

VOCES EXPERTAS

Juan Cruz Perusia

**De la diversidad a la coherencia:
articulando evaluaciones nacionales
y subnacionales en países federales**



De la diversidad a la coherencia: articulando evaluaciones nacionales y subnacionales en países federales¹

Por Juan Cruz Perusia

En las últimas décadas, las evaluaciones educativas a gran escala han experimentado un crecimiento sostenido a nivel global. Este crecimiento es evidente si se observa que en 2010 se registraron **269 operaciones evaluativas** de este tipo en distintos países del mundo, con alcance nacional e internacional, mientras que en 2019 la cifra ascendía a **676 operaciones evaluativas**, reflejando la creciente centralidad que estos instrumentos han adquirido en la gestión y gobernanza educativa.

En la actualidad, no solo son instrumentos para el monitoreo de los resultados de sistemas educativos, sino que también, si están correctamente implementadas, son un instrumento clave para diseñar políticas e intervenciones más efectivas orientadas a la mejora de los aprendizajes.

Evaluación educativa en países federales

En países federales, la gestión de la educación se caracteriza por responsabilidades compartidas entre el gobierno nacional y los gobiernos subnacionales, con grados variables de descentralización. Aunque no existe un modelo único que determine el papel exacto de cada nivel de gobierno en un sistema de evaluación, sí pueden destacarse dos elementos clave.

En primer lugar, la existencia de un sistema nacional de evaluación a gran escala está plenamente justificada incluso en contextos altamente descentralizados.

Estos sistemas permiten monitorear el logro de objetivos de aprendizaje comunes para todo el país —por ejemplo, competencias fundamentales—, garantizando el derecho a la educación y ofreciendo resultados comparables entre jurisdicciones. Además, cuentan con economías de escala en su diseño y aplicación, reduciendo costos y favoreciendo un financiamiento más eficiente de la actividad evaluativa.

En segundo lugar, la diversidad de objetivos de política y de gestión escolar requiere instrumentos de evaluación de distinta naturaleza, cuya implementación resulta más adecuada en diferentes niveles de gobierno. Por ello, la coexistencia de evaluaciones na-

¹ Este artículo toma como base el 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.

cionales y subnacionales no solo es inevitable, sino también necesaria: los sistemas subnacionales se diseñan complementando al sistema nacional, cubriendo necesidades específicas de información que éste no aborda.

La forma en que se combinan las evaluaciones nacionales y subnacionales varía de un país a otro, reflejando sus prioridades y estructuras de gestión. En Brasil, el SAEB (Sistema de Avaliação da Educação Básica) es una evaluación nacional con cobertura censal en determinados grados y áreas, que ofrece resultados por escuela y estado. A nivel subnacional, varios estados han desarrollado evaluaciones propias para monitorear aprendizajes en otros grados o con fines más formativos.

En Australia, el NAPLAN (National Assessment Program – Literacy and Numeracy) evalúa con cobertura censal a todos los estudiantes de 3°, 5°, 7° y 9° año, con reportes detallados por escuela. Su frecuencia anual y su amplitud en grados y áreas hacen que cubra gran parte de las necesidades de información del sistema educativo, por lo que la presencia de evaluaciones subnacionales adicionales es limitada.

En Canadá, el PCAP (Pan-Canadian Assessment Program) tiene diseño muestral y se focaliza en monitorear el desempeño general del sistema. Las provincias, responsables de la educación, implementan sus propias evaluaciones censales o muestrales para responder a necesidades específicas de gestión y rendición de cuentas.

En Estados Unidos, el NAEP (National Assessment of Educational Progress) también tiene un diseño muestral y permite comparaciones entre estados. La mayor parte de las evaluaciones censales se realizan a nivel estatal, con gran diversidad en áreas, grados evaluados y propósitos.

Independientemente de la arquitectura evaluativa adoptada en cada país, la coexistencia de evaluaciones nacionales y subnacionales es un hecho que abre oportunidades —más y mejor información ajustada a distintos niveles de decisión—, pero también impone desafíos de gobernanza: alinear marcos de referencia, coordinar calendarios de aplicación y, sobre todo, integrar de forma sinérgica resultados de diversas fuentes para que la evidencia se traduzca efectivamente en mejoras de los aprendizajes.

Situación actual de Argentina

Argentina es un país federal en el que la responsabilidad principal sobre la gestión de los sistemas educativos recae en las 24 jurisdicciones (23 provincias y la Ciudad Autónoma de Buenos Aires). Si bien desde 1993 el país cuenta con un operativo nacional de evaluación —denominado Aprender desde 2016—, en los últimos años ha crecido significativamente la implementación de evaluaciones provinciales a gran escala.

Entre 2019 y 2024, **21 de las 24 jurisdicciones** desarrollaron dispositivos propios de evaluación de aprendizajes, con distintos niveles de cobertura, periodicidad y propósitos. Algunas provincias han optado por evaluaciones censales para contar con resultados por escuela, mientras que otras han implementado estudios muestrales o evaluaciones focalizadas en áreas y niveles específicos. Estas iniciativas han fortalecido la capacidad de los ministerios jurisdiccionales para generar evidencia propia y retroalimentar la gestión educativa, al tiempo que han permitido atender necesidades específicas de información no cubiertas por las evaluaciones nacionales.

Sin embargo, esta expansión también ha incrementado la diversidad de enfoques y calendarios de aplicación, lo que plantea varios desafíos. **Por un lado, se requiere coordinar los distintos dispositivos provinciales con las evaluaciones nacionales**, de modo que el conjunto de herramientas funcione como un sistema coherente y eficiente para mejorar los aprendizajes. **Por otro, resulta clave garantizar que el diseño y la implementación de las evaluaciones provinciales respondan a criterios técnicos sólidos** —en aspectos como la validez y confiabilidad de las mediciones, la comparabilidad de los resultados o la transparencia de los procesos—, de manera que la información producida sea robusta y útil para la toma de decisiones. **Finalmente, es necesario definir con mayor claridad las estrategias de uso de la información que se genera.** La situación actual se caracteriza por una abundancia de datos, pero en muchos casos sin un plan preciso sobre qué decisiones de política o prácticas pedagógicas se pretende modificar con cada tipo de dispositivo evaluativo, lo que limita su potencial impacto.

Principios para abordar los desafíos para el diseño de sistemas provinciales de evaluación efectivos

La coexistencia de evaluaciones nacionales y provinciales en Argentina ofrece un potencial significativo para fortalecer las políticas educativas, pero solo si estas se articulan en un marco de coordinación que maximice su complementariedad y evite redundancias. Como se señaló, el reto no es únicamente coordinar agendas y marcos de referencia, sino también **orientar estratégicamente el uso de la información**, de modo que cada dispositivo evaluativo tenga un propósito claro y contribuya de manera específica a la mejora de los aprendizajes.

En este sentido, el diseño de un sistema de evaluación efectivo —tanto a nivel nacional como provincial— requiere atender al menos tres definiciones clave (Chatergoon y Marion, 2016; Sigman y Mancuso, 2017):

- 1. Propósito del sistema y teoría de cambio.** Es fundamental establecer con claridad para qué se evalúa y cómo se espera que la información contribuya a generar cambios en las políticas o en las prácticas escolares. El propósito del sistema define su orientación general, mientras que la teoría de cambio

explicita los mecanismos mediante los cuales los resultados de las evaluaciones se transformarán en acciones concretas de mejora, definiendo productos, destinatarios de estos y sus usos esperados.

2. Balance y eficiencia de los dispositivos. Un sistema efectivo debe encontrar un equilibrio adecuado entre la cantidad, el tipo y la frecuencia de los dispositivos evaluativos, considerando tanto las necesidades de información como la capacidad operativa de los actores involucrados. Esto implica evitar superposiciones innecesarias, asegurar que cada dispositivo aporte valor diferenciado y optimizar el uso de recursos humanos, técnicos y financieros.

3. Coherencia del sistema evaluativo. Se refiere a que el currículo, los procesos de enseñanza y la evaluación estén alineados, de modo que todo el sistema educativo trabaje hacia un conjunto común de objetivos de aprendizaje. Este alineamiento asegura que las evaluaciones midan lo que realmente se espera que los estudiantes aprendan, que la enseñanza esté orientada a esos mismos objetivos y que los resultados obtenidos puedan interpretarse y utilizarse de manera consistente para mejorar los aprendizajes.

En conjunto, estas definiciones permiten que los distintos niveles de gobierno, con sus respectivas evaluaciones, no funcionen como piezas aisladas, sino como componentes de un sistema integrado capaz de generar información relevante, confiable y orientada a la acción.

Sin embargo, diseñar sistemas de evaluación efectivos no es solo una cuestión técnica, sino también un ejercicio de sensibilidad y adaptación. Tal como ocurre en otras disciplinas, la evaluación requiere rigor metodológico, pero también la capacidad de interpretar contextos, equilibrar prioridades y ajustar diseños para que respondan a las realidades locales. No existe una fórmula universal; lo que sí existe es el desafío —y la posibilidad— de construir sistemas que, con inteligencia y flexibilidad, conviertan la información en una verdadera palanca para la mejora de los aprendizajes.

