Ciencia Política y Gobierno (Universidad Torcuato Di Tella), Especialista en Educación (Universidad de San Andrés) y Magíster en Políticas Públicas y Administración (London School of Economics and Political Science).

Juan Cruz Perusia

— Investigador principal de Educación en CIPPEC

Economista (Universidad Nacional de Córdoba) y Magíster en Economía (Universidad del CEMA).

Juan Xanthopoulos

— Coordinador de Educación en CIPPEC

Licenciado en Ciencia Política (Universidad de San Andrés) y Maestrando en Políticas Educativas (Universidad Torcuato Di Tella).

---

### Proceso de calidad

Todos los Documentos de Política Pública están sujetos a la política de calidad de CIPPEC que contempla la revisión externa de cada publicación. La revisión por parte de especialistas contribuye a garantizar la calidad técnica, la relevancia y la accesibilidad de los documentos. Un agradecimiento especial para Alejandra Cardini (Instituto Internacional de Planeamiento de la Educación de la UNESCO) y Tamara Vinacur (Banco Interamericano de Desarrollo), quienes colaboraron en ese proceso.

### Agradecimientos

Este documento fue elaborado en el marco del Laboratorio de Innovación y Justicia Educativa del Programa de Educación de CIPPEC (EduLab). Agradecemos el apoyo de las organizaciones que nos acompañan en este proyecto en 2023: Zurich, Fundación Itaú, Banco Industrial y Fundación Telefónica-Movistar. Por otra parte, los autores agradecen a Virginia Lozano por sus aportes a lo largo de la elaboración del documento y a Cecilia Beloqui por su atenta lectura y valiosos comentarios.



### Para citar este documento:

Torre, E., Perusia, J.C. y Xanthopoulos, J. (agosto de 2023). *Hacia una política nacional educativa a gran escala al servicio de la mejora de los aprendizajes*. Documento de Política Pública N° 244. Buenos Aires: CIPPEC.

Por medio de sus publicaciones, CIPPEC aspira a enriquecer el debate público en la Argentina con el objetivo de mejorar el diseño, la implementación y el impacto de las políticas públicas, promover el diálogo democrático y fortalecer las instituciones.

Los Documentos de Políticas Públicas de CIPPEC ofrecen un análisis que sintetiza los principales diagnósticos y tomas de posición sobre un problema o una situación que afecta al país.

CIPPEC alienta el uso y divulgación de sus documentos sin fines comerciales. Las publicaciones de CIPPEC son gratuitas y se pueden descargar en [www.cippec.org](http://www.cippec.org)

---

## ¿QUIÉNES SOMOS?

CIPPEC es una organización independiente, apartidaria y sin fines de lucro que produce conocimiento y ofrece recomendaciones para construir mejores políticas públicas.

## ¿QUÉ HACEMOS?

CIPPEC propone, apoya, evalúa y visibiliza políticas para el desarrollo con equidad y crecimiento, que anticipen los dilemas del futuro mediante la investigación aplicada, los diálogos abiertos y el acompañamiento a la gestión pública.

## ¿CÓMO NOS FINANCIAMOS?

CIPPEC promueve la transparencia y la rendición de cuentas en todas las áreas de la función pública y se rige por esos mismos estándares. El financiamiento de CIPPEC está diversificado por sectores: cooperación internacional, empresas, individuos y gobiernos. Los fondos provenientes de gobiernos se mantienen por debajo del 30 por ciento del presupuesto total.

