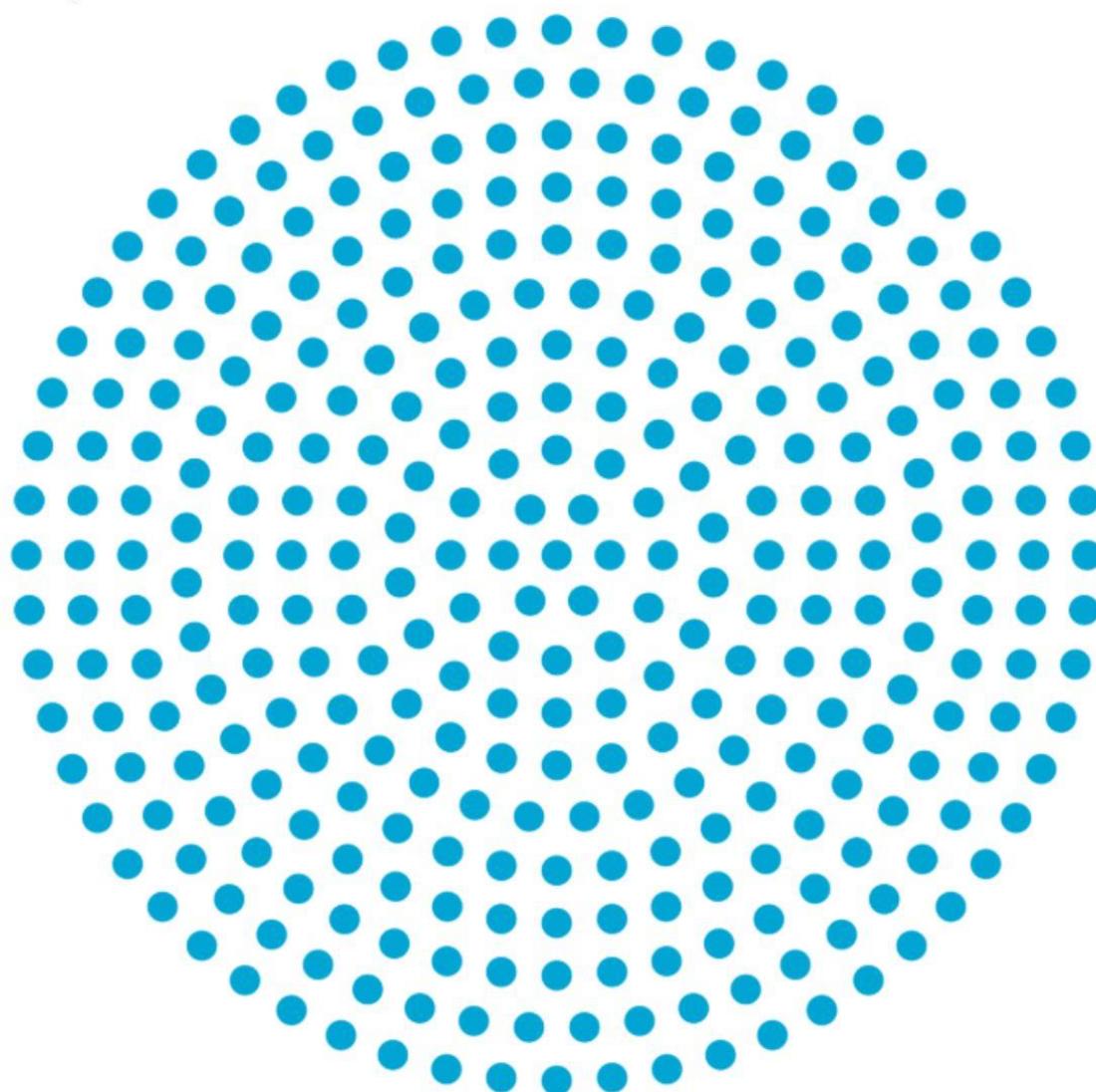


PROGRAMA DE EDUCACIÓN

El diseño de sistemas provinciales efectivos de evaluación de aprendizajes en Argentina

Juan Cruz Perusia
Juan Xanthopoulos



CIPPEC®



Índice

Resumen	3
Introducción	5
1. Las evaluaciones de aprendizaje a gran escala en los países federales	7
2. De un conjunto de evaluaciones educativas a un sistema de evaluación efectivo para el logro de aprendizajes	12
3. La coexistencia de evaluaciones nacionales y subnacionales de aprendizaje a gran escala en el sistema federal argentino	15
4. Lineamientos para construir sistemas de evaluación de aprendizajes provinciales efectivos en Argentina	21
5. Reflexiones finales	27
Bibliografía	29
Anexo	31
Acerca de los autores	35

Resumen

Las evaluaciones de aprendizaje a gran escala se han consolidado como una política educativa cuya implementación ha crecido significativamente en los últimos años en todo el mundo. Su importancia radica en que, si se implementan correctamente, pueden utilizarse tanto para monitorear los aprendizajes de los y las estudiantes como para mejorarlos. En un contexto donde una elevada proporción de niños, niñas y adolescentes que asisten a la escuela no alcanzan resultados satisfactorios, estas evaluaciones se presentan como una herramienta fundamental para impulsar mejoras.

No obstante, diseñar un sistema de evaluación efectivo, es decir, con la capacidad de impulsar mejoras en los aprendizajes, no es una tarea sencilla. Requiere tomar un conjunto de decisiones estrechamente vinculadas entre sí que resulte en un conjunto coherente y balanceado de dispositivos, que respondan a propósitos diferentes, y a su vez generen información oportuna y relevante para que los distintos actores del sistema educativo puedan tomar decisiones orientadas a mejorar los aprendizajes.

En países federales, donde existen responsabilidades compartidas en la planificación y la gestión de la educación, el desafío suele ser mayor dado que conviven sistemas de evaluación nacionales y subnacionales. Esto es lo que se observa, por ejemplo, en países como Australia, Brasil, Canadá, Estados Unidos, España o México. Esta coexistencia de sistemas de evaluación está justificada en ambos niveles de gobierno. Desde la perspectiva nacional, en la definición de objetivos de aprendizaje que son comunes a todos los Estados subnacionales, en la posibilidad de disponer de resultados comparables entre las distintas jurisdicciones estados de un país, y de diseñar una sola evaluación que pueda ser replicada en todo el territorio, en tanto economía de escala. Desde la perspectiva subnacional, en las necesidades particulares de información de esos sistemas educativos específicos. En rigor, es difícil pensar que un solo nivel de la gestión educativa del sistema tenga la capacidad de implementar todos los tipos de evaluaciones necesarias.

En Argentina, la Ley Federal de Educación de 1993 sentó las bases de la evaluación educativa a gran escala en Argentina y dispuso que es responsabilidad del poder ejecutivo nacional, en coordinación con las provincias, evaluar de manera permanente al sistema educativo. Transcurridos más de 30 años desde su sanción, las evaluaciones nacionales se han sostenido a lo largo del tiempo, a pesar de los cambios de gestión, lo cual refleja un consenso sobre la importancia de su implementación. En paralelo, la mayoría de las provincias han desarrollado evaluaciones propias, sobre todo en años recientes: desde 2019 se registra que 21 de las 24 jurisdicciones han realizado evaluaciones propias o planean hacerlo en lo que resta de 2024.

Ahora bien, el auge evaluativo actual en Argentina resulta en una gran cantidad de dispositivos de evaluación diseñados e implementados a nivel nacional y provincial que no necesariamente coexisten en armonía. Existen situaciones de solapamiento que conspiran contra la eficiencia del uso de los recursos públicos y el uso efectivo de la información que emerge de las evaluaciones. Si bien estas superposiciones podrían responder a que algunas provincias estén buscando evaluar contenidos que les son propios o que haya otras necesidades específicas que no son satisfechas por los dispositivos nacionales, es muy probable que

este solapamiento responda más bien a la falta de coordinación entre niveles de gobierno. Por tanto, lograr que la coexistencia de dispositivos se transforme en un sistema de evaluación efectivo para apoyar la mejora de aprendizajes es un gran desafío.

En una publicación previa (Torre et al, 2023), con la mirada puesta en las definiciones a nivel nacional, junto a otras recomendaciones, se subrayó la importancia de contar con una planificación de las evaluaciones nacionales a gran escala que tenga un horizonte temporal más amplio (diez años), de modo de dar mayor previsibilidad y favorecer la coordinación con las jurisdicciones.

Este documento coloca el foco en las provincias y propone cuatro preguntas para orientar el diseño o fortalecimiento de los sistemas provinciales de evaluación de aprendizajes: 1) ¿Cuáles son los objetivos del sistema de evaluación provincial?; 2) ¿Cuáles son los dispositivos de evaluación más adecuados para responder de manera efectiva a los objetivos propuestos?; 3) ¿Existen necesidades de información que pueden satisfacerse con los dispositivos de evaluación nacionales e internacionales?; 4) ¿Qué dispositivos de evaluación provinciales se requieren?

La meta es contar con un sistema provincial sustentado en una teoría de cambio sólida, con dispositivos bien definidos (en términos de cobertura, áreas de conocimiento, grados/años, frecuencia) que, en complementariedad con los esfuerzos evaluativos nacionales, se transformen en una herramienta efectiva para mejorar los aprendizajes.

Introducción

Los resultados de las evaluaciones a gran escala en Argentina revelan una brecha significativa entre los aprendizajes esperados y los alcanzados por nuestros estudiantes. En primaria, según Aprender 2023, 3 de cada 10 estudiantes no alcanzan un nivel satisfactorio de aprendizajes en Lengua, mientras que 5 de cada 10 no lo hacen en Matemática. En el nivel secundario, según Aprender 2022, la proporción sube a 4 de cada 10 en el caso de Lengua y a 8 de cada 10 en Matemática. Esto implica que una proporción elevada de niños, niñas y adolescentes que están dentro del sistema educativo no alcanza aprendizajes satisfactorios.

Las evaluaciones estandarizadas a gran escala permiten conocer los logros de aprendizaje de un grupo de estudiantes en un año determinado y en un número limitado de asignaturas o competencias. A menudo se clasifican como evaluaciones subnacionales, nacionales o internacionales (según el o los sistemas educativos sobre los que se implementen), y la estandarización implica una uniformidad de procesos de implementación, que involucra contenidos, procesos y tiempos de administración, metodología de corrección y generación de resultados. Si se implementan correctamente, además de permitir un seguimiento de la evolución de los aprendizajes, sus resultados pueden ser un insumo esencial para la mejora de la calidad educativa.

No obstante, para que se transformen en una verdadera herramienta para la mejora, es necesario tomar un conjunto de decisiones estrechamente vinculadas entre sí (Torre et al., 2023). 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 los resultados de la evaluación, los productos que se derivarán de estos 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.

En países federales, donde existen responsabilidades compartidas con respecto a la planificación y gestión del sistema educativo, suelen convivir dispositivos de evaluación subnacionales, nacionales e, incluso, internacionales. Tal como mencionamos anteriormente, esta coexistencia no garantiza un funcionamiento armonioso entre las evaluaciones existentes. Lejos de ser una tarea sencilla, lograr que este conjunto de dispositivos se complemente entre sí y se transforme en un sistema de evaluación que contribuya a la mejora de aprendizajes es un gran desafío.

En Argentina, la política nacional de evaluación de aprendizajes se ha mantenido a lo largo de más de treinta años y hoy convive con dispositivos provinciales de evaluación educativa. Una de las debilidades principales de las políticas de evaluación a gran escala en Argentina es la falta de definiciones acerca de cómo se espera que ocurran las mejoras en el sistema educativo a partir de sus resultados. En consecuencia, se ha generado una gran cantidad de información sobre el desempeño del sistema educativo en términos de aprendizajes, pero que, en general, ha sido subutilizada como herramienta para la mejora de su calidad. Por otra parte, se observa una falta de articulación entre las políticas de evaluación nacional y provinciales: hay dispositivos provinciales que evalúan las mismas áreas de conocimiento que los nacionales, el mismo grado o año y hasta a veces se repite el año de implementación.

Aumentar la coordinación entre las políticas nacional y provinciales de evaluación –considerando también las evaluaciones internacionales en las que participa la provincia o el país–

resulta clave para favorecer la complementariedad y evitar superposiciones. Esto permitiría avanzar hacia una visión de integralidad de un sistema de evaluación de aprendizaje donde los distintos niveles de la gestión educativa desempeñen roles definidos.

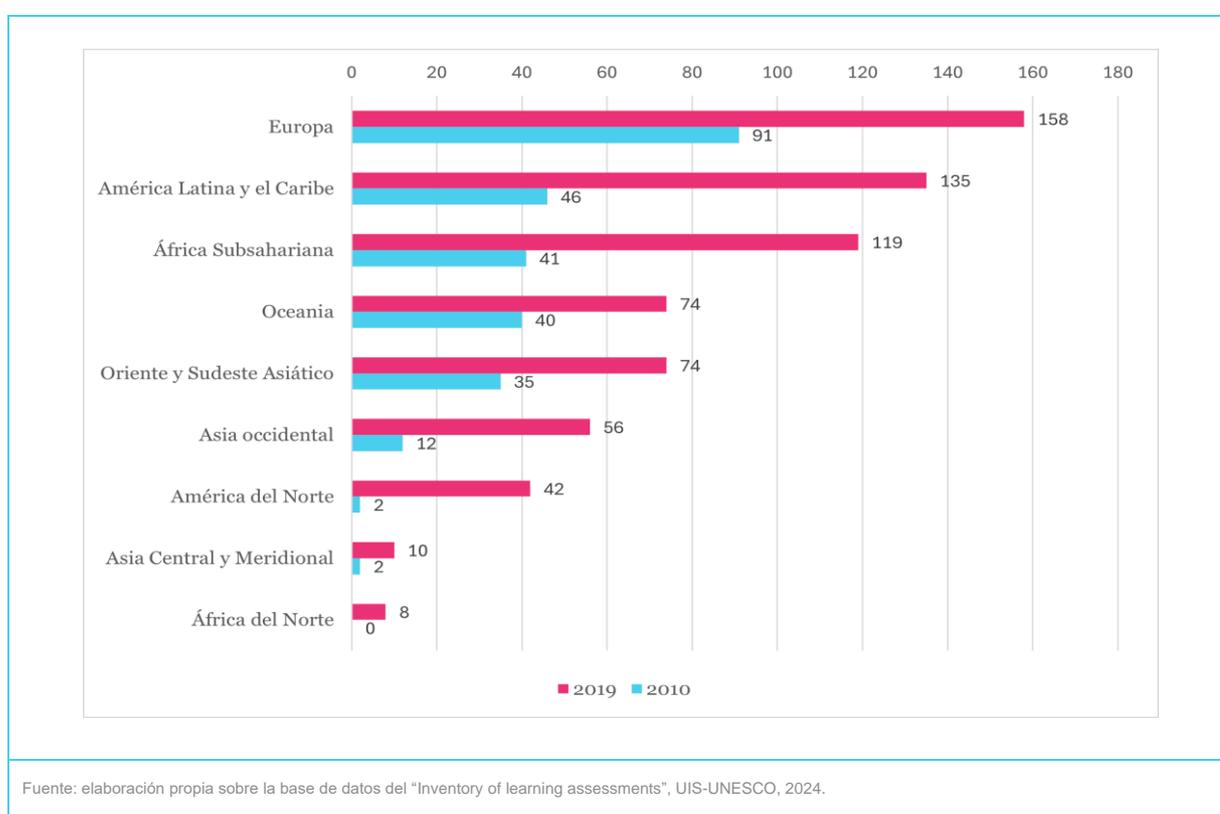
El objetivo de este documento es presentar un conjunto de lineamientos para pensar el diseño o fortalecimiento de un sistema provincial de evaluación que resulte efectivo para la mejora de los aprendizajes, en el marco del sistema educativo federal argentino.

En la primera sección se presenta un panorama sobre las evaluaciones de aprendizaje a gran escala en algunos países federales, mostrando que es usual que coexistan dispositivos nacionales y subnacionales. En la segunda sección se hacen algunas consideraciones teóricas para abordar el desafío de construir un sistema subnacional de evaluación de aprendizajes en el marco de países federales. En la tercera sección se analiza el caso argentino, describiendo las características generales de la política nacional de evaluación y de los dispositivos provinciales de evaluación desarrollados en los últimos años. En la cuarta sección se propone un conjunto de lineamientos para diseñar un sistema provincial de evaluación de aprendizajes que resulte coherente, balanceado y efectivo que contribuya a la mejora de los aprendizajes en Argentina. Por último, en la quinta sección, se proponen algunas reflexiones finales.

1. Las evaluaciones de aprendizaje a gran escala en los países federales

Las evaluaciones de aprendizaje a gran escala constituyen una política educativa cuya implementación ha crecido en forma notoria en los últimos años: mientras que en 2010 se registraron 269 operaciones evaluativas de este tipo en distintos países del mundo con alcance nacional e internacional, esta cifra se incrementó a 676 en 2019. O sea, más del doble en el transcurso de nueve años. En la actualidad, estas evaluaciones están presentes a nivel global, aunque en los últimos 10 años se ha constatado un mayor crecimiento en su implementación en las regiones compuestas mayoritariamente por países en desarrollo (**Gráfico 1**).

GRÁFICO 1. Cantidad de evaluaciones nacionales e internacionales por regiones geográficas (2010 y 2019)



La importancia de esta herramienta de política educativa radica en que, si se implementa correctamente, puede utilizarse no sólo para monitorear los aprendizajes de los y las estudiantes, sino, lo que es más importante, para mejorarlos (UNESCO UIS, 2018). En esa dirección, la medición de los aprendizajes y un mejor seguimiento de estos, junto con la utilización de los resultados de la evaluación para orientar la acción a distintos niveles, son acciones fundamentales para la construcción de sistemas educativos fuertes que permitan enfrentar las crisis de los aprendizajes que acontece a nivel global (Banco Mundial, 2018; Torre et al., 2023).

En países federales, donde existen responsabilidades compartidas con respecto a la planificación y gestión del funcionamiento del sistema educativo entre los niveles de gobierno nacional y subnacionales (comúnmente provincias o estados), los dispositivos de evaluación a gran escala implementados por los niveles subnacionales son usuales. La **Tabla 1** presenta un panorama de la implementación de estos dispositivos de evaluación en un grupo de países federales seleccionados: Australia, Brasil, Canadá, Estados Unidos y España. En todos estos países se verifica la coexistencia entre sistemas de evaluación gestionados por el gobierno central con dispositivos de evaluación provinciales o estatales.

TABLA 1. Dispositivos de evaluación a gran escala nacionales, subnacionales e internacionales en países federales seleccionados (2019-2024)

País	Evaluaciones nacionales	Entidades subnacionales con evaluaciones propias	Evaluaciones internacionales con representatividad en entidades subnacionales
Australia	<p>National Assessment Program – Literacy and Numeracy (NAPLAN). Evalúa alfabetización y competencias aritméticas.</p> <p>National Assessment Program – Civics and Citizenship (NAP–CC). Evalúa educación cívica y ciudadanía.</p> <p>National Assessment Program – Science Literacy (NAP–SL). Evalúa alfabetización científica.</p> <p>National Assessment Program – Information and Communication Technology Literacy (NAP–ICTL). Evalúa las competencias de uso de las tecnologías de la información y las comunicaciones.</p>	Nueva Gales del Sur, Queensland, Australia Occidental, Territorio de la Capital Australiana.	
Brasil	<p>Sistema de Avaliação da Educação Básica (Saeb). Evalúa lengua portuguesa, matemática, ciencias naturales y ciencias sociales.</p> <p>Exame Nacional do Ensino Médio (Enem) (no obligatorio). Evalúa lengua portuguesa y matemática.</p>	Ceará, San Pablo, Río de Janeiro, Minas Gerais, Amapá, Amazonas, Bahía, Distrito Federal (Brasilia), Goiás, Espírito Santo, Maranhao, Mato Grosso, Mato Grosso del Sur, Pará, Paraíba, Paraná, Pernambuco, Piauí, Río Grande del Norte, Río Grande del Sur, Rondonia, Roraima, Santa Catarina, Tocantins.	<p>PISA 2022</p> <p>Hay desagregación de resultados por regiones geográficas (Norte, Noreste, Sur, Sureste, Medio Oeste), aunque no por estados.</p>

Canadá	Pan-Canadian Assessment Program (PCAP). Evalúa lectura, matemática y ciencia.	Columbia Británica, Nueva Escocia, Yukón, Nueva Brunswick, Manitoba, Ontario, Isla Príncipe Eduardo, Saskatchewan, Alberta, Territorios del Noroeste, Nunavut.	<p>PISA 2022</p> <p>Ontario, Quebec, Alberta, Isla Príncipe Eduardo, Manitoba, Nueva Brunswick, Saskatchewan, Terranova y Labrador, Nueva Escocia, Columbia Británica</p> <p>PIRLS 2021</p> <p>Alberta, Columbia Británica, Terranova y Labrador, Ontario, Quebec</p> <p>TIMSS 2019 y 2021</p> <p>Ontario, Quebec</p>
España	<p>Sistema de evaluación definido por Ley de Educación (LOMLOE)</p> <p>Evaluación General del Sistema Educativo (EGS): pruebas estandarizadas nacionales con aplicación muestral, para sexto de primaria (iniciando en 2025) y para cuarto de la educación secundaria (iniciando en 2026).</p> <p>Evalúan competencia lingüística, competencia matemática y STEM, competencia plurilingüe y competencia digital</p>	<p>Evaluaciones de diagnóstico (ED) censales implementadas por comunidades autónomas anualmente y reguladas por Lomloe: 4to. de primaria y 2do. de secundaria sobre la base de un marco común nacional.</p> <p>Varias comunidades han desarrollado por años sus propias evaluaciones de distintos tipos (p. ej. Canarias, Cataluña, País Vasco).</p>	<p>PISA 2022</p> <p>Andalucía, Aragón, Asturias, Islas Baleares, Islas Canarias, Cantabria, Castilla y León, Castilla-La Mancha, Cataluña, Extremadura, Galicia, País Vasco, Comunidad Valenciana, La Rioja, Madrid, Murcia, Navarra, Ceuta, Melilla</p> <p>TIMSS 2019</p> <p>Madrid</p>
Estados Unidos	National Assessment of Educational Progress (NAEP). Evalúa lectura, matemática, educación cívica, historia de los EE.UU. y ciencia.	Alaska, Arizona, Alabama, Arkansas, Colorado, California, Connecticut, Distrito de Columbia (Washington DC), Florida, Georgia, Hawái, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Massachusetts, Michigan, Mississippi, Missouri, Nebraska, Nevada, Nueva Hampshire, Nueva Jersey, Nueva York, Carolina del Norte, Dakota del Norte, Ohio, Oklahoma, Oregon, Pensilvania, Utah, Vermont, Virginia, Estado de Washington, Virginia del Oeste, Wisconsin, Wyoming.	

Fuente: elaboración propia.

Al mismo tiempo, un segundo elemento a destacar en esta coexistencia entre evaluaciones diseñadas e implementadas tanto por el gobierno nacional como por los subnacionales en países federales es que, en ciertas ocasiones, estas últimas incluyen operativos diseñados por organismos internacionales. Los casos de Canadá y España son un ejemplo de esto, donde la evaluación internacional PISA 2022 –organizada por la OCDE– fue realizada no sólo para obtener resultados que den cuenta de los niveles de aprendizajes del país, sino también a cada uno de los sistemas educativos gestionados por provincias canadienses y comunidades autónomas españolas. En estos dos países también se registra la administración de otras evaluaciones internacionales, como TIMSS y PIRLS, las cuales se administraron con representatividad para algunas entidades subnacionales (ver cuarta columna de la **Tabla 1**).

¿Por qué los Estados subnacionales implementan evaluaciones educativas propias a gran escala cuando ya existen pruebas nacionales?

El diseño e implementación de un sistema de evaluación educativa, con la capacidad efectiva de impulsar mejoras de los aprendizajes, requiere definir con claridad los propósitos de los distintos dispositivos que lo componen, identificando destinatarios y usos esperados de la información.

La **Tabla 2** propone una categorización de evaluaciones según tres propósitos generales y detallan ejemplos de posibles usos esperados de los resultados de las pruebas, distinguiendo entre aquellas evaluaciones que se conocen como de “implicancias altas” de las de “implicancias bajas”. Las evaluaciones de implicancias altas refieren a aquellas cuyos resultados tienen una consecuencia directa y tangible sobre los sujetos evaluados o sobre las instituciones educativas involucradas y/o su personal. Por otro lado, en las de implicancias bajas, los resultados no tienen una consecuencia directa sobre los sujetos o entidades evaluadas¹.

TABLA 2. Tipos de evaluaciones de aprendizajes a gran escala según su propósito general

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

¹ Los términos que se utilizan en la literatura de investigación educativa en inglés para las evaluaciones de implicancias altas y bajas son “*high-stake assessments*” y “*low-stake assessments*”, respectivamente.

<p>¿Quiénes son los potenciales destinatarios de la información?</p>	<ul style="list-style-type: none"> • Responsables gubernamentales de formular las políticas educativas. • Organizaciones multilaterales de crédito y de cooperación internacionales. • Supervisores escolares, escuelas, docentes, institutos de formación docente, padres con hijos en edad escolar. 	<ul style="list-style-type: none"> • Distintos niveles del sistema educativo. • Escuelas, docentes, padres con hijos/as en edad escolar. 	<ul style="list-style-type: none"> • Distintos niveles del sistema educativo • Estudiantes
<p>¿Qué usos previstos específicos posibilitan?</p>	<ul style="list-style-type: none"> • Desarrollar políticas de apoyo a las escuelas, generales y focalizadas, para la mejora de los resultados. • Producir conocimiento sobre logros educativos y factores asociados a estos. • Orientar la definición del currículo y la elaboración de materiales pedagógicos. • Fortalecer programas de formación docente. 	<ul style="list-style-type: none"> • Otorgar incentivos económicos a escuelas por participación en evaluaciones. • Otorgar incentivos económicos, a escuelas y/o docentes, según el desempeño de sus estudiantes. • Propiciar un mercado competitivo utilizando los resultados para construir rankings de escuelas según su desempeño y hacerlos públicos. 	<ul style="list-style-type: none"> • Acreditar la finalización exitosa de un nivel educativo. • Establecer un mecanismo de selección para acceder al nivel superior (usualmente terciario o universitario).

Fuente: adaptado de Torre et. al, 2023.

La complejidad en la definición de los propósitos de un sistema de evaluación educativa, destinatarios y usos de los resultados se incrementa en los países federales donde un conjunto importante de atribuciones vinculados a la gestión del sistema educativo se asigna a los Estados subnacionales que lo componen. Esto sucede porque estos gobiernos, al ser responsables de la gestión de sus sistemas educativos, muchas veces fijan objetivos específicos para ellos. De esta manera, estos objetivos pueden conducir a demandas particulares de información, donde los sistemas de evaluación propios son una de las herramientas que les permita generar la evidencia requerida de apoyo a la gestión educativa.

De esta manera, la existencia de un conjunto de dispositivos de evaluación subnacionales se justificará en la medida en que respondan a propósitos específicos jurisdiccionales y tengan la capacidad de complementar de manera efectiva el funcionamiento del sistema nacional de evaluación, evitando una superposición conceptual y operativa entre ambos.

En vínculo con este último punto, vale advertir el riesgo que conlleva la falta de coordinación entre evaluaciones nacionales y subnacionales. En algunas situaciones en el sistema educativo argentino, se han verificado evaluaciones provinciales que en muchas instancias evalúan las mismas áreas de conocimiento que las evaluaciones nacionales, el mismo grado o año, e incluso a veces repitiéndose el año de implementación (Torre et al., 2023,).

La falta de coordinación no sólo conspira contra la eficiencia en el uso de los recursos públicos, sino que deslegitima los instrumentos de evaluación al generar cuestionamientos por parte de educadores, padres y estudiantes, sobre la realización excesiva de pruebas y la consecuente pérdida de valioso a la enseñanza y el aprendizaje (Sigman y Mancuso, 2017).

A esto se agrega que la generación constante de una gran cantidad de información puede atentar contra su uso efectivo. Al dedicar tanto tiempo y recursos a la producción de datos evaluativos, se reducen los espacios y posibilidades para analizar esta información y desarrollar planes efectivos para la mejora de los aprendizajes.

2. De un conjunto de evaluaciones educativas a un sistema de evaluación efectivo para el logro de aprendizajes

Como ya observamos, la coexistencia de evaluaciones nacionales y subnacionales en países federales no garantiza automáticamente el funcionamiento armonioso de estos dispositivos para lograr los objetivos que se proponen. En otras palabras, “un conjunto de evaluaciones no constituye un sistema, así como una pila de ladrillos no constituye una casa” (Coladarci, 2002). El desafío entonces es cómo hacer que esta coexistencia de dispositivos se transforme en un sistema de evaluación efectivo para apoyar la mejora de aprendizajes.

Un sistema de evaluación efectivo puede definirse como un conjunto coherente y balanceado de dispositivos, que responden a propósitos diferentes, y generan información oportuna y relevante para que los distintos actores del sistema educativo puedan tomar decisiones orientadas a mejorar los aprendizajes.

De la definición se desprenden los siguientes puntos a destacar:

- En un sistema de evaluación efectivo hay lugar para distintos tipos de dispositivos evaluativos.
- Los distintos dispositivos permiten responder a necesidades de información asociadas a diferentes fines o propósitos y usos esperados de la información.
- Los usos esperados de la información pueden definirse para los distintos niveles de la gestión del sistema educativo: nacional y subnacionales, pudiendo extenderse a niveles intermedios de supervisión, instituciones educativas y a nivel del aula de clases.

A continuación, la **Tabla 3** presenta tres aspectos identificados por la investigación sobre evaluación educativa como claves para la búsqueda de la efectividad de este sistema.

TABLA 3. Definiciones importantes para diseñar un sistema de evaluación efectivo

Dimensión	Concepto
Propósito y teoría de cambio	<p>Las evaluaciones se desarrollan y diseñan para cumplir un propósito específico. Un sistema de evaluación efectivo debería incluir diferentes tipos de evaluación, alineados con los distintos objetivos, para proporcionar información a educadores y responsables de política en los diferentes niveles del sistema.</p> <p>A la definición de propósito y objetivos específicos sigue la identificación de usuarios y usos esperados de la información, especificación de los mecanismos que se espera movilicen los cambios deseados, y finalmente el monitoreo de la medida en que las necesidades de los usuarios se satisfacen.</p>
Balance de dispositivos y eficiencia	<p>Para equilibrar las distintas evaluaciones se requiere reflexionar sobre qué puede aportar cada prueba y su eficiencia, es decir, aprovechar al máximo los recursos y eliminar las evaluaciones redundantes, no utilizadas y fuera de plazo, permitiendo que cada dispositivo haga aquello para lo que está diseñada.</p>
Coherencia	<p>Refiere a cómo el currículo, el proceso de enseñanza y la evaluación deben estar alineados para garantizar que todo el sistema esté trabajando hacia un conjunto común de objetivos de aprendizaje.</p>

Fuente: elaboración propia en base a Chatergoon & Marion (2016) y Sigman & Mancuso (2017).

Definir lo que un sistema efectivo de evaluación es o debería ser resulta más sencillo que diseñarlo. Esto se debe a que en la organización y gestión de los sistemas educativos intervienen múltiples niveles de decisión y acción cuya coordinación en la práctica suele ser dificultosa.

En contextos de países federales, como ya se expuso, existen responsabilidades concurrentes entre los gobiernos nacional y subnacionales en la planificación y gestión de la educación. El peso que recae respecto de estas responsabilidades sobre esas entidades varía según la organización de cada país, sumado al hecho de que, además, muchas acciones pueden ser descentralizadas en las propias instituciones educativas. Dado esto, en principio, no es posible tener una definición taxativa sobre qué rol debería cumplir cada nivel de gestión en ese sistema de evaluación efectivo. Sin embargo, sí existen dos elementos relevantes a destacar.

Primero, la presencia de los sistemas nacionales de evaluación está plenamente justificada aun en contextos de organización descentralizada de la educación. Esta justificación se sustenta en que estos sistemas son los que permiten monitorear la consecución de objetivos de aprendizajes establecidos para los y las estudiantes, y que son comunes entre los Estados subnacionales –por ejemplo, el desarrollo de competencias en comprensión lectora y matemáticas–, objetivo clave en la responsabilidad del Estado de asegurar el derecho a la educación. Luego, contar con resultados comparables entre los estados o provincias de un país agrega valor a la interpretación de los resultados que se obtienen en cada jurisdicción en pos de evaluar niveles alcanzados y avances. Finalmente, desde un punto de vista operativo de las evaluaciones, existen importantes economías de escala para el diseño de estas, lo que es positivo para la reducción de costos en la gestión del sistema de evaluación y lograr así un esquema eficiente de financiamiento.

Segundo, la multiplicidad de objetivos de política y de gestión escolar que muchas veces se persiguen requieren instrumentos de evaluación de distinta naturaleza, cuya administración

también recae mejor en uno u otro nivel. De esta manera, resulta previsible que la coexistencia de dispositivos evaluativos, está también justificada desde la óptica de que es muy difícil pensar en que un solo nivel de la gestión educativa del sistema tenga la capacidad de implementar todos los tipos de evaluaciones necesarias.

Con respecto a cómo afrontar el diseño de un sistema de evaluación efectivo para responder a distintas necesidades de la gestión educativa, una clave radica en cómo se coordinan las distintas evaluaciones y la utilización de sus resultados, definiendo de la mejor manera sus distintos propósitos, balance entre dispositivos y coherencia, para que en conjunto actúen de manera virtuosa en impulsar una mejora de los aprendizajes.

El diseño de un balance adecuado requiere conocer que existen distintas evaluaciones según el momento del año escolar en que se administran y los contenidos que cubren, por lo que sirven a propósitos diferentes: evaluaciones formativas y evaluaciones sumativas.

Estas tienen propósitos y enfoques distintos en el proceso de enseñanza y aprendizaje. La evaluación formativa es un proceso continuo que se lleva a cabo durante el desarrollo del aprendizaje, con el objetivo de proporcionar retroalimentación tanto al docente como al estudiante para ajustar y mejorar el proceso educativo. Permite identificar fortalezas y áreas de mejora en tiempo real. Por otro lado, la evaluación sumativa se realiza al final de un período de instrucción y su propósito es medir y evaluar el rendimiento final del estudiante en relación con los objetivos de aprendizaje. En su concepción inicial, mientras que la evaluación formativa es diagnóstica y orientada a la mejora, la sumativa se enfoca en la constatación o certificación del aprendizaje adquirido.

Los sistemas de evaluación a gran escala implementados por estados nacionales o subnacionales han estado mayormente compuestos por instrumentos de evaluación de tipo sumativo, que buscan caracterizar los logros educativos de los y las estudiantes al final de un grado o nivel educativo determinado en áreas de conocimiento determinadas.

Sin embargo, en los últimos años también se han desarrollado experiencias de evaluaciones a gran escala formativas que se organizan con planificación y algún nivel de estandarización por parte de los equipos de gestión educativa central. De esta forma, se brindan a los y las docentes los instrumentos de evaluación específicos a utilizar, para luego orientarlos sobre cómo utilizar los resultados de las evaluaciones en el aula para modificar su enseñanza y satisfacer las necesidades de los y las estudiantes que necesitan más apoyo. Adicionalmente, muchas veces tanto docentes como estudiantes también reciben materiales, incluidos libros de texto y guías para docentes, en el marco de esas mismas acciones evaluativas.

En la región, una de las experiencias más reconocidas en el desarrollo de estos sistemas de evaluación formativa –en combinación con evaluaciones sumativas estandarizadas a gran escala–, son los casos de algunos estados de Brasil, como por ejemplo Ceará y São Paulo ([Meraz Velazco, 2021](#))². De todo modos, debe destacarse que durante el período de pandemia varios países en América Latina, como Chile, Colombia, Guatemala, Honduras, Panamá y Uruguay avanzaron en el desarrollo de directrices y plataformas para promover la utilización de la evaluación formativa ([Perusia, 2021](#)). A continuación, en la sección siguiente, se examina el caso de Argentina en profundidad.

² El caso de Ceará y de uno de sus municipios, Sobral, ha sido en los últimos años un caso de estudio en la investigación educativa por los logros alcanzados en resultados de aprendizaje, reconociéndose que uno de los instrumentos de política más relevantes utilizados fueron las evaluaciones educativas. Para una síntesis del caso, puede consultarse [Luna-Bazaldúa y Cruz, 2024](#).

3. La coexistencia de evaluaciones nacionales y subnacionales de aprendizaje a gran escala en el sistema federal argentino

En Argentina, el gobierno nacional y las provincias comparten la responsabilidad principal e indelegable de proveer una educación integral, permanente y de calidad. Para garantizar la calidad de la enseñanza impartida, la Ley Federal de Educación de 1993 dispuso que es responsabilidad del poder ejecutivo nacional, en coordinación con las provincias, evaluar de manera permanente al sistema educativo nacional.

Esta determinación fue ratificada por la Ley de Educación Nacional (Ley 26.206), sancionada en 2006, que también estableció que las jurisdicciones debían participar en el desarrollo e implementación de la evaluación periódica del sistema. Esta ley establece además que la política nacional de evaluación debe ser concebida como un instrumento de mejora de la calidad de la educación.

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 se ha desarrollado y construido desde la visión de las evaluaciones de carácter diagnóstico (Torre et al., 2023).

El marco normativo vigente determina un modelo federal “abierto o de consenso”, en donde se definen el propósito general de la evaluación, pero se deja la definición de usuarios y tipos de usos previstos, cobertura, grados o años y contenidos a evaluar para la concertación en el marco del Consejo Federal de Educación (Pinkasz, 2021). Este organismo, que reúne a la máxima autoridad educativa a nivel nacional y a los ministros de educación de las 24 jurisdicciones, es el responsable de generar acuerdos sobre las características específicas de la política de evaluación nacional.

A continuación, en las dos subsecciones que siguen, se reseña el marco organizativo del sistema de evaluación nacional y los dispositivos provinciales de evaluación de aprendizajes a gran escala existentes en los años recientes³.

El Plan Nacional de Evaluación Educativa: marco de referencia sobre el propósito y las características de las evaluaciones nacionales

En 2021 y 2022 se acordaron los planes nacionales de evaluación educativa 2021-2022 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. En este sentido, representan un avance en términos de planificar y dar cierto marco de previsibilidad a la implementación de los dispositivos de evaluación existentes.

Como se puede ver en la **Tabla 4**, el Plan Nacional de Evaluación Educativa 2023-2024 incluye evaluaciones de aprendizaje a gran escala, tanto sumativas como formativas. Con respecto a las primeras, se preveía, inicialmente, la implementación de dos operativos de la

³ La Ley de Educación Nacional también dio origen al Consejo Nacional de Calidad de la Educación, como un órgano de asesoramiento especializado sobre la temática. Esta institución, sin embargo, cumple un rol de asesoramiento técnico para el diseño e implementación de la política nacional de evaluación y si bien se dispuso su creación en 2006, su puesta en funcionamiento recién se dio en 2020.

evaluación nacional Aprender. El primero, en 2023, con carácter censal para evaluar a los y las estudiantes de 6° grado del nivel primario en Lengua y Matemática. Mientras que el segundo, en 2024, también con carácter censal para evaluar a los y las estudiantes del último año del nivel secundario en Lengua y Matemática. Así, se replicaba las características de los últimos operativos en cuanto a cobertura y asignaturas y estudiantes a evaluar.

No obstante, en 2024, se acordó sumar un nuevo operativo para evaluar a una muestra de estudiantes de 3° grado del nivel primario en Lengua (Res. CFE N° 472/24). Este agregado responde a la definición de la alfabetización inicial, que se desarrolla durante el primer ciclo de la educación primaria, como una prioridad por parte de la nueva gestión de gobierno. Por lo tanto, su implementación responde a la necesidad de contar con indicadores de logros de aprendizajes actualizados para realizar un diagnóstico certero del estado de la alfabetización inicial en todo el país.

TABLA 4: Dispositivos de evaluación y productos esperados en el Plan Nacional de Evaluación (2023-2024)

Dispositivo	Año de implementación	Nivel educativo evaluado	Estudiantes evaluados/as	Asignaturas evaluadas	Cobertura	Productos esperados
Aprender	2023	Primario	6° grado	Lengua y Matemática	Censal	Informe nacional Informes jurisdiccionales
	2024	Secundario	5°/6° año	Lengua y Matemática	Censal	Reporte por escuela* Informes temáticos nacionales
	2024	Primario	3° grado	Lengua	Muestral	Recomendaciones para la enseñanza
Evaluación formativa en línea - Acompañar	2023-2024	Primario Secundario	1°, 2° y 3° grado 2° y 3° año	Lengua, Matemática, Ciencias Naturales y Sociales	No aplica, plataforma de uso voluntario	Instrumentos de evaluación digitales en Lengua, Matemática, Ciencias Sociales y Ciencias Naturales adaptadas a los procesos de enseñanza y aprendizaje que se desarrollen en las aulas de los docentes que utilicen la plataforma.

Fuente: elaboración propia sobre la base de Res. CFE N° 472/24.

*Nota: en el caso de Aprender 2024 para el nivel primario en Lengua no se enviará un reporte de resultados por escuela dado su carácter muestral.

En cuanto a las evaluaciones formativas, el plan incluye, al igual que su antecesor, un dispositivo de evaluación formativa en línea: la plataforma Acompañar. Desarrollada por el gobierno nacional, habilita a las y los docentes a elaborar instrumentos de evaluación digitales en Lengua, Matemática, Ciencias Sociales y Ciencias Naturales adaptados a los procesos de enseñanza y aprendizaje que se desarrollan en sus aulas. Desde esta perspectiva, Acompañar permite evaluar para conocer qué y cómo se está aprendiendo, así como brindar retroalimentación en función de las estrategias implementadas en cada curso.

Por otra parte, a diferencia de su predecesor del año 2021-2022, el último plan nacional avanzó en precisar el conjunto de productos que se generarían a partir de estas evaluaciones. Esta definición es un aspecto positivo ya que anticipa de qué manera y en qué momento se

difundirán los resultados. No obstante, 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.

El plan nacional también menciona dos evaluaciones internacionales: ERCE y PISA. Si bien ambos dispositivos se implementarán en 2025, son parte de este plan dado que en el período comprendido (2023-2024) estaban previstas sus respectivas pruebas piloto. Esto evidencia la voluntad de continuar participando en estas dos evaluaciones internacionales, lo cual Argentina viene haciendo de forma ininterrumpida desde 1997 en el primer caso y desde el 2000 en el segundo.

Caracterización de las evaluaciones subnacionales de aprendizaje en Argentina

Las evaluaciones jurisdiccionales de aprendizajes en Argentina 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. Dos de las jurisdicciones con mayor tradición en la implementación de evaluaciones propias son la Ciudad Autónoma de Buenos Aires y Córdoba. En la primera de ellas, se destaca la continuidad de evaluaciones sumativas desde 2011 hasta la actualidad con FESBA en un inicio, luego TESBA y también FEPBA. En Córdoba, por su parte, las de las pruebas PRISMA desde 2013.

Como consecuencia de la pandemia, surgieron nuevos dispositivos de evaluación en varias jurisdicciones (Torre et al., 2023) en cuya gran mayoría 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.

Más recientemente, el Plan Nacional de Alfabetización (Decreto 579/2024) ha generado otro impulso con el desarrollo de nuevos dispositivos de evaluación en una gran cantidad de provincias. Esto se debe a que uno de sus ejes de trabajo es el monitoreo y la evaluación de sus resultados. Por lo tanto, muchos de estos nuevos dispositivos ponen el foco en la fluidez y la comprensión lectora. Un testimonio virtuoso de ello puede ser la experiencia mendocina de implementación del Censo de Fluidez Lectora que, a partir del 2021, resulta muy valiosa a la hora de pensar el diseño de los nuevos dispositivos.

Como resultado de los impulsos sucesivos a las evaluaciones provinciales, la mayoría de ellas han diseñado sus propios dispositivos. De hecho, como se observa en la **Tabla 5**, 21 de las 24 jurisdicciones han llevado a cabo evaluaciones propias desde 2019 o tienen previsto hacerlo en lo que resta de 2024⁴. A su vez, muchas de ellas cuentan con más de un dispositivo propio.

⁴ De acuerdo con el relevamiento realizado, no se encontró información pública sobre la existencia de dispositivos de evaluación de aprendizajes a gran escala propios en Neuquén, Río Negro y Tierra del Fuego para el período considerado.

TABLA 5. Resumen de las evaluaciones jurisdiccionales de aprendizajes implementadas en Argentina para los niveles Inicial, Primario y Secundario de Educación Común*, por jurisdicción y características seleccionadas entre 2019 y 2023 y planeadas para 2024

Jurisdicción	Niveles y estudiantes evaluados	Áreas definidas por los dispositivos evaluativos	Tipo de evaluaciones utilizadas**
Buenos Aires	Primario: 1°, 2°, 3°, 4° y 6° grado	Prácticas del Lenguaje, Matemática y Escritura	Formativas
Catamarca	Secundario: 6° año	Lengua, Competencia Lingüística y Matemática	Sumativas y formativas
Chaco	Inicial: Salas de 4 y 5 Primario: 3°, 6° y 7° grado Secundario: 5° año	Lengua, Matemática, Ciencias Naturales, Ciencias Sociales, Fluidez Lectora, Comprensión Lectora, Oralidad y Pro-sodia	Sumativas y formativas
Chubut	Primario: 3° y 6° grado Secundario: 1° año	Lengua y Matemática	Sumativas
Ciudad Autónoma de Buenos Aires	Inicial: Sala de 5 Primario: 1°, 2°, 3°, 4°, 6° y 7° grado Secundario: 3° año y estudiantes de 15 años	Lengua, Fluidez y Comprensión Lectora, Escritura, Matemática, Ciencias, Inglés y Pensamiento Computacional	Sumativas y formativas
Córdoba	Primario: 2° y 3° grado Secundario: 1°, 2°, 3°, 5°, 6° año y estudiantes de 15 años	Lengua y Literatura, Fluidez y Comprensión Lectora, Matemática, Ciencias, Saberes Digitales	Sumativas y formativas
Corrientes	Inicial: sala de 5 Primario: 2°, 3° y 6° grado Secundario: 1°, 2°, 3°, 6° y 7° año (técnica)	Lengua y Literatura, Fluidez y Comprensión Lectora, Matemática, Formación General, Científico Tecnológico y Formación Técnico Específico	Sumativas y formativas
Entre Ríos	Inicial: no especificado Primario: 2° grado	Vocabulario, Conciencia Fonológica y Fluidez Lectora	Formativas
Formosa	Primario: 3° y 6° grado	Lengua y Matemática	Formativas
Jujuy	Inicial: sala de 5 Primario: 1° y 3° grado Secundario: 3° año	Lengua, Lectura, Escritura, Comprensión Lectora y Matemática	Sumativas y Formativas
La Pampa	Primario: 3°, 5° y 6° grado Secundario: 3° año	Lengua y Matemática	Sumativas

La Rioja	Inicial: sala de 5 Primario: 3°, 5° y 7° grado Secundario: 1°, 2° y 3° año	Lengua, Alfabetización en Prácticas del Lenguaje, Matemática, Ciencias Sociales y Ciencias Naturales, Formación Ética y Ciudadana	Formativas
Mendoza	Primario: 2°, 3°, 4°, 5°, 6° y 7° grado Secundario: 1°, 2°, 3°, 4° y 5° año	Lengua, Fluidez, Comprensión Lectora, Matemática y Ciencias	Sumativas y formativas
Misiones	Primario: 4° grado Secundario: 3° año	Lengua y Matemática	Sumativas
Salta	Primario: 2° y 4° grado	Fluidez lectora	Formativas
San Juan	Primario: 3° y 6° grado Secundario: 1° año	Fluidez Lectora	Formativas
San Luis	Inicial: sala de 5 Primario: 1°, 2°, 3°, 4°, 5° y 6° grado Secundario: 1°, 2°, 3°, 4°, 5° y 6° año	Lengua, Fluidez, Comprensión Lectora, Escritura, Matemática y Ciencias Naturales	Sumativas y formativas
Santa Cruz	Primario: 2° grado	Lectoescritura	Sumativas
Santa Fe	Primario: 2° grado	Fluidez Lectora	Formativas
Santiago del Estero	Primario: 3°, 4°, 6° y 7° grado	Fluidez Lectora, Comprensión de Texto y Escritura	Formativas
Tucumán	Primario: 3° y 6° grado	Lengua, Lectura, Escritura y Matemática	Formativas

* Durante la investigación se identificaron provincias donde los planes de evaluación para 2024 incluyen dispositivos de evaluación que van más allá de la Educación Común. Por ejemplo, las provincias de Córdoba y La Rioja prevén la realización de evaluaciones en programas de formación Docente, mientras que Córdoba desarrolló otro dispositivo para Educación de Adultos. Sin embargo, la tabla aquí presentada no cubre esa información por fuera de la Educación Común.

** Esta clasificación se realiza considerando mayormente cómo cada jurisdicción define el propósito de sus evaluaciones. Sin embargo, este estudio no indagó en profundidad sobre el diseño de cada dispositivo, por lo tanto, no se puede descartar que existan situaciones en donde alguno de ellos posea un diseño y/o implementación que no se ajuste con precisión a la clasificación otorgada.

Fuente: elaboración propia sobre la base de Ministerio de Educación de la Nación (2023) y Res. CFE N° 471/24.

Con respecto a las características de las evaluaciones, en primer lugar, lo que se observa en los últimos años es el gran impulso de las evaluaciones formativas (ver **Anexo**). Esto es una novedad dado que los sistemas de evaluación provinciales han estado mayormente compuestos por instrumentos de evaluación sumativa desde su surgimiento. Por ejemplo, esto se evidencia en los sistemas de evaluación provinciales con mayor tradición como, por ejemplo, el de la Ciudad Autónoma de Buenos Aires y el de Córdoba, donde en los últimos años se han sumado evaluaciones formativas a las sumativas ya existentes.

En segundo lugar, los dispositivos provinciales suelen poner el foco en el nivel primario y secundario, por lo que predominan las evaluaciones formativas en el primer caso y las evaluaciones sumativas en el segundo, si bien coexisten ambos tipos en los dos niveles. A su vez,

lo que se observa en el último año, a partir del Plan Nacional de Alfabetización, es el surgimiento de evaluaciones formativas en el nivel inicial que buscan promover la alfabetización temprana de los y las estudiantes.

En tercer lugar, en lo que refiere a asignaturas evaluadas, se puede observar que la mayoría de los dispositivos evalúan Lengua y Matemática. A su vez, el impulso de las evaluaciones formativas ha permitido poner el foco en el desarrollo de algunas áreas específicas de los contenidos curriculares. La fluidez y la comprensión lectora son las más habituales, seguidas por la escritura. También se observa algún impulso incipiente de evaluaciones formativas que buscan promover competencias matemáticas. Por ejemplo, este es el caso de Mendoza, que en 2024 implementó una evaluación muestral sobre aprendizajes de Matemática, con el objetivo de realizar un diagnóstico y llevar adelante distintas acciones remediales. Excepcionalmente, se evalúan otras asignaturas como las Ciencias Naturales (Chaco, La Rioja, San Luis y CABA, Córdoba y Mendoza a través de PISA) o las Ciencias Sociales (Chaco y La Rioja). Algunas provincias, por su parte, han implementado dispositivos que responden a objetivos de gobierno específicos como el aprendizaje del Inglés y de Pensamiento Computacional en CABA o Saberes Digitales en Córdoba.

En resumen, lo que se observa al analizar las características de las evaluaciones provinciales de aprendizaje es, en el nivel inicial, el surgimiento de nuevos dispositivos de evaluación formativa con foco en la alfabetización temprana para sala de 4 y de 5. En el nivel primario, la coexistencia de evaluaciones formativas, principalmente en el primer ciclo con el objetivo de promover el desarrollo de la fluidez y la comprensión lectora, por medio de evaluaciones sumativas en 3° y 6° grado que evalúan Lengua y Matemática. Lo mismo ocurre en el nivel secundario, en donde conviven las evaluaciones formativas con foco en la fluidez y la comprensión lectora en los primeros años y las evaluaciones sumativas en 3° y 5°/6° año.

Aunque existen esfuerzos recientes por fortalecer la complementariedad, si se las compara con las evaluaciones nacionales, en muchos casos los dispositivos provinciales adquirieron características similares, evidenciando cierto grado de solapamiento. 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 redundaría 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.

Si bien estos solapamientos podrían responder a que algunas provincias estén buscando evaluar contenidos que les son propios –no tan bien reflejados en los que son cubiertos por las evaluaciones nacionales–, o que haya otras necesidades específicas de la información no satisfechas por el dispositivo Aprender, es muy probable que la superposición responda a veces a la falta de coordinación entre niveles de gobierno. En estos contextos, se debe reiterar que fortalecer la articulación entre la política nacional y las jurisdiccionales resulta clave.

4. Lineamientos para construir sistemas de evaluación de aprendizajes provinciales efectivos en Argentina

La información presentada en la sección anterior permite afirmar que la actividad evaluativa actual en el sistema educativo argentino es intensa: a la administración nacional de las pruebas sumativas estandarizadas a gran escala –que han estado presentes con bastante regularidad en el país desde 1993– se han ido agregado distintos dispositivos de evaluación provinciales, al punto tal que hoy que apenas 3 provincias no han implementado este tipo de intervenciones educativas durante el período 2019-2024.

Dada la organización específica del sistema educativo argentino, donde el Estado nacional tiene asignada la responsabilidad principal sobre la evaluación del sistema, y las provincias poseen las mayores atribuciones vinculadas a la gestión directa de las instituciones educativas bajo su jurisdicción, a continuación, presentamos las cuatro preguntas que deberían orientar el diseño o fortalecimiento de los sistemas provinciales de evaluación de aprendizajes.

1. ¿Cuáles son los objetivos del sistema de evaluación?

La identificación de los objetivos de las evaluaciones es el inicio de la discusión para diseñar un nuevo sistema de evaluación o mejorar el sistema existente.

Entre estos objetivos usualmente aparecen propósitos definidos de manera amplia: la rendición de cuentas, el desarrollo de una cultura de evaluación, el diagnóstico para la mejora de los aprendizajes. Este último es el que aparece con mayor fuerza en el marco organizativo del sistema de evaluación nacional definido en el artículo 94 de la Ley de Educación Nacional, aunque el concepto de transparencia también se hace presente allí.

La definición de objetivos de un sistema de evaluación requiere avanzar un poco más allá de la declaración de un propósito general de, por ejemplo, “diagnóstico para la mejora”, de manera que estos objetivos se transformen en definiciones operativas que posibiliten avanzar con el diseño de evaluaciones efectivas. En otras palabras, se busca definir las consecuencias esperables en términos de decisiones o acciones que se prevén tomar a partir de los resultados.

Aquí es donde aparece entonces la relevancia de diseñar una teoría de cambio que incluya la definición de destinatarios y usos esperados de la información, que consiste en el delineamiento de los mecanismos que deberían movilizar los cambios deseados. Estos usos esperados incluyen la identificación de productos como también el diseño de las estrategias de utilización de estos (por ejemplo, promoción de procesos de reflexión e instancias de trabajo en las instituciones educativas).

Con respecto a los destinatarios de los datos generados por las evaluaciones, es fundamental realizar la distinción entre audiencias y usuarios de la información a los fines de clarificar el

diseño de la teoría de cambio⁵. Los usuarios de los resultados de las evaluaciones son aquellas personas a las que se les asigna responsabilidades específicas desde la evidencia, dándoles así un rol en la política (Pinkasz, 2021). Deben recibir la información de manera clara y tener las capacidades necesarias para realizar las acciones que se espera de ellos.

Por su parte, las audiencias se constituyen en destinatarios pasivos. Aquí se ubica la ciudadanía en general y, mediante la comunicación de los resultados de las evaluaciones, entran en juego los conceptos de rendición de cuentas y transparencia en el funcionamiento de la gestión pública educativa.

2. ¿Cuáles son los dispositivos de evaluación más adecuados para responder de manera efectiva a los objetivos y teoría de cambio del sistema de evaluación propuesto?

La identificación de los dispositivos de evaluación más adecuados para la consecución de los objetivos propuestos consiste en implementar un trayecto de decisiones para definir las características del operativo de evaluación a gran escala requerido: deberán delimitarse cuáles serán los productos de la evaluación para los destinatarios identificados, determinar si es necesario implementar el operativo con una cobertura censal o muestral, decidir sobre los grados o años a evaluar y qué áreas de conocimiento se evaluarán, además de definir la periodicidad con la que será necesario realizar las evaluaciones (Torre, et al., 2023).

A continuación, la **Tabla 6** presenta algunos ejemplos sencillos de cómo la definición de objetivos de la evaluación, destinatarios y usos de la información afecta la elección del diseño más adecuado para el dispositivo de evaluación a implementar.

TABLA 6. Relación entre objetivos de una evaluación a gran escala, destinatarios y usos de la información, y tipo de dispositivo de evaluación recomendado

Objetivo de la evaluación	Potenciales destinatarios y usos esperados de la información	Dispositivo de evaluación recomendado
Evaluaciones de implicancias bajas		
Diagnóstico para la mejora	Usuarios: responsables de política educativa en el ministerio. Se espera que información guíe el diseño de políticas e intervenciones para fortalecer aprendizajes.	Evaluación sumativa muestral para dar cuenta de los niveles de aprendizaje y cuestionarios de información de contexto para identificar factores asociados al rendimiento académico.
	Usuarios: supervisores y directores de las escuelas. Se espera que estos actores lideren la elaboración de planes de mejora a mediano plazo a partir de un	Evaluación sumativa censal que da cuenta de niveles de aprendizaje por escuela en grados seleccionados.

⁵ En la actualidad, esta distinción es muy relevante en los medios de comunicación, considerando que el primer término se ha utilizado tradicionalmente para referirse a individuos o colectivos que consumen contenido de los medios, como lectores, oyentes o espectadores (Livingstone, 2015); mientras que los usuarios son vistos como participantes activos que interactúan con las plataformas y servicios de medios de diversas maneras (Lomborg y Mortensen, 2017).

	diagnóstico contextualizado de sus resultados en las pruebas.	
	Usuarios: docentes. Se espera que estos se capaciten para mejorar prácticas pedagógicas en áreas que estudiantes muestran sistemáticamente mayores debilidades.	Evaluación sumativa muestral que cubra la mayor cantidad de contenidos posible del currículo para asignaturas seleccionadas.
	Usuarios: docentes. Se espera que estos vayan priorizando su labor sobre estudiantes que no están adquiriendo las competencias esperadas durante el transcurso del año escolar.	Evaluación formativa censal que mide incorporación de contenidos específicos priorizados dentro de cada asignatura. Es clave devolución de resultados oportuna y a nivel del estudiante individual.
Evaluaciones de implicancias altas *		
Certificación de finalización de nivel educativo	Usuarios: estudiantes, instituciones educativas. El examen certifica la conclusión exitosa de un nivel educativo (por ej., secundario) y el o la estudiante puede usarlo con múltiples propósitos (postulación a estudios superiores, búsqueda laboral, etc.).	Evaluación sumativa censal que da cuenta de niveles de aprendizaje de estudiantes de último grado de nivel educativo seleccionado.
Asignación de recursos financieros las escuelas	Usuarios: autoridades ministeriales responsables de asignar financiamiento a las escuelas. Parte de los recursos se asignan por medio de algún tipo de fórmula que utiliza resultados de las evaluaciones (por ej., pagos proporcionales a niveles de participación; por alcanzar determinados niveles de aprendizaje según objetivos, etc.).	Evaluación sumativa censal que da cuenta de niveles de aprendizaje de estudiante en último grado de nivel educativo seleccionado, más cuestionarios de información complementaria para caracterizar contexto y ajustar fórmulas de distribución (por ej., por nivel socioeconómico).
<p>* En Argentina no hay experiencias de evaluaciones de implicancias altas con los objetivos consignados como ejemplos en esta tabla. Para ambos ejemplos, el diseño de dispositivos de evaluación adecuados debe tener características y precauciones muy específicas para el objetivo que se proponen. Por ejemplo, la asignación de recursos sobre la base de alguna fórmula que incorpore los resultados de las pruebas siempre se hace con un conjunto de recaudos que buscan evitar el aumento de las desigualdades de resultados, y no simplemente asignando más recursos allí donde los resultados son más altos.</p> <p>Fuente: elaboración propia.</p>		

En la práctica es muy común que un sistema de evaluación se proponga múltiples propósitos (por ejemplo, rendición de cuentas, mejora de la enseñanza y el monitoreo). Responder a esta multiplicidad de propósitos requiere el diseño de un proceso evaluativo balanceado y eficiente (**Tabla 3**), que maximice los beneficios a partir de la inversión realizada.

Para cumplir con todos estos propósitos, se requiere una planificación cuidadosa sobre qué datos se priorizarán para cada uno de esos objetivos, porque una mala utilización de la información puede llevar a tener consecuencias negativas, entre las cuales una de las más

conocidas es la de generar en las escuelas un efecto de reducción del currículo sobre el que se enseña en la búsqueda de “enseñar para las pruebas” (Chattergoon y Marion, 2016).

Finalmente, se debe resaltar que el diseño correcto de un dispositivo de evaluación de aprendizaje a gran escala es un proceso complejo que requiere cumplir con un conjunto de requisitos técnicos para garantizar que los resultados que se obtengan sean válidos y confiables. Entre estos requisitos se incluye, por ejemplo, la definición clara del marco de evaluación, una construcción rigurosa de ítems que compondrán los instrumentos del dispositivo, un muestreo con representatividad adecuada de la población a evaluar, o procedimientos estandarizados de aplicación, entre muchos otros. La falta de cumplimiento de estos requisitos resultará en que la información obtenida pierda calidad para el uso deseado⁶.

3. ¿Existen necesidades de información que puedan satisfacerse con los dispositivos de evaluación nacionales e internacionales?

Teniendo claro los objetivos de las evaluaciones educativas y los dispositivos requeridos para alcanzarlos, la acción clave que sigue para el desarrollo de un sistema provincial efectivo es la identificación de las necesidades específicas de información que pueden abordarse con las evaluaciones nacionales e internacionales ya planificadas⁷.

Para esto, cada gobierno provincial debería saber cuál será la información generada por el sistema nacional de evaluación a la que podrá acceder. Para eso, como punto de partida, contar con una política nacional de evaluación a largo plazo, de por ejemplo 10 años, donde se establezcan tipos de instrumentos de evaluación, grados y materias, es clave para la planificación evaluativa provincial (Torre et al., 2023).

Los resultados de las evaluaciones nacionales Aprender, según sean censales o muestrales, brindan evidencia para trabajar a distintos niveles del sistema educativo, por lo que la elaboración de estrategias que saquen el máximo provecho a esa evidencia es clave en el desarrollo de un sistema de evaluación provincial efectivo y eficiente.

En segundo lugar, las evaluaciones internacionales en las que participa Argentina son otra oportunidad para las provincias de acceder a datos sobre el desempeño de sus estudiantes y sus contextos. Estos estudios internacionales, como ERCE de la UNESCO o PISA de la OCDE, además de permitir medir resultados de aprendizaje en áreas básicas de conocimiento en comparación con otros países del mundo, dan una gran cantidad de información contextual y muchas veces incluyen módulos de medición sobre competencias específicas. Por ejemplo, ERCE 2019 consideró la medición de algunas competencias socioemocionales para estudiantes de 3er y 6to grado.

Por lo tanto, una decisión a evaluar es la participación en esos estudios con muestras que permitan obtener resultados representativos a nivel de la provincia⁸. En el caso de Argentina,

⁶ Los requisitos de calidad para el diseño e implementación de las evaluaciones a gran escala es una temática que se encuentra por fuera del alcance de este documento. Sin embargo, la bibliografía sobre esto es variada. Una referencia reciente sobre esta temática es la siguiente: Clarke, Marguerite; Luna-Bazaldua, Diego. 2021. [Primer on Large-Scale Assessments of Educational Achievement](#). National Assessments of Educational Achievement. Washington, DC: World Bank.

⁷ La Ley Nacional de Educación de 2006 establece, en su 96, que “La política de información y evaluación se concertará en el ámbito del Consejo Federal de Educación. Las jurisdicciones participarán en el desarrollo e implementación del sistema de evaluación e información periódica del sistema educativo, verificando la concordancia con las necesidades de su propia comunidad en la búsqueda de la igualdad educativa y la mejora de la calidad”.

⁸ Como ya se expuso anteriormente, hay países que deciden implementar estos estudios internacionales con representatividad para cada una de las entidades subnacionales que lo componen (Tabla 1).

existen algunos antecedentes de jurisdicciones que han hecho esto, el más reciente es la implementación de PISA 2022, estudio en el que Ciudad Autónoma de Buenos Aires, Córdoba y Mendoza participaron como “regiones adjudicadas” en el estudio, lo que significa que diseñaron muestras de instituciones educativas que permiten obtener resultados representativos para esas jurisdicciones ([Ministerio de Educación de la Nación, 2023](#))⁹. La Rioja, por su parte, tomó la decisión de implementar en 2023 el dispositivo ERCE Escuelas en una escala piloto¹⁰.

4. ¿Qué dispositivos de evaluación provinciales se requieren?

La determinación de los dispositivos provinciales que compondrán el sistema de evaluación surge de observar los objetivos que se hayan fijado para el sistema de evaluación que no pueden ser respondidos por los dispositivos evaluativos nacionales ya existentes¹¹. A partir de aquí, se abre una multiplicidad de posibilidades sobre la base de posibles objetivos de política y necesidades de información provenientes de las evaluaciones para el monitoreo de esas políticas.

La observación de los planes de evaluación educativa nacionales de los últimos años permite identificar cuáles son las áreas o espacios curriculares en las que los dispositivos propios provinciales pueden ser un aporte complementario, no solapado, a las evaluaciones ya existentes:

- Evaluaciones de áreas esenciales del nivel inicial, tales como precursores de la lectura como la conciencia fonológica y reconocimiento de letras.
- Evaluaciones en grados iniciales de los niveles primario y secundario, o de finales de ciclos iniciales, considerando que las evaluaciones nacionales Aprender se vienen implementando regularmente 6to. de primaria y el último de la secundaria¹².
- Evaluaciones sumativas del final del nivel primario en sistemas educativos que cuentan con una duración de 7 años, considerando que Aprender se implementa en 6to grado de ese nivel.
- Evaluaciones sobre áreas de conocimiento distintos a Lengua y Matemática, que son las usualmente evaluadas por Aprender.
- Evaluaciones a gran escala formativas en primaria y secundaria, complementando a la información que posibilita obtener la plataforma nacional “Acompañar”.

Como observamos, la **Tabla 5** muestra cómo algunas jurisdicciones ya han avanzado con sus propias evaluaciones en las áreas de vacancia identificadas. Esto significa un progreso hacia sistemas provinciales de evaluación eficientes en el aprovechamiento de recursos, aunque

⁹ PISA permite que se obtengan resultados representativos para regiones específicas, llamadas “Regiones Adjudicadas”. En Argentina, se ofrece a los gobiernos jurisdiccionales la posibilidad de participar bajo esta modalidad en caso de que sea de su interés contar con resultados representativos a nivel de la jurisdicción. Desde 2012, participó en dicha condición la Ciudad Autónoma de Buenos Aires (CABA), mientras que en 2018 se sumaron por única vez las provincias de Buenos Aires y Tucumán, junto a la provincia de Córdoba, que reiteró su participación en la edición actual junto a CABA y Mendoza, que lo hizo por primera vez.

¹⁰ ERCE Escuelas una iniciativa a cargo del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE), perteneciente a la OREAL/UNESCO Santiago, que tiene como propósito poner a disposición de los docentes de tercer grado una herramienta pedagógica que les permita contar con un diagnóstico de los aprendizajes de los estudiantes al inicio y al término de tercer grado en las asignaturas de Lectura y Matemática. Es aplicada por cada docente de aula, sus resultados son de carácter interno, y tienen el propósito de orientar las decisiones pedagógicas de los docentes de cada aula.

¹¹ Eventualmente, también se puede considerar dispositivos de evaluación internacionales a los que es posible adherir.

¹² En la sección 3 se explicó que en 2024 se prevé una evaluación nacional muestral de 3er. grado de primaria de comprensión lectora, sin embargo, al momento de la realización de este informe no hay claridad sobre la continuidad de esa evaluación en los próximos años.

esta eficiencia es apenas uno de los componentes de los que hemos definidos como sistemas efectivos para el desarrollo de los aprendizajes.

La complementariedad respecto de áreas no cubiertas por las evaluaciones nacionales es uno de los criterios a considerar para la decisión sobre qué dispositivos implementar para la evaluación de aprendizajes. Sin embargo, la identificación de estas áreas de vacancias podría realizarse sobre la base de otros criterios: se puede pensar en la necesidad de evaluar poblaciones con características específicas, modalidades educativas particulares que requieran dispositivos adaptados a sus características específicas, o incluso otro tipo de evaluaciones distintas a las de aprendizaje que se administran a estudiantes.

Si bien el abordaje de estas otras evaluaciones está por fuera del alcance de este documento, es necesario mencionar que pueden formar parte de un sistema provincial de evaluación educativa. Por ejemplo, en la provincia de Córdoba, actualmente se plantean tres componentes adicionales a las evaluaciones a gran escala de aprendizaje de los estudiantes: evaluaciones de políticas educativas, evaluación institucional y evaluación de desempeño docente.

Este tipo de enfoques evaluativos, que incorpora otras formas de evaluación, busca tener una mirada más amplia del funcionamiento y resultados de aprendizaje del sistema educativo a partir de la generación de información sobre distintos componentes que hacen al sistema.

5. Reflexiones finales

El desarrollo que se constata en Argentina de las distintas evaluaciones provinciales desde 2019, y con mayor fuerza desde 2021 luego de la pandemia, lo ubican dentro de un grupo de países federales dentro de los que la presencia de evaluaciones diseñadas e implementadas por los gobiernos subnacionales constituyen el escenario más común: para el período 2019-2024, se ha constatado la implementación de evaluaciones a gran escala de algún tipo en 21 de las 24 jurisdicciones.

Esta tendencia de desarrollo de las evaluaciones provinciales en el país es coherente con la organización federal de Argentina y la descentralización de la gestión educativa en las provincias. Esta lógica organizativa impone a las provincias la necesidad de contar con información oportuna que responda a prioridades de política específicas y sirva como insumo para el diseño de estrategias propias para abordar las problemáticas jurisdiccionales.

Un desafío de implementar instrumentos de evaluación provinciales a gran escala radica en cómo coordinar su implementación con las evaluaciones existentes a nivel nacional, de manera que todos los dispositivos actúen en su conjunto como un sistema efectivo para impulsar la adquisición de aprendizajes. En el caso argentino, se verifican algunos solapamientos entre evaluaciones nacionales y provinciales, pero también deben reconocerse que existe una multiplicidad de esfuerzos por fortalecer esta complementariedad entre sistemas para tener un balance de evaluaciones de distintos tipos.

En la línea de lograr esta complementariedad se reconocen distintas iniciativas: evaluaciones en el nivel inicial, en grados (o años) del nivel primario (o secundario) no evaluados por el gobierno nacional, evaluaciones sobre áreas de conocimiento no abordadas por las pruebas nacionales, y el desarrollo de evaluaciones formativas a gran escala.

Las evaluaciones formativas impulsadas por varias provincias e incluso por el gobierno nacional, por medio de la plataforma “Acompañar” –que se administran a gran escala y con cierto nivel de estandarización de los instrumentos evaluativos utilizados– constituyen un desarrollo novedoso tanto para el país como para la región. En Argentina, este tipo de evaluaciones en las provincias busca sobre todo apoyar el desarrollo de la fluidez y comprensión lectora en educación primaria, aunque algunas experiencias también alcanzan la educación secundaria.

Esta tendencia creciente de las evaluaciones formativas impulsadas centralmente es valiosa, aunque vale la pena reflexionar sobre la naturaleza de la intervención y en qué medida será escalable para impulsar los aprendizajes en general. La evaluación formativa, por su naturaleza, debería ser parte de las estrategias docentes que estos implementan de manera continua, a lo largo del año escolar, como parte del proceso de enseñanza regular. Dado esto, si bien estas evaluaciones formativas a gran escala son susceptibles de implementarse para impulsar los aprendizajes de alguna competencia particular –como ocurre con la fluidez lectora, por ejemplo–, estos procesos encuentran grandes dificultades para ser replicados para los distintos contenidos curriculares que se espera sean cubiertos durante el proceso educativo. El camino realmente efectivo aquí sería, en realidad, fortalecer la formación docente inicial y continua en la utilización de la evaluación formativa como parte del proceso de enseñanza.

Un segundo elemento para destacar es que, en algunas experiencias jurisdiccionales, se puede observar cómo los resultados de las evaluaciones a gran escala se utilizan como parte de los parámetros de decisión para tomar decisiones de asignación presupuestaria. Por ejemplo, la Ciudad de Buenos Aires cuenta con el programa “Escuelas en Foco” que prioriza 500

instituciones que reciben recursos adicionales. De manera similar, los resultados de algunas evaluaciones se han utilizado en la provincia de Córdoba para dotar de recursos adicionales a las instituciones priorizadas en el marco del Programa “Maestro+Maestro”, como para entrega de libros como parte del “Compromiso Alfabetizador Córdoba”.

La utilización de los resultados de las evaluaciones para distribuir recursos a escuelas o a docentes, por sí solos o combinados con otros datos educativos, es un aspecto sobre el que algunos sistemas educativos en la región han avanzado mucho. Por ejemplo, en el estado de São Paulo se ha utilizado un bono por resultados asociado al logro de determinados objetivos de aprendizaje; y el estado de Ceará asigna recursos a los municipios con una fórmula que posee un componente educativo ligado a los resultados de las evaluaciones (Meraz Velasco, 2021).

El diseño de mecanismos de asignación de recursos a las escuelas y a los docentes sobre la base de parámetros que consideren la participación de las escuelas en las evaluaciones educativas y los resultados que obtienen allí es un camino inexplorado en Argentina. Avanzar en esa dirección requiere precaución, dado que la utilización de resultados de las evaluaciones con este propósito no puede realizarse de manera directa, y debe conectarse con otros datos que permitan contextualizar la situación de los estudiantes y las escuelas, en el marco de las legislaciones de financiamiento educativo existentes. Sin embargo, vale la pena revisar en profundidad algunas experiencias de otros países en la materia para luego evaluar su posible aplicación en determinados contextos jurisdiccionales.

Un tercer elemento a resaltar es que la proliferación de dispositivos de evaluación que se ha dado en el país desde 2021, luego de la pandemia, se da en el contexto de los avances tecnológicos que posibilitan una digitalización muy eficiente de la información recolectada. La digitalización de la administración de las evaluaciones de aprendizaje permite recopilar y procesar datos muy rápidamente (en tiempo real, en algunos casos), lo que facilita un monitoreo oportuno del rendimiento estudiantil y permite realizar ajustes rápidos en las políticas educativas.

Tecnologías como la inteligencia artificial y el aprendizaje automático permiten optimizar el análisis de grandes volúmenes de datos, identificar patrones de riesgo y personalizar intervenciones en función de las necesidades específicas de los estudiantes. A través de sistemas de alerta temprana y plataformas de visualización accesibles, los resultados de las evaluaciones se pueden integrar a sistemas de gestión educativa, donde se combinan con otros datos escolares para mejorar la precisión de las estrategias de apoyo y prevención del abandono escolar.

Considerando todo esto, un gran desafío para los sistemas de evaluación futuros es el de tener la capacidad de integrar estos avances a los procesos más tradicionales de evaluaciones a gran escala, para así facilitar la integración de la información de las evaluaciones en todos los niveles de política educativa, en pos de una planificación y gestión basada en evidencia.

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>
- 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
- 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
- CFE. (mayo de 2024). Resolución N° 471/2024. *Plan Nacional y Planes Jurisdiccionales de Alfabetización*. Obtenido de <https://www.argentina.gob.ar/educacion/consejofederal/documentos-consejo-federal-de-educacion/resolucion-cfe-ndeg-47124-plan>
- CFE. (mayo de 2024). Resolución N° 472/2024. *Plan Nacional de Evaluación Educativa 2023-2024*. Obtenido de <https://www.argentina.gob.ar/educacion/documentos-cfe/resoluciones>
- Chatergoon, R. y Marion, S. (Enero de 2016). *Not as Easy as It Sounds: Designing a Balanced Assessment System*. National Association of State Boards of Education.
- Coladarci, T. (2002). Is It a House... Or a Pile of Bricks? Important Features of a Local Assessment System. *The Phi Delta Kappan*, 83(10), 772–774. <http://www.jstor.org/stable/20440251>
- Decreto del Poder Ejecutivo Nacional (julio de 2024). Decreto N° 579/2024. *Plan Nacional de Alfabetización*. Obtenido de <https://www.argentina.gob.ar/normativa/nacional/decreto-579-2024-401148>
- Ley N° 24.195 de 1993. Ley Federal de Educación.
- Ley N° 26.206 de 2006. Ley Educación Nacional.
- Livingstone, S. (2015) Audiences and publics: reflections on the growing importance of mediated participation. *Festschrift for Jay Blumler (pp.132-140)*, edited by S. Coleman, G. Moss and K. Parry.
- Lomborg, S., & Mortensen, M. (2017). Users across media: An introduction. *Convergence*, 23(4), 343-351. <https://doi.org/10.1177/1354856517700555>
- Luna-Bazaldua, D. y Cruz, L. (2024). When an assessment system works to improve learning: the case of Sobral. *World Bank Blogs*. <https://blogs.worldbank.org/en/education/when-an-assessment-system-works-to-improve-learning--the-case-of>
- Meraz Velazco, A. (2021). El uso de los resultados de las pruebas de evaluación de los aprendizajes en el planeamiento de las políticas educativas en Brasil. Buenos Aires: IIPE-UNESCO. Obtenido de: <https://www.buenosaires.iipe.unesco.org/es/publicaciones/el-uso-de-los-resultados-de-las-pruebas-de-evaluacion-de-los-aprendizajes-en-el-o>
- Ministerio de Educación de la Nación (2023). *Dispositivos de evaluación de aprendizajes de estudiantes: experiencias jurisdiccionales*. Ciudad Autónoma de Buenos Aires: Ministerio de Educación de la Nación.

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/

Pinkasz, D. (2021). 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: IPE-UNESCO. Obtenido de: <https://unesdoc.unesco.org/ark:/48223/pf0000379593>

Sigman, D. y Mancuso, M. (2017). Designing a comprehensive assessment system. WestEd.

UNESCO. (2024). Inventory of Learning Assessments. <https://tcg.uis.unesco.org/inventory-of-learning-assessments/>

Torre, E., Perusia, J. y Xanthopoulos, J. (2023). *Hacia una política nacional de evaluación educativa a gran escala al servicio de la mejora de los aprendizajes*. CIPPEC. <https://www.cippec.org/wp-content/uploads/2023/08/DPP-244-EDU-Hacia-una-politica-nacional-de-evaluacion-educativa-08.23.pdf>

UNESCO UIS. (2018). Making the Case for a Learning Assessment. Quick Guide No. 2. https://uis.unesco.org/sites/default/files/documents/quick-guide2-making-case-learning-assessments-2018-en_2.pdf

Anexo

TABLA 7. Evaluaciones jurisdiccionales de aprendizajes implementadas en Argentina, por jurisdicción y características seleccionadas entre 2019 y 2023 y planeadas para 2024

Jurisdicción	Nombre del dispositivo	Tipo de evaluación	Año/s de implementación	Estudiantes evaluados	Asignaturas y competencias evaluadas	Cobertura
Buenos Aires	Prueba Escolar	Formativa	2022 2023	Nivel primario 2022: 3º y 6º grado 2023: 3º, 6º y 4º grado	Prácticas del Lenguaje Matemática	Censal
	No especificado	Formativa	2024	1º y 2º grado	Escritura	No especificado
Catamarca	Pensar en Aprender - Aprender a Pensar	Sumativa	2019	Nivel secundario 6º año	Lengua Matemática	Censal
	No especificado	Formativa	2024	No especificado	Competencia lingüística	No especificado
Chaco	Dispositivo de Evaluación Chaqueño	Sumativa	2022 2023 2024	Nivel primario y secundario 2022-2023: 7º grado y 2º año 2024: 6º grado y 5º año	2022-2023: Matemática, Ciencias Naturales, Lengua y Ciencias Sociales 2024: Lengua y Matemática	Censal
	No especificado	Formativa	2024	Nivel inicial Salas de 4 y 5	Oralidad	No especificado
	No especificado	Formativa	2024	3º grado	Fluidez lectora, comprensión lectora y prosodia	Muestral
Chubut	No especificado	Sumativa	2024	Nivel primario y secundario 3º y 6º grado 1º año	Lengua y Matemática	Muestral
Ciudad Autónoma de Buenos Aires	FEPBA	Sumativa	2019 2021 2023	Nivel primario 7º grado	Prácticas del Lenguaje Matemática	Censal
	Evaluación de Pensamiento Computacional	Sumativa	2023 2024	Nivel primario 4º y 6º grado	Pensamiento computacional	No especificado
	TESBA	Sumativa	2019 2021 2023	Nivel secundario 3º año	Lengua y Literatura Matemática	Censal
	Evaluación de Inglés	Sumativa	2023	Nivel primario y secundario 7º grado y 3º año	Inglés	No especificado
	PISA	Sumativa	2022	Estudiantes de 15 años	Lectura Matemática Ciencias	Muestral
	Evaluación de alfabetización temprana en salas de 5 años	Formativa	2024	Nivel inicial Sala de 5	Lengua y Matemática	Muestral
	Fluidez y comprensión lectora	Formativa	2022 2023 2024	Nivel primario y secundario 2022: 4º grado 2023: 3º grado 2024: 3º y 6º grado	Fluidez y comprensión lectora	2022: Muestral. 2023 y 2024: Censal (gestión estatal) y muestral (gestión privada)
	Evaluación en base a Progresiones	Formativa	2022 2023	Nivel primario 2022: 1º, 2º y 3º grado	Lectura y escritura	2022: Muestral

			2024	2023: 1º y 2º grado 2024: 1º y 2º grado		2023: Muestral (gestión privada) y Censal (gestión estatal) 2024: Muestral (gestión privada) y Censal (gestión estatal)
	Pausas evaluativas	Formativa	2021 2022 2023 2024	Nivel primario 2021-2023: 3º y 6º grado 2024: 4º y 6º grado	Matemática y prácticas del lenguaje	Muestral (gestión privada) y Censal (gestión estatal)
Córdoba	Prisma 3	Sumativa	2019 2021 2023	Nivel primario y secundario 3º grado y 3º año	Lengua y Literatura Matemática	Muestral
	Prisma 2FP	Sumativa	2023	Nivel secundario 2º año	Lengua y Literatura Matemática Saberes Digitales	Muestral
	Prisma 5	Sumativa	2022	Nivel secundario 5º año	Lengua y Matemática	Muestral
	Prisma 6	Sumativa	2023	Nivel secundario 6º año	Lengua y Literatura Matemática	Muestral
	PISA	Sumativa	2022	Estudiantes de 15 años	Lectura Matemática Ciencias	Muestral
	Fluidez y comprensión lectora	Formativa	2024	Nivel primario y secundario 3º grado y 1º año	Fluidez y comprensión lectora	Muestral
Corrientes	Corrientes Operativo Provincial de Evaluación (COPE)	Sumativa	2022 2023 2024	Nivel primario y secundario 3º y 6º grado 3º y 6º año	Matemática Lengua y Literatura	Muestral
	COPE Digital Educación Técnica	Sumativa	2024	Nivel secundario 7º año	Formación General, Científico Tecnológico y Formación Técnico Específico	No especificado
	No especificado	Formativa	2024	Nivel inicial Sala de 5	Lenguaje oral, vocabulario y escritura	Muestral
	Estudio comparativo de aprendizaje	Formativa	2022	Ciclo Básico	Lengua y Literatura Matemática	Muestral
	Operativo de fluidez lectora	Formativa	2022	Nivel primario 2º grado	Fluidez lectora	Muestral
	Censo de Fluidez y Comprensión Lectora	Formativa	2023	Nivel primario 3º grado	Fluidez y comprensión lectora	Censal
	Fluidez y comprensión lectora	Formativa	2024	Nivel primario y secundario 6º grado 1º y 2º año	Fluidez y comprensión lectora	Censal
Entre Ríos	No especificado	Formativa	2024	Nivel inicial No especificado	Vocabulario y conciencia fonológica	Muestral
	ALEER	Formativa	2023 2024	Nivel primario 2º grado	Fluidez lectora	2023: Muestral 2024: Censal
Formosa	Evaluaciones procesuales	Formativa	2019 2020 2021 2022	Nivel primario 3º y 6º grado	Lengua Matemática	Muestral
Jujuy	Evaluación Saber Más	Sumativa	2021	Nivel secundario	Lengua	Censal

				3° año	Matemática	
	Evaluar-me	Formativa	2024	Nivel inicial y primario Sala de 5 y 1° y 3° grado	Lectura, escritura y comprensión lectora	Muestral
La Pampa	Monitoreo de Aprendizajes Pampeanos (MAP)	Sumativa	2022 2023 2024	Nivel primario y secundario 2022: 3° y 5° grado 2023: 6° grado y 3° año 2024: 3° y 6° grado	Lengua y Matemática	2022: Muestral 2023: Muestral (primaria) y Censal (secundaria) 2024: Censal
La Rioja	Evaluación formativa en la finalización de sala de 5 años	Formativa	2024	Sala de 5	Alfabetización en Prácticas del Lenguaje y Matemática	Muestral
	Evaluación Formativa para la Etapa de Intensificación de la Enseñanza	Formativa	2022 2023 2024	Nivel primario 3°, 5° y 7° grado	Lengua Matemática Ciencias Sociales Ciencias Naturales	Censal
	Evaluación Formativa de los Aprendizajes en el Ciclo Básico	Formativa	2023	Nivel secundario 1° y 3° año	Matemática Lengua Ciencias Naturales (Biología - Física y Química) Historia y Geografía Formación Ética y Ciudadana	No especificado
	No especificado	Formativa	2024	Nivel secundario 2° año	Lengua y Matemática	Muestral
	ERCE Escuelas	Formativa	2023	Nivel primario 3° grado	Lengua y Matemática	Muestral
Mendoza	Relevamiento Provincial de Aprendizaje (RPA)	Sumativa	2021 2022 2023	Nivel primario y secundario 2021: 4° y 5° grado y 2° y 3° año 2022 y 2023: 4° grado y 2° año	2022: Lengua 2021 y 2023: Lengua Matemática	Muestral
	Censo de Fluidez y Comprensión Lectora	Formativa	2021 2022 2023 2024	Nivel primario y secundario 2021: 4° y 7° grado y 1° año 2022: 3°, 5 y 7° grado y 1° año 2023: 2° a 7° grado y todos los años del nivel secundario 2024: 3° y 6° grado y 1° año	2021 a 2023: Fluidez lectora 2024: Fluidez y comprensión lectora	Censo
	Mendoza Mejora Aprendiendo Matemática	Formativa	2024	Nivel primario y secundario 4° y 7° grado 3° y 5° año	Matemática	Muestral
	PISA	Sumativa	2022	Estudiantes de 15 años	Lectura Matemática Ciencias	Muestral
Misiones	Evaluación Provincial de Logros de Aprendizaje	Sumativa	2021	Nivel primario y secundario 4° grado y 3° año	Lengua Matemática	Muestral
Salta	Alfabetización Temprana con Perspectiva Intercultural	Formativa	2023 2024	Nivel primario 4° grado	Fluidez lectora	Muestral

	¿Cómo leemos en 2º grado?	Formativa	2022 2023 2024	Nivel primario 2º grado	Fluidez lectora	2022 y 2023: Muestral 2024: Censal
San Juan	Censo de Fluidez Lectora	Formativa	2024	Nivel primario y secundario 3º y 6º grado y 1º año	Fluidez lectora	Muestral
San Luis	Evaluación de Calidad Educativa	Sumativa	2019	Nivel primario y secundario Lengua y Matemática: 2º grado a 6º año Ciencias Naturales: 3º grado a 6º año	Lengua Matemática Ciencias Naturales	Muestral
	Plataforma de lectura	Formativa	2019 2020 2021 2022 2023 2024	Nivel primario y secundario Todos los grados y años	Fluidez y comprensión lectora	No especificado
	Queremos aprender	Formativa	2024	Nivel inicial y primario Sala de 5 y 1º, 2º y 3º grado	Letoescritura	Censal
Santa Cruz	No especificado	Sumativa	2024	Nivel primario 2º grado	Letoescritura	Muestral
Santa Fe	Test de lectura	Formativa	2024	Nivel primario 2º grado	Fluidez lectora	Censal
Santiago del Estero	Evaluación muestral de Alfabetización	Formativa	2024	Nivel primario 3º y 6º grado	Fluidez lectora y comprensión de textos	Muestral
	Evaluación censal de Alfabetización	Formativa	No especificado	Nivel primario 4º a 7º grado	Comprensión lectora y escritura	Censal
Tucumán	Evaluación para la mejora de la Calidad Educativa (EMeCE)	Formativa	2024	Nivel primario 3º y 6º grado	Lengua y Matemática	Muestral
	No especificado	No especificado	No especificado	No especificado	Lectura y escritura	Muestral

Fuente: elaboración propia sobre la base de Ministerio de Educación de la Nación (2023) y Res. CFE N° 471/24.



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

Acerca de los autores

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 Lilia Toranzos y Daniel Capistrano, 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: Banco Industrial, Boldt, Fundación Telefónica-Movistar, Santillana y Zurich. Por otra parte, los autores agradecen a Armando Acuña por sus aportes a lo largo de la elaboración del documento y a Gabriela Brandán por su atenta lectura y valiosos comentarios.



Para citar este documento:

Perusia, J.C. y Xanthopoulos, J. (noviembre de 2024). *El diseño de sistemas provinciales de evaluación de aprendizajes efectivos en Argentina*. Documento de Política Pública N° 247. 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

¿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.

