

EVALUACIÓN EN PANDEMIA:

¿Cómo diagnosticamos las pérdidas de aprendizajes para informar los esfuerzos de recuperación?

FEBRERO 2022

Anna Herrero Tejada, Marina López Leavy, Micaela Finoli,
Felipe Hevia de la Jara, Sarah Stanton y Ariel Fiszbein

© 2022, Diálogo Interamericano. Primera Edición.
Diálogo Interamericano
1155 15th St. NW, Suite 800 Washington, DC 20005
Tel: + 1 202-822-9002 / Fax: 202-822-9553
Correo electrónico: education@thedialogue.org

Autores:

Anna Herrero Tejada, Marina López Leavy, Micaela Finoli, Felipe Hevia de la Jara, Sarah Stanton y Ariel Fiszbein

Imagen de Portada: Thomas Park, Unsplash, <https://unsplash.com/photos/6MePtA9EVDA> (CC0) La imagen ha sido editada a blanco y negro con una cubierta azul.

Diagramación: Victoria Tosi












Agradecimientos

Este informe representa un trabajo colaborativo del Programa de Educación del Diálogo Interamericano y el Banco Mundial. Los autores agradecen a todos los expertos que compartieron sus conocimientos y experticias con nosotros por medio de entrevistas y correos, y también los participantes en el taller de retroalimentación para validar el análisis y desarrollar las recomendaciones presentadas aquí. Del equipo del Diálogo Interamericano, reconocemos el trabajo infatigable de Alejandra Vargas Durango para gestionar los detalles logísticos del proyecto. Extendemos un agradecimiento especial a los colegas del Banco Mundial, quienes hicieron posible este trabajo, participaron activamente en los talleres y ofrecieron retroalimentación sobre el borrador, en particular Juan Diego Alonso, Emanuela di Gropello, María José Ramírez y María José Vargas.

Índice

AGRADECIMIENTOS	3
INTRODUCCIÓN	7
METODOLOGÍA	8
1. Evaluaciones estandarizadas	9
<i>a. Dispositivos nacionales</i>	9
<i>b. Dispositivos regionales e internacionales</i>	18
2. Evaluaciones formativas	21
<i>a. Eje I: Enfoque del contenido</i>	23
<i>b. Eje II: Alineación con programas de recuperación y aceleramiento</i>	26
<i>c. Eje III: El rol docente</i>	30
<i>d. Eje IV: Alianzas y financiamiento</i>	33
<i>e. Consideraciones para la implementación y replicabilidad de los casos</i>	35
3. Recomendaciones	36
BIBLIOGRAFÍA	39
ANEXOS	41

Lista de figuras

- 
MAPA 1. DECISIONES ADOPTADAS POR LOS PAÍSES DE LA REGIÓN RESPECTO A LA APLICACIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN EN 2020
- 
MAPA 2. DECISIONES ADOPTADAS POR LOS PAÍSES DE LA REGIÓN RESPECTO A LA APLICACIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN EN 2021
- 
GRÁFICO 1. DISTRIBUCIÓN DE LAS EVALUACIONES IMPLEMENTADAS EN EL AÑO 2020 EN LAC, SEGÚN EL TIPO DE IMPLICANCIAS
- 
GRÁFICO 2. DISTRIBUCIÓN DE LAS EVALUACIONES IMPLEMENTADAS EN EL AÑO 2021 EN LAC, SEGÚN EL TIPO DE IMPLICANCIAS
- 
RECUADRO 1. ANÁLISIS DEL CASO DE URUGUAY CON ARISTAS (2020)
- 
DIAGRAMA 1. SÍNTESIS DE CONTINUIDADES Y CAMBIOS EN DISTINTOS ELEMENTOS DE LAS EVALUACIONES NACIONALES IMPLEMENTADAS EN 2020 Y 2021
- 
RECUADRO 2. ANÁLISIS DEL CASO DE COLOMBIA CON LAS PRUEBAS SABER 3°, 5°, 9° Y 7°
- 
RECUADRO 3. ANÁLISIS DEL CASO DE CHILE CON SIMCE
- 
MAPA 3. PROYECCIÓN INFORMADA POR LOS PAÍSES DE LA REGIÓN RESPECTO A LA APLICACIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN EN 2022
- 
TABLA 1. APLICACIONES DE DISPOSITIVOS REGIONALES EN 2020 Y 2021
- 
TABLA 2. APLICACIÓN DE DISPOSITIVOS INTERNACIONALES EN 2020 Y 2021
- 
TABLA 3. INFORMACIÓN GENERAL DE LOS CASOS SELECCIONADOS
- 
DIAGRAMA 2. LOS CUATRO EJES ANALÍTICOS
- 
TABLA 4. ENFOQUE DE CONTENIDO DE LAS INICIATIVAS ANALIZADAS
- 
RECUADRO 4. LA APUESTA PARA LA EVALUACIÓN DIAGNÓSTICA DE APRENDIZAJES BÁSICOS: EL PROYECTO DE MIA
- 
TABLA 5. EL ENFOQUE DE RECUPERACIÓN DE LOS CASOS
- 
RECUADRO 5. COMPARABILIDAD Y ALINEACIÓN CURRICULAR: EVALUACIONES DIAGNÓSTICAS MEC - CAED
- 
RECUADRO 6. PROGRAMA DE TUTORÍAS DE LA FUNDACIÓN CARVAJAL
- 
TABLA 6. EJES I Y II, COMPARADOS
- 
TABLA 7. EL ROL DEL DOCENTE EN LOS CASOS ANALIZADOS
- 
RECUADRO 7. CAPACITACIÓN PARA LA EVALUACIÓN FORMATIVA EFECTIVA: PROYECTO CREER
- 
RECUADRO 8. APALANCANDO LA CAPACIDAD TÉCNICA EN SOCIEDAD CIVIL: NUEVA PRUEBA NACIONAL DIAGNÓSTICA EN LA PLATAFORMA IQ

Lista de acrónimos

ALC	América Latina y el Caribe
ANEP	Administración Nacional de Educación Pública
ASER	Annual Status of Education Report
BJAT	Belize Junior Achievement Test
BM	Banco Mundial
BNCC	Base Curricular Nacional Común (Base Nacional Comum Curricular)
CAEd	Centro de Políticas Públicas y Evaluación Educativa
CAPE	Caribbean Advanced Proficiency Examinations
CCSLC	Caribbean Certificate of Secondary Level Competence
CPEA	Caribbean Primary Exit Assessment
CSEC	Caribbean Secondary Education Certificate
CXC	Consejo de Examinaciones del Caribe (Caribbean Examinations Council)
DIA	Diagnóstico Integral de Aprendizajes
ECE	Evaluación Censal de Estudiantes
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
ERCE	Estudio Regional Comparativo y Explicativo
GRADE	Grupo de Análisis para el Desarrollo
IAD	Diálogo Interamericano
ICCS	Estudio Internacional sobre Educación Cívica y Ciudadana (International Civic and Citizenship Education Study)
LLECE	Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación
MEC	Ministerio de Educación
MEJOREDUC	Comisión Nacional para la Mejora Continua de la Educación
MIA	Medición Independiente de Aprendizajes
ODS	Objetivo de Desarrollo Sostenible
OECD	Organisation for Economic Co-operation and Development
PDT	Prueba de Transición para la Admisión Universitaria
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PSE	Primary School Examination
SAEB	Sistema de Evaluación de la Educación Básica (Sistema de Avaliação da Educação Básica)
SEA	Sistema de Evaluación de Aprendizajes
SEP	Secretaría de Educación Pública
SIMCE	Sistema de Medición de la Calidad de la Educación
SNEPE	Sistema Nacional de Evaluación del Proceso Educativo
TIMSS	Trends in International Mathematics and Science Study
UGEL	Unidad de Gestión Educativa Local
UNESCO	United Nations Educational, Scientific and Cultural Organization

INTRODUCCIÓN

La pandemia generada por el COVID-19 impactó fuertemente a los sistemas educativos de América Latina y el Caribe, donde la duración de los cierres de escuelas ha sido la más larga del mundo. Las consecuencias han afectado tanto las trayectorias educativas como las oportunidades de aprendizaje de todos los estudiantes, pero en particular – y desproporcionadamente – a aquellos en situaciones de mayor vulnerabilidad. Previa a la emergencia sanitaria, importantes brechas en la adquisición de los aprendizajes ya afectaban a los países de la región. La pobreza de aprendizajes, que define el % de población de 10 años incapaz de leer y comprender un texto simple, ya era del 53% y, según las estimaciones, aumentaría al 63%, pero los nuevos datos demuestran un incremento real que llega al 70% (World Bank, 2021:7).

La evaluación educativa es una herramienta clave y esencial para comenzar un largo proceso de medición de rezagos educativos y recuperación de aprendizaje después de la crisis. El objetivo de este informe es analizar las estrategias de evaluación implementadas en este periodo de pandemia, destacar buenas prácticas y ofrecer recomendaciones para seguir apoyando los esfuerzos de diagnóstico y recuperación a futuro.

Este informe se elabora en el marco de la alianza “Recuperando la escolarización y los aprendizajes después de la pandemia” (en inglés, *Recovering Schooling and Learning after COVID-19*) entre el Banco Mundial (en adelante BM) y el Diálogo Interamericano (en adelante IAD). El objetivo es apoyar la implementación de acciones que promuevan mayores niveles de escolarización y recuperación de los aprendizajes en la región.

En el marco de esta colaboración, se publicarán dos documentos. Este informe constituye el primero, donde se presentará información acerca de los dispositivos de evaluación de los aprendizajes estandarizados y las iniciativas de evaluación formativa que tuvieron lugar durante la pandemia. El segundo informe sistematiza evidencia de programas remediales en la región y a nivel internacional para la recuperación de los aprendizajes.

Este documento se estructura en tres apartados. El primero, presenta una caracterización general de los dispositivos de evaluación estandarizada nacionales, regionales e internacionales, las principales tendencias y las proyecciones para el 2022. Dentro del primer apartado, la primera sección analiza los cambios que se han introducido en los dispositivos nacionales y la segunda se enfoca en dispositivos regionales y nacionales. El primer apartado concluye cada sección con las tendencias principales para cada grupo de dispositivos, igual que sus proyecciones a futuro. El segundo apartado describe y analiza una selección de iniciativas de evaluación formativa gubernamentales y no gubernamentales, las cuales ilustran una variedad de modelos que pueden contribuir a los esfuerzos de recuperación de aprendizaje. Este apartado se divide en cuatro secciones, cada una enfocada en uno de cuatro ejes analíticos que definen las experiencias: el enfoque del contenido, su alineación con programas de recuperación, el rol del docente, y las alianzas y financiamiento del cual dependen. Finalmente, el tercer y último apartado ofrece cuatro recomendaciones de política que resultaron de la investigación y el taller.

METODOLOGÍA

Las siguientes preguntas guiaron la investigación y el análisis presentado en este informe:

1. ¿Qué pasó en 2020 y en 2021 con la aplicación de (i) los dispositivos estandarizados de evaluación de los aprendizajes nacionales, regionales e internacionales; y (ii) los dispositivos de evaluación formativa?
2. ¿Qué cambios e innovaciones se han introducido a partir de la pandemia?
3. ¿Cuáles son las proyecciones para las próximas ediciones de los dispositivos de evaluación?
4. ¿Cuál es el rol de ambos tipos de dispositivos en el proceso diagnóstico de pérdidas y brechas de aprendizaje, pensando en un escenario de recuperación y de resiliencia después de la pandemia?

El estudio releva y analiza información sobre evaluaciones de los aprendizajes en los niveles primario y secundario de la modalidad de educación común o regular¹ en los sistemas educativos de la región. Entre los dispositivos de evaluación relevados, el informe se focalizará en los siguientes grupos de evaluaciones:

Las evaluaciones estandarizadas (primer apartado):

- Exámenes con fines de certificación de un curso o nivel educativo (primario o secundario) o de selección/admisión para los niveles superiores. Este grupo comprende a aquellas evaluaciones cuyas implicancias son altas y las consecuencias asociadas a sus resultados son “duras”².
- Evaluaciones con fines de monitoreo, cuyas implicancias pueden ser altas y sus consecuencias más duras (por ejemplo, en aquellos casos donde los resultados se utilizan para la rendición de cuentas por parte de las instituciones educativas, o existen incentivos asociados a los resultados); o bien pueden tener implicancias bajas y consecuencias “blandas”, en la medida en que los resultados se utilizan para

realizar diagnósticos o seguimiento de distintos indicadores del sistema educativo en su conjunto.

Esta clasificación en base a la finalidad de las evaluaciones es realizada únicamente para facilitar la descripción y análisis de los dispositivos de evaluación. Además, en el primer apartado se distinguirá el alcance de los distintos dispositivos (nacional, regional, o internacional).

Las evaluaciones formativas (segundo apartado):

- Se incluyen evaluaciones diseñadas a nivel central y distribuidas a nivel del aula, así como ejercicios de diagnóstico que forman parte de programas más amplios de intervención pedagógica.
- En este informe, las evaluaciones formativas se caracterizan por su objetivo de relevar datos para informar acciones de recuperación y retroalimentación en el corto plazo, de forma paródica y regular, con un énfasis en el rol del docente, pero también por parte de las autoridades.

El relevamiento de datos para informar este estudio se llevó a cabo entre los meses de noviembre de 2021 y enero de 2022. La metodología de investigación se basó, en gran parte, en la búsqueda y revisión de documentos públicos y comunicaciones oficiales de las autoridades educativas y entidades evaluadoras.

En particular, para documentar los casos y experiencias de evaluación formativa, la revisión de información pública fue complementada por consultas a referentes específicos, incluyendo siete entrevistas a informantes clave, responsables de los casos analizados. Los hallazgos preliminares de la investigación fueron presentados en un taller, en el cual un grupo de expertos y líderes en evaluación de la región validaron los resultados y ofrecieron una serie de recomendaciones con base en sus propias experiencias.

Debido a la duración del periodo de investigación, y la metodología basada en revisión de información pública,

los hallazgos del estudio presentan ciertas limitaciones. En la mayoría de los casos, no se localizó información validada sobre el costo de las iniciativas destacadas. Adicionalmente, en algunos casos (indicados en el documento), no se confirmó el número exacto de alumnos que fueron evaluados, lo cual puede comprometer la comparabilidad entre iniciativas.

1. Evaluaciones estandarizadas: El uso de dispositivos nacionales, regionales e internacionales en pandemia

Las características de los instrumentos y las estrategias involucradas en las evaluaciones estandarizadas las posicionan como los dispositivos más apropiados para medir aprendizajes a gran escala, especialmente a nivel nacional o regional. Al ser estandarizadas, utilizan un mismo diseño y procedimientos de administración, corrección y reporte de resultados. Por lo tanto, sus resultados pueden ser interpretados como un indicador relevante de los aprendizajes de los estudiantes.

Por estos motivos, entre otros, la evaluación estandarizada de los aprendizajes se constituye como una herramienta que potencialmente puede ser utilizada para estimar las pérdidas y brechas en los aprendizajes. Para hacerlo, se requieren dispositivos con diseños que permitan la comparabilidad en el tiempo. A su vez, la información relevada, articulada y complementada con otros datos, puede orientar el diseño de intervenciones que avancen hacia la recuperación y mejora de los aprendizajes. Asimismo, las características que los dispositivos de evaluación asuman serán importantes para favorecer su implementación, y generar información válida y confiable que se constituya como un insumo más que pueda ser utilizado para informar la toma de decisiones de políticas de prácticas educativas en general y, muy especialmente, de programas de recuperación de aprendizajes (ej. remedial).

Esta primera sección analiza el rol de la aplicación, suspensión o modificación de los **dispositivos nacionales**, regionales e internacionales de evaluación educativa

a gran escala durante la pandemia, en 2020 y 2021. Se estructura en dos subsecciones. La primera centra su mirada en el análisis de los dispositivos nacionales de evaluación de los aprendizajes, a partir de un relevamiento de información basado en fuentes oficiales (ver Anexo 1). La segunda sección se centra en los dispositivos de evaluación de los aprendizajes a **nivel regional e internacional**, implementados en los países de la región (ver Anexo 2). Tanto dentro de los dispositivos de evaluación a nivel nacional como a nivel regional, se distinguen:

- **Los exámenes con fines de certificación** en el nivel primario o secundario **y los exámenes con fines de selección o admisión** a la educación superior, donde agentes de evaluación indican que los estudiantes han adquirido los conocimientos y competencias estipulados para dicho año de estudio o nivel, y por lo tanto pueden continuar o acceder al siguiente.
- **Las evaluaciones con fines de monitoreo**, a través de las cuales se busca analizar la información de los sistemas educativos en su conjunto, y cuyos resultados pueden contribuir a la reflexión sobre los logros y desafíos en la adquisición de los aprendizajes.

A. DISPOSITIVOS NACIONALES

El análisis presentado a continuación se basa en un relevamiento realizado sobre veinticinco (25) dispositivos nacionales de evaluación de los aprendizajes³ en veinte (20) países de la región⁴ (ver Anexo 1).

APLICACIÓN DE EVALUACIONES NACIONALES EN 2020 Y 2021

En el año 2020, la mayoría de los países de la región decidieron posponer o suspender los dispositivos nacionales de evaluación de los aprendizajes. Del universo de veinticinco (25) evaluaciones relevadas, veintidós (22) de ellas correspondía que fueran implementadas en dicho año dada su frecuencia de aplicación⁵.

Las evaluaciones efectivamente aplicadas fueron solo siete (7/22), que representan aproximadamente el 32% de los dispositivos relevados, mientras que las quince (15/22) restantes, que constituyen el 68%, fueron suspendidas o pospuestas (ver Mapa 1). Para visualizar la distribución y el detalle de las evaluaciones según su aplicación en 2020, ver el Anexo 3.

MAPA 1: DECISIONES ADOPTADAS POR LOS PAÍSES DE LA REGIÓN RESPECTO A LA APLICACIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN EN 2020

Fuente: Elaboración propia con base en las fuentes de información oficiales de los Ministerios de Educación.



Al año siguiente, sin embargo, se observa una inversión entre el porcentaje de dispositivos nacionales de evaluación suspendidos o pospuestos, frente a aquellos que efectivamente fueron implementados.

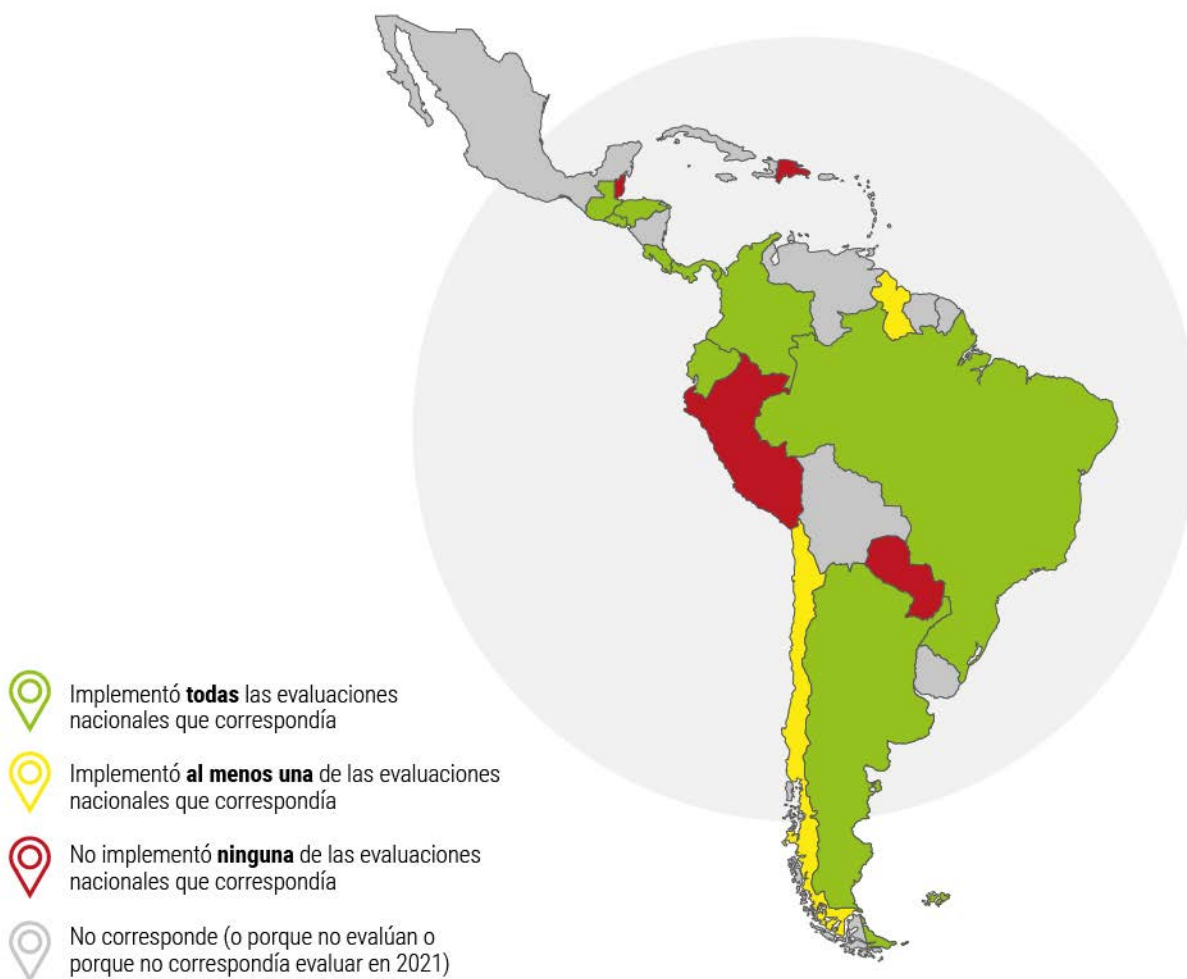
En 2021, del total de las veinticinco (25) evaluaciones nacionales relevadas, solamente una (1) no correspondía ser implementada dada su frecuencia ⁶. De las veinticuatro (24) restantes, nueve (9/24) fueron suspendidas o postergadas, lo cual representa el 37% del total de las evaluaciones que correspondía aplicar en dicho año. Las quince (15/24) restantes se aplicaron y representan el 63% de los dispositivos nacionales que correspondía fuesen

aplicados en dicho año (ver Mapa 2). Para visualizar la distribución y el detalle de las evaluaciones según su aplicación en 2021, dirigirse al Anexo 4.

Del total de 16 países que correspondía que aplicaran al menos 1 dispositivo nacional de evaluación de los aprendizajes en 2020 y en 2021, 3 de ellos (19%) no logró implementar ninguna de las evaluaciones nacionales durante la pandemia ⁷.

MAPA 2: DECISIONES ADOPTADAS POR LOS PAÍSES DE LA REGIÓN RESPECTO A LA APLICACIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN EN 2021

Fuente: Elaboración propia con base en las fuentes de información oficiales de los Ministerios de Educación.



PRINCIPALES TENDENCIAS OBSERVADAS EN LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN DURANTE LA PANDEMIA

A partir del análisis preliminar de la información relevada se identifican tres tendencias que se relacionan, por un lado, a los factores que influyeron en la implementación (o suspensión) de los dispositivos en cada año. Por otro lado, se observan puntos de encuentro tanto en los elementos de las evaluaciones en los cuales se han introducido (o se prevén) cambios e innovaciones, como en aquellos que se han mantenido sin modificaciones.

En primer lugar, se identifica una tendencia relacionada con la finalidad de las evaluaciones nacionales. En segunda instancia, se registra una tendencia relacionada con las continuidades y rupturas en las características de los dispositivos. Por último, se examina una tercera tendencia relacionada con los motivos compartidos entre aquellas evaluaciones que no fueron aplicadas en 2020 ni en 2021.

Para cada una de las tendencias observadas, se aporta un caso de estudio que permite observar con mayor profundidad cómo se comporta la tendencia en un contexto específico.

Primera tendencia: en 2020 se sostuvieron las evaluaciones nacionales con fines de certificación y admisión/selección, y se pospusieron o suspendieron las evaluaciones con fines de monitoreo, para no afectar las trayectorias educativas estudiantiles.

En el apartado anterior se caracterizó la distribución del total de los dispositivos nacionales de evaluación, evidenciando una amplia diferencia entre el número de evaluaciones aplicadas en 2020 (alrededor del 32%) en comparación con el año siguiente (63%).

Esta primera tendencia indaga, dentro del conjunto de evaluaciones implementadas en cada año, ¿cuál fue la proporción de dispositivos con fines de certificación, admisión o selección -de altas implicancias- respecto a aquellos con fines de monitoreo de los aprendizajes -en su mayoría de bajas implicancias-?

A partir del análisis de la información, es posible observar que, en el año 2020, caracterizado por el cierre de escuelas

en gran parte de los sistemas educativos de la región, la tendencia fue posponer o suspender las evaluaciones con fines de monitoreo e intentar sostener aquellas de altas implicancias que afectan directamente la continuidad de las trayectorias escolares en el sistema de educación formal (tanto en la certificación de los estudios como en la admisión al siguiente nivel educativo).

En el año 2020, del conjunto de evaluaciones implementadas en la región, aquellas con fines de certificación y admisión/selección representaba el 86% (6/7), mientras que el conjunto de dispositivos de monitoreo comprendía el 14% restante (1/7) (ver Gráfico 1). Para el 2021, sin embargo, que fue un año en gran medida caracterizado por la reapertura parcial de escuelas, la diferencia entre los dos grupos de evaluaciones -de certificación/admisión y de monitoreo- se reduce, siendo el 53% (8/15) de los dispositivos de evaluación aplicados de certificación y admisión con consecuencias "duras", y el 47% (7/15) con fines de monitoreo (ver gráfico 2).

GRÁFICO 1: DISTRIBUCIÓN DE LAS EVALUACIONES IMPLEMENTADAS EN EL AÑO 2020 EN ALC, SEGÚN EL TIPO DE FINALIDAD DE LAS EVALUACIONES

Fuente: Elaboración propia con base en las fuentes de información oficiales de los Ministerios de Educación.

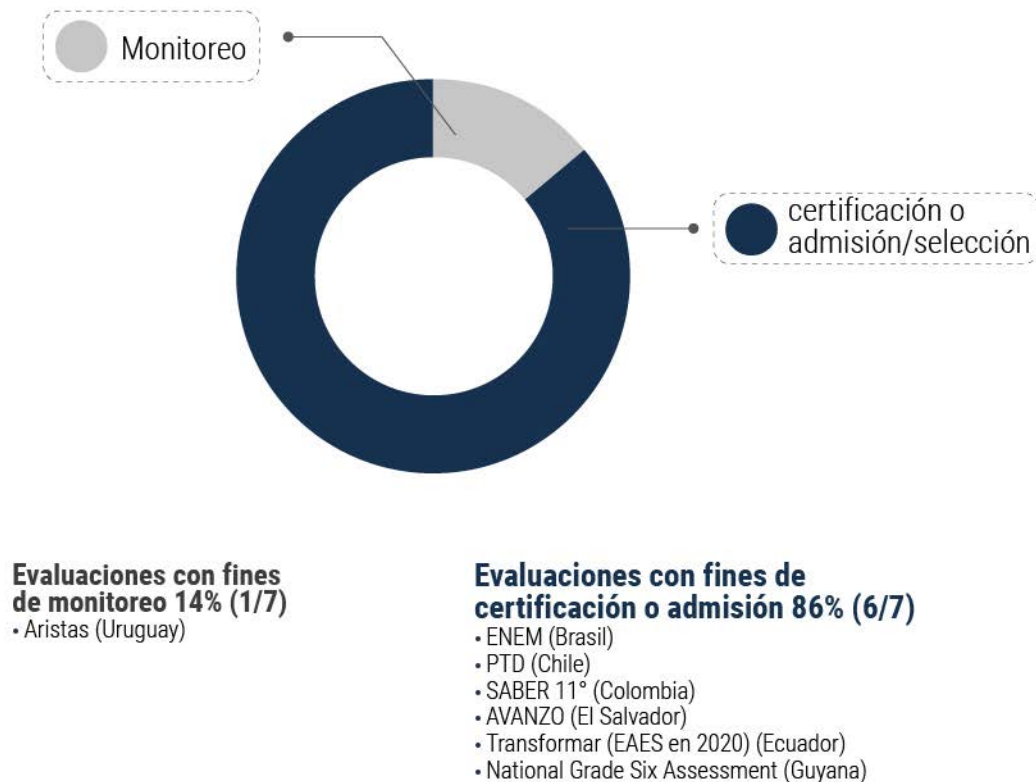
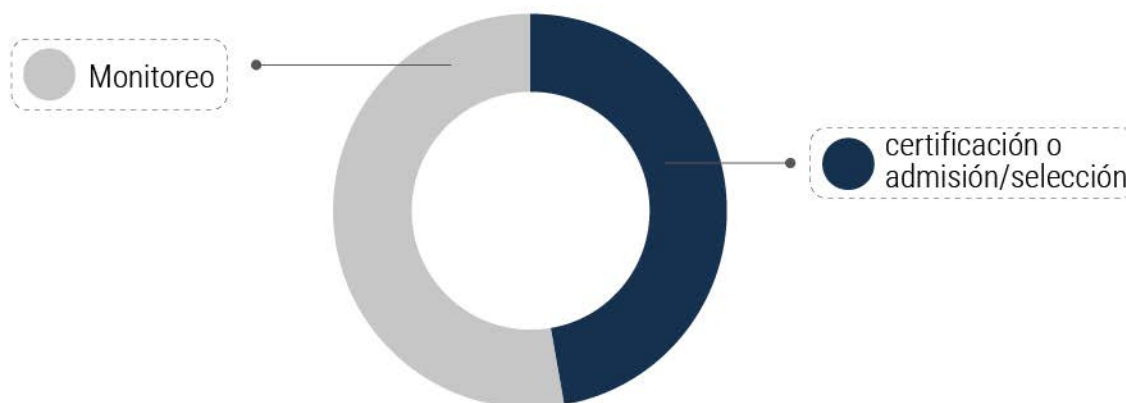


GRÁFICO 2: DISTRIBUCIÓN DE LAS EVALUACIONES IMPLEMENTADAS EN EL AÑO 2021 EN ALC, SEGÚN EL TIPO DE FINALIDAD DE LAS EVALUACIONES

Fuente: Elaboración propia con base en las fuentes de información oficiales de los Ministerios de Educación.



Evaluaciones con fines de monitoreo 53% (8/15)

- Aprender (Argentina)
- SAEB (Brasil)
- SABER 3°5°9° (Colombia)
- Ser Estudiante (Ecuador)
- Prueba para graduandos (Guatemala) *
- Pruebas Diagnósticas (El Salvador)
- Evaluaciones nacionales de desempeño estudiantil (Honduras)
- Crecer (Panamá)

Evaluaciones con fines de certificación o admisión/selección 47% (7/15)

- ENEM (Brasil)
- PDT (Chile)
- SABER 11° (Colombia)
- FARO (Costa Rica)
- Transformar (ex EAES) (Ecuador)
- AVANZO (El Salvador)
- National Grade Six Assessment (Guyana)

* Si bien la evaluación tenía altas implicancias antes de la pandemia, en 2021 el dispositivo tuvo bajas implicancias, con un alcance muestral y la participación fue opcional.

RECUADRO 1. ANÁLISIS DEL CASO DE URUGUAY CON ARISTAS (2020)

Aristas 2020 se presenta como un caso excepcional ya que fue el único de bajas implicancias implementado en el 2020.

Factores relevantes que favorecieron la implementación de la evaluación en el 2020:

Se identifican dos tipos de factores relevantes que, combinados, podrían ser explicativos del caso: aquellos relacionados a las condiciones externas (retorno a la presencialidad y acceso a la tecnología) y aquellos relativos a las características del instrumento (alcance muestral, formato digital y aplicación presencial).

Características del instrumento de evaluación

- Alcance muestral: el dispositivo de evaluación Aristas tiene un alcance muestral, no censal. Esto implica que los estudiantes que participan de la evaluación son un grupo reducido de alumnos de 3° y 6°, seleccionado de forma aleatoria para que la muestra sea representativa del universo total de estudiantes de 3° y 6°. En 2020, del total de estudiantes matriculados, 48.234 en 3° y 47.836 en 6°, entraron en la muestra 5.524 y 5.800 estudiantes de cada año de estudio.

- Formato digital y aplicación presencial: el formato de las pruebas Aristas es digital, no en papel. Esta característica simplifica los aspectos logísticos involucrados en la implementación de estos dispositivos y se combina con el componente presencial de la aplicación, favoreciendo así las condiciones de aplicación.

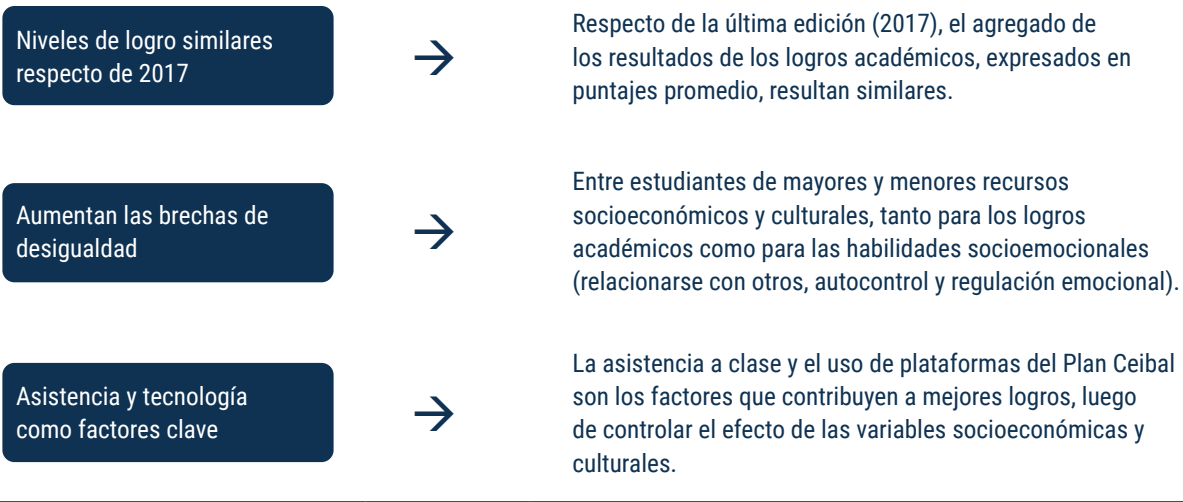
Condiciones externas

- Retorno temprano a la presencialidad: Uruguay fue uno de los países de la región donde hubo una reapertura de escuelas temprana en el 2020, realizándose de forma escalonada. Desde el mes de abril, la enseñanza en las escuelas rurales volvió a ser presencial, y desde el mes de junio se retomaron las clases presenciales en todas las escuelas.
- Amplio acceso a la tecnología: los resultados de la evaluación Aristas 2020 muestran que solamente el 10% de los estudiantes de 3° y 6° que fueron evaluados declaró no contar con computadora o dispositivo electrónico durante la pandemia. Este porcentaje también se encuentra entre los más bajos de la región.

Resultados de Aristas 2020

Considerando que fue el único dispositivo diagnóstico de evaluación implementado en 2020, resulta relevante examinar los resultados obtenidos tanto académicos como socioemocionales.

Este diagnóstico, además de ser el único disponible de acceso público, instala un antecedente para pensar en algunas de las dimensiones que las futuras acciones remediales debieran atender en dicho contexto



Segunda tendencia: tanto en 2020 como en 2021 se introducen algunas modificaciones en el formato de la aplicación de los instrumentos de evaluación y en la información complementaria relevada.

La segunda tendencia también se basa en el conjunto de evaluaciones que fueron implementadas entre el 2020 y el 2021. Se identifican, por un lado, las continuidades,

entendidas como aquellos elementos comunes que se mantuvieron con leves o sin modificaciones en la mayoría de los casos, y por otro, las rupturas o aquellos cambios o innovaciones que fueron introducidos en las aplicaciones de las evaluaciones.

En términos generales, se destaca la prevalencia de las continuidades por sobre las rupturas. Entre los elementos

que se mantuvieron estables, se encuentran: (a) las áreas evaluadas y los destinatarios, y (b) la modalidad de aplicación presencial. Aquí es importante aclarar que, si bien en algunos casos las evaluaciones tomaron como objeto de evaluación o referente los contenidos priorizados para la enseñanza, las áreas generalmente evaluadas en la mayoría de los casos se mantuvieron constantes.

Por otro lado, entre los cambios o innovaciones introducidas se destacan: (a) el relevamiento de información complementaria sobre la enseñanza y el aprendizaje durante la pandemia (por ejemplo:

modalidad de estudio durante la pandemia, frecuencia de cursada, habilidades socioemocionales, entre otras) y (b) sustitución del formato del instrumento en papel hacia digitales o mixtos (digital y papel). Este último elemento tiene la potencialidad de consolidarse en futuras aplicaciones.

A continuación, se presenta un esquema gráfico que sintetiza las continuidades y los cambios e innovaciones en los dispositivos implementados durante la pandemia (ver Diagrama 1).

DIAGRAMA 1. SÍNTESIS DE CONTINUIDADES Y CAMBIOS EN DISTINTOS ELEMENTOS DE LAS EVALUACIONES NACIONALES IMPLEMENTADAS EN 2020 Y 2021

CONTINUIDADES	CAMBIOS E INNOVACIONES
ÁREAS EVALUADAS Y DESTINATARIOS	FORMATO MIXTO O DIGITAL
Únicamente modificadas en Saber 359 (Colombia) donde se introduce Educación Ambiental (aunque no está relacionado al contexto de pandemia) y Aprender (Arg) que en 2021 se iba a aplicar en el Nivel Secundario y se implementó en el Nivel Primario	Se cambia el formato del dispositivo en Enem (Brasil), Saber 359 (Colombia), Prueba para graduandos (Guatemala), Avanzo (ex PAES, El Salvador) y Transformar (Ecuador)
MODALIDAD DE APLICACIÓN PRESENCIAL	INFORMACIÓN COMPLEMENTARIA RELEVADA SOBRE LA ENSEÑANZA Y EL APRENDIZAJE DURANTE LA PANDEMIA
Únicamente modificada en Transformar (Ecuador), donde el dispositivo (que ya se aplicaba en formato digital) se aplicó desde los hogares	Aprender (Argentina), Enem (Brasil), Aristas (Uruguay)

RECUADRO 2. ANÁLISIS DEL CASO DE COLOMBIA CON LAS PRUEBAS SABER 3°, 5°, 9° Y 7°

SABER 3°, 5°, 9° fue la evaluación que realizó más cambios en 2021. Se trata de una evaluación muestral, anual y de bajas implicancias. A continuación, se presentan las dimensiones que sufrieron mayores modificaciones y se detalla el cambio o innovación introducido:

- **Áreas evaluadas:** se introduce una evaluación en educación ambiental, además de Ciencias Naturales.
- **Destinatarios:** se incluye a 7° grado con una prueba piloto.
- **Formato:** digital, utilizando el equipamiento de las escuelas y con provisión adicional en los casos que se requiriese. De todas formas, se mantuvo la aplicación en formato papel de los cuestionarios que relevan información complementaria.
- **Uso de resultados:** se prevé que los resultados de la prueba de 7° se utilicen para realizar el primer estudio de seguimiento de trayectorias escolares. Este implica evaluar y hacer seguimiento del contexto familiar, institucional y social del estudiante a través del tiempo y en el tránsito del nivel de primaria a secundaria. Por este motivo, la aplicación se implementará cada dos años.

Tercera tendencia: en los casos en los que no se aplicó en 2020 ni en 2021, se lo atribuye a motivos sanitarios y en menor medida a motivos técnicos o económicos.

La tercera tendencia examina los motivos comunes por los cuales un grupo de siete (7) dispositivos nacionales de evaluación en cuatro (4) países no fueron implementados en 2020 ni en 2021. Los casos que comprenden este grupo son: Chile con SIMCE; República Dominicana con sus Evaluaciones Nacionales y Evaluaciones Diagnósticas; Perú con la ECE y la EM, y Belice con las PSE y BJAT.

Entre el conjunto de casos que suspendieron la aplicación, en todos se presenta como razón explicativa las

condiciones sanitarias. Además, en Chile, Perú y Belice se presentan motivos técnicos relacionados a la dificultad para garantizar la validez, confiabilidad e incluso la comparabilidad de la información relevada. Estas razones se basan en la prevalencia de la excepcionalidad que supuso la no presencialidad en la modalidad de enseñanza con motivo de la pandemia que, consecuentemente, afectaría el rendimiento estudiantil en las evaluaciones.

Con excepción de Belice, con las pruebas PSE y BJAT, el resto de los casos declaró que las evaluaciones de periodicidad anual serán aplicadas en 2022.

RECUADRO 3. ANÁLISIS DEL CASO DE CHILE CON SIMCE

El dispositivo de evaluación SIMCE es uno de los que no se implementaron en 2020 ni en 2021. Las dos razones que se dieron fueron a) las condiciones sanitarias y b) que si la cobertura no podía garantizarse, los requisitos técnicos de validez y confiabilidad del instrumento no podían asegurarse.

A diferencia de la tendencia entre la mayoría de los sistemas educativos de la región, el caso de Chile es uno donde las evaluaciones nacionales que se aplicaban regularmente perdieron protagonismo frente a la aparición o al fortalecimiento de relevamientos de evaluación alternativos, con otras características. Un ejemplo de esto fue la creación de la plataforma Diagnóstico Integral de Aprendizajes (DIA).

A continuación, se mencionan las principales características de DIA, junto con el crecimiento que tuvieron entre el 2020 y 2021:

La plataforma DIA incluye tres dispositivos de evaluación:

- **Diagnóstico:** cómo llegan los estudiantes al año escolar.
- **Monitoreo intermedio:** cómo avanzan los estudiantes.
- **Evaluación de cierre:** cuál fue el progreso de aprendizaje.



A partir de los resultados DIA 2020 y 2021, la Agencia de la Calidad ofrecerá un nuevo sistema de evaluación para que los establecimientos públicos puedan abordar las brechas de aprendizaje.

APLICACIÓN DE EVALUACIONES NACIONALES PREVISTA PARA EL 2022

A partir de la revisión y consulta a fuentes oficiales de información realizada hasta el mes de enero del 2022, la mayoría de los países prevé la implementación de al menos una o de todos los dispositivos nacionales de evaluación de los aprendizajes (ver Mapa 3).

MAPA 3. PROYECCIÓN INFORMADA POR LOS PAÍSES DE LA REGIÓN RESPECTO A LA APLICACIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN EN 2022

Fuente: Elaboración propia con base en las fuentes de información oficiales de los Ministerios de Educación.



B. DISPOSITIVOS REGIONALES E INTERNACIONALES

APLICACIÓN DE DISPOSITIVOS REGIONALES EN 2020 Y 2021

Durante el 2020 y el 2021 no correspondía aplicar el Estudio Regional Comparativo y Explicativo (ERCE), del Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) de UNESCO, ya que había sido aplicado en 2019. Se trata de una evaluación que mide los logros de aprendizaje de estudiantes que asisten a los sistemas educativos de la región, con fines diagnósticos y de bajas implicancias.

Sin embargo, tanto en 2020 como en 2021 correspondía que se aplicaran las evaluaciones de altas implicancias (también llamadas exámenes) de la región del Caribe, desarrolladas por el Consejo de Examinaciones del Caribe (CXC, por su nombre en Inglés *Caribbean Examinations Council*).

La evaluación CPEA (por sus siglas en inglés, *Caribbean Primary Exit Assessment*) es la única implementada para el nivel primario. Se trata de una prueba que se toma a estudiantes que estén finalizando el nivel primario para certificar sus competencias en Matemática, Lengua, Educación Cívica y Ciencias.

En el nivel secundario, los estudiantes que están finalizando el ciclo básico (generalmente en el tercer año del nivel) y

los adultos que ya lo han concluido participan de la prueba CCSLC, *Caribbean Certificate of Secondary Level Competence*, que evalúa los conocimientos, las competencias generales, actitudes y valores en hasta siete áreas (aunque cada estudiante debe demostrar ser competente en cinco). Los estudiantes que están finalizando el ciclo avanzado del nivel secundario (usualmente en el quinto año del nivel) participan del CSEC, *Caribbean Secondary Education Certificate*, para acreditar que han culminado el nivel. Aquellos estudiantes que desean continuar con estudios superiores, participan además del CAPE, *Caribbean Advanced Proficiency Examinations*).

Todas las evaluaciones del Consejo de Examinaciones del Caribe se aplicaron en 2020 y en 2021, aunque no en todos los países miembros del Consejo. Sin embargo, cada sistema educativo, junto con el CXC tomó decisiones independientes respecto a la implementación de cada una de las evaluaciones, así como la postergación de las fechas. El criterio epidemiológico primó sobre otros motivos a la hora de posponer o cancelar la aplicación de alguna de las evaluaciones en alguna de sus ediciones.

Si bien todas las pruebas tienen fines de acreditación o certificación de los aprendizajes, las consecuencias derivadas del uso de sus resultados son más “duras” en algunas de ellas, que además de certificar o acreditar ciertos aprendizajes o competencias, son utilizadas para la continuación de un trayecto académico o profesional (como es el caso de la CSEC) o la transición del nivel secundario a los estudios superiores (CAPE).

TABLA 1. APLICACIONES DE DISPOSITIVOS REGIONALES EN 2020 Y 2021

Fuente: Elaboración propia en base a las fuentes de información oficiales de los agentes de evaluación LLECE-UNESCO y CXC.

	ERCE	CPEA	CCSLC	CSEC	CAPE
PAÍSES PARTICIPANTES	Argentina, Brasil, Colombia, Costa Rica, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Paraguay, Panamá, República Dominicana, Uruguay y Perú (última edición 2019)	Antigua y Barbuda, Barbados, Belice, Dominica, Granada, Guyana, Jamaica, San Vicente y las Granadinas, San Cristóbal y Nieves, Santa Lucía y Trinidad y Tobago* (participan de una o más exámenes, según el año)			
APLICACIÓN 2020	N/C	Se aplicó en todos los países con excepción de Granada	Se aplicó en todos los países con excepción de Granada	Se aplicó	Se aplicó
APLICACIÓN 2021	N/C	Se aplicó	Se aplicó	Se aplicó	Se aplicó

Notas: N/C: no corresponde.

* Se excluyeron Anguila (UK), las Islas Caimán (UK), Monserrat (UK), las Islas Turcas y Caicos (UK), las Islas Vírgenes Británicas (UK), Saba (Países Bajos), Surinam (Países Bajos) y San Martín (Países Bajos).

APLICACIÓN DE DISPOSITIVOS INTERNACIONALES EN 2020 Y 2021

En 2020 no correspondía aplicar evaluaciones internacionales, según la frecuencia de los dispositivos de evaluación involucrados.

En el 2021 tenían que implementarse PISA y PIRLS. En el caso de PISA, se pospuso para relevar las problemáticas del período post pandemia. No obstante, PIRLS se aplicó de forma presencial y con la posibilidad de optar entre un formato digital o en papel. Además, se lleva a cabo una modificación en la asignación de instrumentos en función de los puntajes obtenidos en aplicaciones previas.

TABLA 2. APLICACIÓN DE DISPOSITIVOS INTERNACIONALES EN 2020 Y 2021, SEGÚN EL AÑO OFICIAL DE LA EDICIÓN DE LA EVALUACIÓN

Fuente: Elaboración propia en base a las fuentes de información oficiales de los agentes de evaluación OECD e IEA.

	PISA	ICCS	PIRLS	TIMSS	ICILS
PAÍSES PARTICIPANTES (ÚLTIMA EDICIÓN)	Brasil, Chile, Colombia, Costa Rica, El Salvador, Guatemala, México, Panamá, Perú, República Dominicana, y Uruguay (2018)	Chile, Colombia, México, Perú y República Dominicana (2016)	Brasil y Chile (2021)	Chile (2019)	Chile y Uruguay (2018)
APLICACIÓN 2020	N/C	N/C	N/C	N/C	N/C
APLICACIÓN 2021	No se aplicó*	N/C	Se aplicó	N/C	N/C

Notas: N/C: no corresponde.

*En 2021 correspondía que se implementase el estudio PISA. Sin embargo, con motivo de la pandemia la prueba fue pospuesta para el 2022. En 2021, los países participantes del hemisferio sur participaron de la evaluación piloto de PISA 2022.

TENDENCIAS IDENTIFICADAS EN LAS EVALUACIONES REGIONALES E INTERNACIONALES

A continuación, se presentan una serie de tendencias identificadas a partir del relevamiento de los dispositivos de evaluación a nivel regional e internacional.

Cabe destacar que, en el caso de las evaluaciones internacionales (PISA, ICCS, PIRLS y TIMSS) se trata de tendencias que son anteriores al período de la pandemia, pero que han sido sostenidas a pesar de ella.

Primera tendencia: *el uso de la tecnología para innovar y fortalecer los instrumentos de evaluación y el procesamiento de resultados.*

Todos los dispositivos de evaluación están tendiendo a incorporar a) un diseño adaptativo según nivel de complejidad de las respuestas del grupo (TIMSS, PIRLS), b) ítems

interactivos (PISA, TIMSS), c) formato digital (PISA, TIMSS, PIRLS - ofrece formato mixto en 2021), y d) inclusión de ítems de respuesta abierta que complejizan los mecanismos de corrección para apreciar de forma más integral las interacciones de los estudiantes con las consignas para su resolución (TIMSS y PISA con las rúbricas).

Segunda tendencia: *la incorporación de nuevos objetos de evaluación en los instrumentos.*

Se destaca la incorporación de nuevos objetos: pensamiento creativo (PISA), valoración sobre tendencias en las prácticas cívicas y de ciudadanía – populismo, autoritarismo y ODS 4 – (ICCS) y tareas de resolución de problemas e investigación (PSIs, por sus siglas en inglés) (TIMSS), y pensamiento computacional (CT), programación y ciudadanía digital (ICILS).

PROYECCIONES PARA FUTURAS EDICIONES

En función de la información disponible y de acceso público de los organismos y agentes que desarrollan las evaluaciones regionales e internacionales, se prevé la aplicación de los dispositivos en los años 2022, 2023 y 2026.

Regionales

- ERCE: se desarrollará un estudio cualitativo en profundidad en algunos países desde 2022. Se prevé el desarrollo de un estudio cualitativo para profundizar en las razones de los resultados ERCE 2019.
- Según las fuentes oficiales disponibles por parte del Consejo de Examinaciones del Caribe (CXC), no se especifican cambios o modificaciones en las próximas aplicaciones de las pruebas CPEA, CCSLC, CSEC o CAPE.

Internacionales

- PISA: se prevé su próxima edición para el 2022, con los resultados dándose a conocer en diciembre de 2023. Se evaluará Pensamiento Creativo (Creative Thinking). Se incluyen ítems abiertos e interactivos con un nuevo diseño visual.
- ICCS: se prevé su próxima aplicación para el 2022. Se prevé ampliar las temáticas relacionadas con el ODS 4 (cambio climático) y tendencias: populismo y autoritarismo.
- PIRLS: se prevé su próxima aplicación para el 2026. Se introduce un nuevo esquema de asignación de pruebas según nivel de complejidad, basado en resultados previos.
- TIMSS: se prevé su próxima aplicación para el 2023. Se espera que las próximas aplicaciones sean en formato digital y con un diseño adaptativo.
- ICILS: se prevé su próxima aplicación para el 2023. Se prevé ampliar las temáticas relacionadas con pensamiento computacional (CT), programación y ciudadanía digital.

C. SÍNTESIS Y REFLEXIONES

La mayoría de los países de la región cuenta con al menos un dispositivo nacional de evaluación con fines de

monitoreo que tienen la potencialidad de identificar los logros y desafíos del sistema educativo para garantizar la adquisición de los aprendizajes. En 2020, sin embargo, aquellos países que también cuentan con exámenes nacionales con fines de certificación y de selección, priorizaron la implementación de estos dispositivos por sobre los anteriores.

A partir del 2021, las evaluaciones con fines de monitoreo volvieron a aplicarse en la mayoría de los países de la región. La información ofrecida por estas evaluaciones aún está siendo analizada y en muchos casos los resultados están pendientes de publicación. Como consecuencia, los datos sobre el impacto de la pandemia en los aprendizajes, y las brechas entre distintos grupos de estudiantes, son todavía escasos e incipientes.

Estas evaluaciones con fines de monitoreo, por su escala y potencial para producir información comparable en el tiempo, de haber sido priorizadas, podrían haber dado cuenta de los efectos de la pandemia en los sistemas educativos, para la planificación de los ciclos lectivos posteriores. La tendencia, sin embargo, pareciera indicar que a partir de 2022 se hará un mayor uso de estas herramientas y, deseablemente, de sus resultados. Como desafío, queda evaluar si y cómo se utilizará la información de estos dispositivos para la planificación educativa de los próximos ciclos lectivos y en adelante.

Los dispositivos nacionales de evaluación se constituyen como una herramienta relevante para el diagnóstico de las pérdidas de aprendizaje durante la pandemia, aunque no suficiente. Se observa una preocupación por desarrollar instancias complementarias de relevamiento de información que permitan explicar “la película” de la “foto” que presentan los resultados de las evaluaciones. Esta tendencia se observa que, tanto a nivel nacional como regional, se están desarrollando o prevé el desarrollo de estudios cualitativos a partir de los resultados que ofrezcan expandir y complementar la información obtenida a partir de las evaluaciones estandarizadas.

Por último, para las evaluaciones internacionales se prevén innovaciones para las próximas ediciones, en su mayoría relacionadas a la inclusión de la tecnología en los distintos procesos involucrados en la evaluación de los aprendizajes. Un nuevo interrogante de cara al futuro es si la situación de la emergencia sanitaria por la pandemia afectará estos dispositivos o si, a partir de ella, se incluirán cambios en los mismos.

2. Evaluaciones formativas: casos ilustrativos y ejes analíticos

Las evaluaciones formativas han ganado relevancia en la última década, convirtiéndose en una pieza crítica del ecosistema educativo en los países de la región. El urgente proceso de recuperación para responder a la pérdida de aprendizajes causada por la pandemia enfatiza la importancia de los diagnósticos formativos. Además, con las demoras en la aplicación de múltiples de los dispositivos nacionales analizados en la sección anterior, las evaluaciones formativas van a tener un rol aún más crítico en los próximos años. Dentro las categoría de diagnósticos formativos, se incluyen tanto evaluaciones diseñadas a nivel central y distribuidas a nivel del aula, como ejercicios de diagnóstico que forman parte de programas más amplios de intervención pedagógica. En este informe, las evaluaciones formativas se caracterizan por su objetivo de relevar datos para informar acciones de recuperación y retroalimentación en el corto plazo, de forma

paródica y regular, con un énfasis en el rol del docente, pero también por parte de las autoridades.

A fin de ilustrar la proliferación de este tipo de esfuerzo y su rol en el marco de recuperación en la post-pandemia, esta sección destaca y analiza nueve casos de evaluación formativa en América Latina y el Caribe (ALC). Los criterios principales para la selección de los casos fueron la representación geográfica, la diversidad de sectores y entidades implementadoras (gubernamentales y de la sociedad civil), su escala, y sus funciones dentro del ecosistema educativo. Más allá de su diversidad geográfica, cuatro de las iniciativas provienen mayoritariamente de gobiernos, tres de sociedad civil, y dos dependen de una cercana colaboración entre entidades gubernamentales y no gubernamentales. Cuatro de las iniciativas se implementaron por primera vez durante la pandemia o como respuesta a esta. Varios de los casos ilustran el uso innovador de sistemas digitales y plataformas de evaluación integradas, mientras que algunos ofrecen ejemplos de pruebas alineadas con enfoques remediales efectivos o cursos de formación docente para la mejor aplicación de metodologías de evaluación diagnóstica (ver tabla 3).

TABLA 3. INFORMACIÓN GENERAL DE LOS CASOS SELECCIONADOS

Fuente: Elaboración propia con información recolectada durante la investigación.

NOMBRE	PAÍS	PRINCIPAL ENTIDAD IMPLEMENTADORA	FECHAS DE IMPLEMENTACIÓN	ESCALA Y ALCANCE (DURANTE 2021)
DIALECT Y DIAMAT, EN DIA+	Chile	Ministerio de Educación ⁸	Primera aplicación en 2021	64,333 niños/as evaluados/as (33,839 en Dialect, 30,833 en Diamat)
EGRA Y EGMA, TUTORÍAS	Colombia	Fundación Carvajal	Desde 2017 y dos veces al año	16,000 niños/as evaluados/as
EVALUACIONES DIAGNÓSTICAS Y FORMATIVAS	Brasil	Ministerio de Educación (MEC) y Centro de Políticas Públicas y Evaluación Educativa (CAEd)	Lanzamiento en 2021	Aún no se han publicado datos oficiales
EVALUACIONES DIAGNÓSTICAS MEJOREDU	México	La Secretaría de Educación Pública (SEP) y la Comisión Nacional para la Mejora Continua de la Educación (MEJOREDU)	Lanzamiento ciclo escolar 2021-2022	516,000 alumnos de más de 2,470 escuelas de enseñanza básica

Continúa en la siguiente página >>.

NOMBRE	PAÍS	PRINCIPAL ENTIDAD IMPLEMENTADORA	FECHAS DE IMPLEMENTACIÓN	ESCALA Y ALCANCE (DURANTE 2021)
MEDICIÓN INDEPENDIENTE DE APRENDIZAJES (MIA)	México	Medición Independiente de Aprendizajes	Desde 2014	300,000 niños/as evaluadas
PAUSAS EVALUATIVAS	Ciudad de Buenos Aires, Argentina	Dirección de Educación Primaria del Ministerio de Educación, gobierno de la Ciudad de Buenos Aires	Desde 2016 y anualmente, menos en 2020	31,400 niños/as de 3º y 6º grado evaluados en matemáticas
PROYECTO CREER	Perú	GRADE	2019-2021 (primera fase), a partir de 2022 (segunda fase e implementación)	Capacitaciones en 40 escuelas multigrado (1-4 docentes por escuela)
PRUEBA DIAGNÓSTICA DE PLATAFORMA IQ	República Dominicana	Fundación Inicia, Ministerio de Educación	2021, aplicación única	30,000 estudiantes evaluados
SISTEMA DE EVALUACIÓN DE APRENDIZAJES (SEA)	Uruguay	Administración Nacional de Educación Pública (ANEP) ⁹	Desde 2011 (de 3º a 6º grado), anualmente. En 2014 se incluyen grados de secundaria.	Abierto a escuelas públicas y privadas ¹⁰

Para analizar y comparar los casos, se definen cuatro ejes clave: el enfoque del contenido de los instrumentos, su alineación con enfoques remediales, el rol del docente en su implementación y uso, y las alianzas y financiamiento del cual dependen. En el análisis presentado a continuación, los ejes permiten clasificar las nueve iniciativas y evaluar su

viabilidad en el contexto regional, así como el potencial para su escalabilidad. Además, dentro de las cuatro dimensiones se identifican recomendaciones y mensajes clave para la agenda de evaluación formativa en la post-pandemia, en base a las experiencias destacadas.

DIAGRAMA 2. LOS CUATRO EJES ANALÍTICOS

Dominio de la evaluación	Alineación con programas de recuperación
¿Está la prueba alineada con aprendizajes esperados o se enfoca en aprendizajes reales?	¿Existen orientaciones generales o un programa de recuperación alineado a los resultados de la prueba?
Rol docente	Alianzas y financiamiento
¿Cuál es el rol de los docentes en el proceso de desarrollo, aplicación, corrección, análisis y uso de la información?	¿De qué alianzas depende el desarrollo, aplicación y la escalabilidad de la prueba?

A. EJE I: DOMINIO DE LA EVALUACIÓN

La efectividad de cualquier diagnóstico de aprendizajes depende de una definición apropiada de los contenidos a evaluar. Por lo general, la evaluación de aprendizajes se ha construido sobre los aprendizajes esperados, producto del currículo vigente en cada país (como en el caso de las evaluaciones estandarizadas), o una síntesis a nivel regional (como en el caso de ERCE). Por otro lado, existen evaluaciones que se generan sin una relación explícita con los currículos, como los instrumentos de PISA, que definen sus propios niveles de logro y habilidades (OECD, 2000), o el caso de las evaluaciones dirigidas por ciudadanos en varios países del sur global (Hevia y Vergara-Lope, 2019).

El enfoque que deberían tener las evaluaciones ha sido una de las principales tensiones en debates educativos, especialmente en referencia a ejercicios de evaluación formativa y diagnóstica. La distinción entre la evaluación de aprendizajes fundamentales versus aprendizajes esperados o curriculares se remonta a antes de la pandemia y está ligada a la discusión sobre la construcción de los currículos (Coll y Martín, 2006). En las últimas décadas, buena cantidad de reformas educativas se centraron en ajustar el currículo, buscando asegurar los

aprendizajes clave, fundamentales o básicos (UNESCO-IBE, 2015). En muchos países, se constataba que el currículo era muy extenso y detallado, dificultando su aplicación real, y generando una situación de rezago de aprendizajes (Pritchett, 2013). Por ello, tener una evaluación más ajustada a los aprendizajes fundamentales, más que al contenido curricular, fue la principal necesidad de instituciones como Pratham, al desarrollar evaluaciones como ASER en India (Banerji, 2013).

La propuesta de evaluar un núcleo de aprendizajes fundamentales, más allá de las propuestas curriculares, tomó fuerza con el desarrollo del ODS 4, y en particular de la meta 4.1.1., la cual obligó a generar estándares mínimos. En la región de ALC, las discusiones llevaron a identificar el nivel II de ERCE como el indicador técnico que pudiera medir estos aprendizajes fundamentales (UNESCO, 2019). Por un lado, algunas de las iniciativas analizadas en este estudio enfatizan la importancia de priorizar la medición de aprendizajes fundamentales en áreas académicas clave con el fin de producir resultados estandarizados, con metas priorizadas e inmediatas, y no limitadas por la clasificación etaria que imponen los currículos nacionales. Los ejercicios alineados a las bases curriculares, por otro lado, pueden llegar a apoyar la labor docente en estructurar sus acciones remediales alrededor de las mallas curriculares que deben cumplir (ver tabla 4).

TABLA 4. DOMINIO DE LAS INICIATIVAS ANALIZADAS

Fuente: Elaboración propia con información recolectada durante la investigación.

	ENFOQUE EN APRENDIZAJES CURRICULARES O ESPERADOS	ENFOQUE EN APRENDIZAJES FUNDAMENTALES O REALES
DIALECT Y DIAMAT, EN DIA+ <i>Chile</i>	Habilidades de lectoescritura y matemáticas de 1º y 2º grado.	
EGRA Y EGMA, TUTORÍAS <i>Colombia</i>		Competencias básicas en lectoescritura y matemáticas, 2º – 5º grado.
EVALUACIONES DIAGNÓSTICAS Y FORMATIVAS MEC-CAED <i>Brasil</i>	Currículo de matemáticas, portugués e inglés hasta 6º grado.	De 6º a 9º grado (secundaria), el enfoque de las pruebas, como el BNCC ¹¹ , son las áreas temáticas organizadas por niveles de logro, y no grados.

	ENFOQUE EN APRENDIZAJES CURRICULARES O ESPERADOS	ENFOQUE EN APRENDIZAJES FUNDAMENTALES O REALES
EVALUACIONES DIAGNÓSTICAS MEJOREDU <i>México</i>	Currículo de lenguaje y matemáticas, por grado, de 2º primaria a 3º secundaria.	
MEDICIÓN INDEPENDIENTE DE APRENDIZAJES (MIA) <i>México</i>		Aprendizajes fundamentales en lectura y matemáticas.
PAUSAS EVALUATIVAS <i>Argentina</i>	Currículo de lenguaje y matemáticas de 3º y 6º grado. Se añade prueba de inglés en 2021.	
PRUEBA DIAGNÓSTICA DE PLATAFORMA IQ <i>República Dominicana</i>	Currículo de lengua y matemáticas, uno en primaria y uno en secundaria (12 reactivos y 20 reactivos, respectivamente).	
SISTEMA DE EVALUACIÓN DE APRENDIZAJES (SEA) <i>Uruguay</i>	Currículo en múltiples áreas de estudio para todos los grados de educación primaria y media.	

Notas: Proyecto CREER no cuenta con un instrumento propio, sino con un curso de capacitación para la aplicación de modelos de evaluación formativa.

Inevitablemente, este debate ha ganado relevancia en la crisis de la pandemia, en el cual la recuperación de aprendizajes va a requerir un ejercicio de nivelación muy significativo, y en muchos casos, independiente o desalineado con los grados o grupos etarios que correspondían antes de la pandemia. Dada la situación, en algunos países hubo discusiones importantes sobre la necesidad de contar con un currículo prioritario, volviéndose central la idea de contar con evaluaciones que pudieran informar respecto de estos aprendizajes priorizados.

Respecto a las iniciativas analizadas, la mayoría de esfuerzos implementados desde gobiernos nacionales o a nivel regional aún buscan medir los logros en relación con los objetivos curriculares predeterminados. En varios de los casos, no se efectuaron cambios significativos sobre el contenido de las pruebas formativas a causa de la pandemia, como con los instrumentos en la plataforma SEA o las Pausas Evaluativas en Buenos Aires. Estas iniciativas siguieron alineadas a los planes de estudios determinados por área y grado. En el caso de Brasil, por ejemplo, se lanzó una nueva plataforma para la evaluación formativa y diagnóstica de los aprendizajes, respondiendo

particularmente a los cambios más recientes del currículo nacional base.

En ciertos contextos, como el de Chile y Colombia, las brechas de acceso y aprendizaje sí forzaron discusiones sobre la adaptación de los objetivos curriculares nacionales. En el caso de Chile, en 2020, las actividades disponibles a través de la plataforma Aprendo en Línea, y las pruebas correspondientes en la plataforma DIA, se adaptaron a un currículo nacional priorizado. En 2021, la plataforma DIA+ empezó a ofrecer ejercicios diagnósticos alineados al plan nacional Leo y Sumo Primero, el cual incluye módulos y actividades para 1º y 2º grado (durante la pandemia, incluidos también en Aprendo en Línea). En el marco de este plan nacional, DIA+ ofrece dos instrumentos más amplios de evaluación formativa, con un foco en las competencias básicas de lectura y matemáticas por grupo etario: el Dialect o el Diamat (UCE, 2021). Aunque el informe analítico automatizado de los resultados se alinea a los objetivos básicos del currículo nacional chileno, el objetivo de las pruebas es producir un diagnóstico temprano de las habilidades en estudiantes de 1º y 2º de básica. Los instrumentos miden en términos de niveles de logro y el Dialect utiliza la escala

Lexile – una escala internacional y estandarizada para clasificar la complejidad de textos que un lector puede leer y comprender. En 2021, la plataforma se hizo disponible para 5,000 escuelas (públicas y subvencionadas) y 64,333 alumnos tomaron al menos una de las pruebas.

En cambio, en el contexto de la pandemia, las iniciativas desarrolladas desde la sociedad civil tienden a enfatizar su foco en aprendizajes fundamentales, con el objetivo de poder abordar las brechas más consecuentes en primera instancia. En general, estas iniciativas también

están asociadas a la medición del impacto de programas pedagógicos específicos. MIA, especialmente, continúa siendo proponente activo de la evaluación diagnóstica de aprendizajes fundamentales, principalmente en matemáticas y lectura. La Fundación Carvajal, similarmente, utiliza una versión adaptada de EGRA y EGMA para identificar los estudiantes que tienen el mayor rezago en lengua y matemáticas y poder incluirlos en su programa de tutorías, paralelo a su currículo primario. EGRA y EGMA permiten un diagnóstico inicial rápido y la opción de medir el progreso nuevamente al final del programa, de manera estandarizada (Devtech, 2021).

RECUADRO 4. LA APUESTA PARA LA EVALUACIÓN DIAGNÓSTICA DE APRENDIZAJES BÁSICOS: EL PROYECTO DE MIA – MÉXICO

El proyecto MIA en México una evaluación dirigida por ciudadanos (*citizen-led assessment*) bajo la responsabilidad del Centro de Investigaciones y Estudios Superiores en Antropología Social – CIESAS y la Universidad Veracruzana.

Las evaluaciones que lleva a cabo este proyecto miden aprendizajes fundamentales (más que esperados) en los dominios de lectura, (fluidez lectora y comprensión inferencial), y aritmética básica (identificación de números, operaciones aritméticas y resolución de problemas).

Estas evaluaciones tienen tres objetivos principales. En primer lugar, medir el rezago de aprendizajes, es decir, la carencia de aprendizajes esperados respecto de la edad y grado escolar de los educandos. En segundo lugar, elaborar diagnósticos para cada niño y niña, buscando identificar el nivel efectivo de aprendizajes con el que cuenta, más allá de su edad, asistencia u otro atributo, y, en tercer lugar, evaluar la efectividad de cursos de recuperación y aceleramiento de los aprendizajes.

Los instrumentos consisten en una prueba de lectura con seis reactivos: sílabas, palabras, enunciados, historia, comprensión I y comprensión II. Y en una prueba de matemáticas, que contiene nueve reactivos: identificación de números de dos dígitos, sumas sin transformación, sumas con transformación, restas sin transformación, restas con transformación, divisiones, problemas con apoyo visual, problema sin apoyo visual y suma de fracciones.

Estos instrumentos se pueden aplicar a niños entre 5-17 años, y se puede aplicar en hogares, escuelas, o centros comunitarios. Debido a la simpleza de los instrumentos, éstos pueden ser aplicados por docentes, tutores, o voluntarios previamente capacitados. Los instrumentos se aplican de manera oral, uno a uno, y se aplica hasta el ítem donde los sujetos pueden responder. Los resultados se recuperan en lápiz y papel o en aplicaciones digitales, identificando el porcentaje de acierto de cada ítem y pudiendo agruparse en sumatorias para fines estadísticos.

Antes de la pandemia, estas evaluaciones se aplicaron a casi 20,000 niñas y niños en 6 estados del sureste mexicano (Veracruz, Tabasco, Campeche, Yucatán, Quintana Roo y Puebla). En 2020, los instrumentos se aplicaron en varios estados del sureste para medir el rezago de aprendizajes (Velásquez-Durán, Vergara-Lope, and Hevia 2021). Gracias a estas mediciones, se pudo estimar la pérdida de aprendizajes en México producto de la pandemia en dos estados del sureste mexicano (Hevia et al. 2022).

Al inicio del ciclo escolar 2021-2022, los docentes aplicaron estos instrumentos en zonas rurales de Tabasco, a cerca de 34,000 niños entre 3º y 6º de primaria, y en Nuevo León a cerca de 289,000 niños entre 2º y 4º de primaria. Además, se aplicaron a cerca de 2,500 participantes de campamentos de aprendizaje en Guerrero.

En la práctica, y especialmente en el contexto de recuperación, el enfoque curricular y el enfoque en aprendizajes básicos no son totalmente contrapuestos. En la mayoría de los casos analizados en esta sección, los enfoques en aprendizajes curriculares incluyen, de algún modo, aprendizajes básicos. La diferencia clave tiene que ver con priorización de ciertos aprendizajes por encima de los objetivos curriculares. Es decir, algunas evaluaciones están centradas en aprendizajes esperados según el currículo nacional, mientras que otras se enfocan en evaluar los aprendizajes actuales o reales de los estudiantes en el contexto actual. Aunque todas las iniciativas evalúen algún aspecto de los aprendizajes fundamentales, solamente algunas se enfocan específicamente en relevar datos sobre los mismos.

Finalmente, el objetivo de todas las iniciativas es dar información en tiempo real a docentes y/o tutores para enfocar la enseñanza a las realidades de los aprendizajes de cada estudiante. Sin embargo, para informar los esfuerzos de recuperación, la adaptación y priorización de contenidos a evaluar debe considerarse como una práctica esencial, especialmente en las iniciativas gubernamentales. En los casos en los que se ha optado por la priorización curricular y del contenido, el objetivo ha sido apoyar la labor del docente, agilizando la aplicación y facilitando el análisis de los datos para la toma de decisiones inmediatas. Para ciertas iniciativas de la sociedad civil, el uso de evaluaciones formativas para medir aprendizajes fundamentales, y no aprendizajes esperados, ha sido esencial en el marco de recuperación – especialmente en contextos donde las brechas son profundas y ya no corresponden a grupos etarios o grados delineados en los currículos nacionales.

B. EJE II: ALINEACIÓN CON PROGRAMAS DE RECUPERACIÓN Y ACELERAMIENTO

Especialmente en el contexto de recuperación de aprendizajes, el valor de los instrumentos de evaluación formativa depende de su buena alineación con esfuerzos de recuperación de los aprendizajes. El uso guiado y efectivo de los resultados a través de enfoques pedagógicos alineados puede no solamente empoderar al personal educativo para adaptar sus acciones y abordar brechas, sino al mismo tiempo, promover su interés en implementar la prueba misma – especialmente, cuando la evaluación se aplica a través de plataformas digitales nuevas y desconocidas por el docente y los alumnos, en un periodo de alta carga y estrés laboral (Vaillant et al., 2022).

A partir de los casos analizados en este estudio, se observan dos principales estrategias de alineación con esfuerzos remediales (ver tabla 5). Generalmente, las iniciativas gubernamentales o de mayor alcance se enfocan en apoyar un diagnóstico rápido y a escala, alineado con orientaciones prediseñadas o estandarizadas para docentes y/o estudiantes. Los esfuerzos de menor escala, varios por parte de la sociedad civil, generalmente construyen sus ejercicios de evaluación formativa sobre una estrategia pedagógica y remedial ya existente.

TABLA 5. EL ENFOQUE DE RECUPERACIÓN DE LOS CASOS

Fuente: Elaboración propia con información recolectada durante la investigación.

	INICIATIVAS	ENTREGA DE INFORMACIÓN A DOCENTES/ESCUELAS	ACCIONES ESPECÍFICAS DE RECUPERACIÓN
ESFUERZOS GUBERNAMENTALES	DIA+, SEA, Pruebas MEC-CAEd	Docentes/escuelas reciben resultados de manera automatizada, con función de análisis de datos. Devoluciones a nivel de centro educativo, para el uso de los resultados para la formación docente.	Guías orientativas a los docentes y talleres
	Pausas Evaluativas, Pruebas Diagnósticas MEJOREDU	Sistema digital para subir y documentar datos (unidireccional) ¹² , y repositorio de pruebas. Los docentes corrigen y usan las respuestas de los alumnos de manera directa.	Guías orientativas a los docentes.
ESFUERZOS DE SOCIEDAD CIVIL	MIA, EGRA Y EGMA/TUTORÍAS, PROYECTO CREER	Organizaciones recopilan, sistematizan y devuelven la información a los docentes y escuelas.	Tutorías, campamentos con enfoque <i>Teaching at the Right Level</i> , y/o capacitación docentes.

Dentro de los esfuerzos gubernamentales de gran escala, hay un espectro de opciones para producir guías pedagógicas orientativas y alineadas a las pruebas formativas. En los casos de Chile y Uruguay, por ejemplo, las plataformas DIA y SEA cuentan con sistemas digitales integrados que permiten que los estudiantes completen la prueba en la plataforma y automáticamente producen reportes diagnósticos con los resultados. En estas plataformas, los docentes pueden consultar guías orientativas (orientadas con planes curriculares por grado

y materia) para diseñar sus propias acciones remediales. En el caso de DIA+, el sistema también produce rutas personalizadas para estudiantes, las cuales abordan las áreas con mayores deficiencias. La nueva plataforma de evaluaciones del MEC, desarrollada por CAEd, tiene una funcionalidad similar al DIA+ o SEA. Esta plataforma para docentes ofrece una variedad de instrumentos para áreas y grados, y ofrece la oportunidad de mapear los resultados contra los objetivos del BNCC¹³

RECUADRO 5. COMPARABILIDAD Y ALINEACIÓN CURRICULAR: EVALUACIONES DIAGNÓSTICAS MEC - CAED - BRASIL

La nueva plataforma de evaluación en Brasil se desarrolló a partir de una licitación de servicio que abrió el MEC y ganó CAEd. La plataforma se lanzó en 2021, y aunque se empezó a diseñar antes de la pandemia, la crisis aceleró el proceso. CAEd fue responsable por el desarrollo de las pruebas, así como el diseño tecnológico de la plataforma.

El objetivo de la plataforma es poder ofrecer un repositorio nacional de instrumentos de evaluación formativa alineadas con el BNCC y apoyar el uso efectivo de los datos por parte de los docentes. Simultáneamente, al contar con un sistema centralizado, la plataforma es capaz de recolectar datos diagnósticos a tiempo real y generar comparabilidad por niveles (con acceso limitado para las autoridades pertinentes) – dentro de la misma escuela, a nivel de municipio, de estado o nacional.

Uno de los desafíos principales al desarrollar esta plataforma fue crear nuevos instrumentos de evaluación formativa alineados al currículo nacional para la educación secundaria. En los primeros grados, las pruebas se agrupan por año y por asignatura. A partir del 6º grado, el currículum se construye en áreas o temáticas de estudio, y no necesariamente por grados (la estructura está basada en el concepto de teaching at the right level). Para reflejar esa distribución, la plataforma ofrece varias pruebas dentro de cada área, según distintos niveles de logro.

Aunque no ofrece actividades específicas con los resultados, la plataforma es altamente sofisticada y tiene varias funcionalidades para apoyar a los docentes a diseñar sus acciones remediales. Inicialmente, los docentes eligen e imprimen las pruebas, las aplican, e introducen las respuestas de sus alumnos en la plataforma. El sistema corrige las pruebas, analiza y sintetiza los resultados para el uso docente. Una de las funcionalidades más innovadoras de la plataforma es que mapea los resultados de la clase contra los objetivos curriculares, y permite visualizar el progreso de los estudiantes contra su plan de estudio. El sistema también identifica las áreas críticas de apoyo en cada prueba (o competencias clave dentro de cada unidad y asignatura), para alertar a los docentes de los temas que deberían reforzar. La plataforma incluye un repositorio de capacitaciones (módulos con videos) para los docentes.

En sus esfuerzos de promover el uso de esta nueva herramienta, el MEC la ha presentado como una pieza crítica en el ecosistema de evaluación nacional y lo ha introducido como una iniciativa a incluir en proyectos y alianzas ya activas. En general, se espera que el uso de la plataforma crezca en los próximos meses (MEC, 2021).

En el caso de las evaluaciones formativas de MEJOREDU o las Pausas Evaluativas, las plataformas digitales son un recurso del docente y no necesariamente cuentan con sistemas integrados para el análisis o corrección automatizada. Los instrumentos desarrollados por MEJOREDU se pueden aplicar a través de varios medios (impreso, digital, aplicación colectiva), y los docentes reciben guías para identificar las áreas de logro y diseñar secuencias didácticas que mejoren los aprendizajes esperados (MEJOREDU, 2021). El caso de Buenos Aires es similar, pero los docentes tienen el rol adicional de documentar y subir los resultados de las pruebas a la plataforma en línea para uso administrativo a nivel de la dirección educativa de la Ciudad y la toma de decisiones.

El caso de MIA, aunque proviene de la sociedad civil y propone un enfoque en aprendizajes fundamentales más que curriculares, tiene un alcance significativo en los sistemas escolares en diferentes entidades federativas. Los instrumentos desarrollados por MIA se aplican en más de seis estados mexicanos – en algunos casos a través de docentes y en colaboración con gobierno, y en otros dentro

de su propio esfuerzo con la sociedad civil y voluntarios capacitados. En ambos modelos de implementación, uno de los enfoques de MIA son los programas que acompañan las pruebas (ej. campamentos TaRL) y las capacitaciones para docentes sobre esfuerzos remediales (Hevia y Vergara-Lope, 2016). Algunas organizaciones civiles de la región, como Escuela Nueva en Colombia, o fundación Glasswing en El Salvador, están adaptando estos instrumentos en algunas de sus intervenciones de recuperación.

A menor escala y desde la sociedad civil, las iniciativas analizadas incluyen enfoques remediales muy completos y prometedores. El proyecto CREER orienta sus capacitaciones a docentes y los apoya a entender cómo los criterios de evaluación formativa propuestos en su programa se alinean con mapas de competencias y estándares de aprendizaje y progresión, dentro de sus currículos nacionales. En Colombia, la Fundación Carvajal implementa un programa anual de tutorías para estudiantes con mayor rezago, con base en los resultados de una aplicación inicial de EGRA y EGMA.

RECUADRO 6. PROGRAMA DE TUTORÍAS DE LA FUNDACIÓN CARVAJAL - COLOMBIA

Desde el 2017, la Fundación Carvajal implementa su modelo de tutorías en escuelas públicas de Colombia. El programa se diseñó en el marco del proyecto Aula Global y fue evaluado positivamente en un estudio de Harvard University y Vanderbilt University, el cuál midió su impacto sobre 5,000 niños y niñas en 8 escuelas públicas (Fundación Carvajal, 2020).

La intervención empieza con la aplicación de EGMA y EGRA a estudiantes de una aula o escuela. La Fundación Carvajal utiliza una versión de EGMA y EGRA adaptada por la Fundación Luque al castellano y al contexto colombiano. El uso de las pruebas requiere de aplicadores capacitados. Para que el modelo funcione, Fundación Carvajal contrata a un equipo de aplicadores y los capacita con un curso en línea construido para ser completado en una semana. Las pruebas toman unos 25 minutos por niño/a. Los resultados se entregan en porcentajes y permiten identificar a un grupo pequeño de alumnos con el mayor rezago.

En el programa remedial, se imparten 24 tutorías, en sesiones de una hora, dos veces por semana, con 8-10 estudiantes y en jornada escolar. Los tutores que lideran las sesiones también son contratados y capacitados por la Fundación Carvajal. Al final del año, se aplica de nuevo las pruebas EGRA y EGMA para evaluar el progreso del estudiante y el impacto de las tutorías.

Durante la crisis del Covid-19, tanto la aplicación de las pruebas como las tutorías mismas se hicieron en línea y de forma individual. A fin de adaptarse a la modalidad virtual, se rediseñó el modelo y se hicieron unas 12,000 aplicaciones remotas – la mayoría por videollamada de WhatsApp. Además, se hizo un esfuerzo significativo, junto con el Ministerio de Educación Nacional (MEN) de Colombia, para conectar a los estudiantes que no tenían acceso a internet; o en ciertos casos, para imprimir y distribuir las guías necesarias. A finales del 2021, la iniciativa se estaba implementando en las mismas escuelas, remota (30%) y presencialmente (70%). A futuro, la Fundación Carvajal espera mantener la aplicación remota, inicialmente para justificar su inversión en adaptar el modelo, pero también para fortalecer la alfabetización digital y el trabajo autónomo de los alumnos que promueve el aprendizaje a distancia.

El programa de tutorías es solamente una parte del modelo. La intervención de Fundación Carvajal tiene dos ejes: identificar y apoyar a los alumnos con mayor rezago a través de sus tutorías y simultáneamente capacitar a sus maestros con metodologías para *enseñar en niveles* en el aula (*teaching at the right level*, en inglés). La capacitación docente es corta, de 24 horas, y generalmente costo-efectiva. Aparte de asegurar una mejor transición de los estudiantes a sus aulas después de las tutorías, el eje de capacitación también ayuda a conectar con los docentes y promover su interés en participar en la iniciativa.

Dado el éxito de la iniciativa en Colombia, el Banco Interamericano de Desarrollo financió un piloto del modelo en Venezuela (Devtech, 2021). Inicialmente, se ha implementado el curso desarrollado para aplicadores de las pruebas y para tutores.

El nivel de alineación con programas de recuperación específicos depende de las distintas prioridades y objetivos de las pruebas. Para iniciativas que tienen como objetivo crear un diagnóstico rápido y a gran escala, la creación de un programa personalizado, alineado a sus instrumentos, es menos viable. Evidentemente, el costo de programas remediales complejos y personalizados es mucho más alto que la creación de rutas orientativas – tanto a nivel financiero como de capacidad. Según las experiencias analizadas, el gasto principal de los programas de tutorías alineados a evaluaciones formativas es la contratación y capacitación de aplicadores o docentes, e incluso de los mismos formadores. Por ejemplo, para la Fundación Carvajal, el costo promedio por alumno en su programa de tutorías es de 40 a 45 dólares.

Adicionalmente, en el contexto de la pandemia, en muchos sistemas educativos se ha sobrecargado y agotado la capacidad humana existente. En algunos casos, ha supuesto un reto encontrar nuevos formadores para capacitar o incentivar a los docentes para que adopten programas de recuperación adicionales y nuevas evaluaciones formativas. Sin embargo, los enfoques remediales que proponen las iniciativas no-gubernamentales analizados (o no estrictamente curriculares) aportan posibles modelos a explorar para su implementación a escala (ver tabla 6). Es posible desarrollar un modelo remedial sintetizado y alineado, con base en el currículo nacional, y con objetivos y actividades pedagógicas relevantes y específicas a los resultados de las pruebas.

TABLA 6. EJES I Y II, COMPARADOS

Fuente: De creación propia con información recolectada durante la investigación.

	GUÍAS ORIENTATIVAS	PROGRAMAS REMEDIALES ALINEADOS
ENFOQUE EN CURRÍCULO Y PLANES DE ESTUDIO	<ul style="list-style-type: none"> ● Pruebas DIA+ ● Pruebas SEA ● Pruebas MEJOREDU ● Pausas Evaluativas ● Pruebas Diagnósticas MEC - CAEd ● Prueba Plataforma IQ 	
ENFOQUE EN APRENDIZAJES BÁSICOS		<ul style="list-style-type: none"> ● MIA ● EGRA y EGMA/Tutorías (Fundación Carvajal) ● Proyecto CREER

C. EJE III: EL ROL DOCENTE

A lo largo de la crisis, el rol del docente ha sido aún más central en los esfuerzos de medición y recuperación de aprendizajes. Entre sus múltiples tareas, los docentes son generalmente los principales aplicadores de instrumentos de evaluación formativa, y en muchos casos, los usuarios prioritarios de los datos que resultan. Durante la pandemia, también han estado a cargo de documentar la información recolectada para la toma de decisiones administrativas por parte de las autoridades. Las demandas del sistema para los docentes han creado una sobrecarga laboral y con ella, el desafío de incentivar al personal para aplicar nuevos instrumentos de evaluación (y herramientas de enseñanza a distancia) y usar efectivamente los resultados (Vaillant et al., 2022). Por eso, la capacitación y acompañamiento de los docentes en este rol son críticos para asegurar que las estrategias de evaluación formativa cumplan con sus objetivos.

La escala de la iniciativa, al igual que el sector implementador, generalmente determina el nivel de involucramiento del docente en los esfuerzos de

evaluación formativa. Por un lado, en las iniciativas gubernamentales, los docentes han sido actores clave para la aplicación de las pruebas, y en ciertos casos, también en la corrección, análisis y documentación de los resultados en plataformas digitales. En la plataforma de MEC-CAEd, al igual que en la plataforma de SEA o el sistema de MEJOREDU, el docente puede elegir el instrumento que quiere aplicar, en el área y grado de interés. En la plataforma SEA+, existe incluso la opción de construir su propio instrumento con los reactivos disponibles. Las dos pruebas en DIA+ son únicas y predeterminadas, y no ofrecen opciones para elegir. Tanto las pruebas de MEC-CAEd como las Pausas Evaluativas, requieren que el docente imprima los instrumentos y los aplique a sus estudiantes. Esta opción también está disponible con el sistema de MEJOREDU, mientras que la plataforma SEA o DIA+ requiere que las pruebas se completen en línea. Con los instrumentos impresos, generalmente los docentes deben corregirlos o introducir manualmente las respuestas en un sistema en línea el cual permite analizar y producir informes, así como recopilar información a nivel nacional para la toma de decisiones (ver tabla 7).

TABLA 6. EJES I Y II, COMPARADOS

Fuente: De creación propia con información recolectada durante la investigación.

	SELECCIONAN INSTRUMENTO	APLICAN EL INSTRUMENTO (MANUALMENTE)	CORRIGEN LAS RESPUESTAS (MANUALMENTE)	ANALIZAN LOS RESULTADOS (MANUALMENTE)	REPORTAN LOS DATOS EN UN SISTEMA	USAN LOS DATOS PARA CREAR NUEVAS ESTRATEGIAS
DIALECT Y DIAMAT, EN DIA+ <i>Chile</i>	x	x	x	x	x	✓
EGRA Y EGMA, TUTORÍAS <i>Colombia</i>	x	✓ (tutores)	✓ (tutores)	x	x	✓
EVALUACIONES DIAGNÓSTICAS Y FORMATIVAS MEC-CAED <i>Brasil</i>	✓	✓	x	x	x	✓
EVALUACIONES DIAGNÓSTICAS MEJOREDU <i>México</i>	✓	✓	✓	x	✓	✓
MEDICIÓN INDEPENDIENTE DE APRENDIZAJES (MIA) <i>México</i>	x	✓	✓	✓	x	✓
PAUSAS EVALUATIVAS <i>Argentina</i>	x	✓	✓	x	✓	✓
PRUEBA DIAGNÓSTICA DE PLATAFORMA IQ <i>República Dominicana</i>	x	x	x	x	✓	x
SISTEMA DE EVALUACIÓN DE APRENDIZAJES (SEA) <i>Uruguay</i>	✓	x	x	x	✓	✓

Por otro lado, varias de las iniciativas de la sociedad civil incluyen a los docentes en los procesos de evaluación, pero generalmente también cuentan con mayor apoyo técnico adicional. En el caso de Fundación Carvajal y algunos de los proyectos de MIA, tanto la aplicación de las pruebas como la implementación de enfoques remediales dependen del reclutamiento y capacitación de tutores y/o voluntarios, además del apoyo docente. Por un lado, los instrumentos de MIA son intencionalmente sencillos y fáciles de aplicar, lo cual permite que puedan ser usados por distintos actores capacitados, no solamente maestros (Hevia y Vergara-Lope, 2016). Por otro lado, la aplicación de EGMA y EGRA requiere de una capacitación muy específica, lo cual generalmente libera a los docentes de esa tarea. Incluso en las iniciativas que lideran tutores y no docentes, las organizaciones generalmente cuentan con estrategias de acompañamiento directo a los docentes para apoyar el uso de la información que resulte del ejercicio. Aparte de las capacitaciones de MIA

y Fundación Carvajal, el Proyecto CREER se enfoca en el desarrollo de un curso de formación para docentes, con un eje clave sobre evaluación formativa (ver recuadro 7).

En la mayoría de las iniciativas, se considera que, dado su rol, es imperativo acompañar la aplicación de la evaluación diagnóstica con capacitaciones para los docentes y técnicos educativos involucrados. Dentro del espectro de opciones de capacitación, en los casos estudiados se ofrecen desde seminarios orientadores pre-aplicación, hasta formaciones extensas y acompañamiento directo en la implementación de programas remediales. En las experiencias de mayor escala, como SEA, MEJOREDU o DIA, se ofrecen seminarios para orientar a los docentes antes de aplicar las pruebas o para dar uso a la plataforma y a los informes analíticos. La plataforma de MEC-CAEd incluye módulos de capacitación integrados para apoyar a los profesores en la interpretación y uso de resultados (MEC, 2021).

RECUADRO 7. CAPACITACIÓN PARA LA EVALUACIÓN FORMATIVA EFECTIVA: PROYECTO CREER – PERÚ

GRADE, con financiamiento de la Old Dart Foundation, inició la primera fase del Proyecto CREER en 2019. El objetivo principal de este proyecto es fortalecer las capacidades de los docentes y técnicos educativos asociados que trabajan con escuelas rurales, multigrado y monolingüe de Perú, con un enfoque en la enseñanza para la diversidad. Uno de los dos ejes del proyecto consiste en acompañar a los docentes en sus prácticas pedagógicas con un curso de formación sobre evaluación formativa.

La primera fase del proyecto se enfocó en levantar información sobre las necesidades del personal en las escuelas, desarrollar el curso en base a ellas y pilotarlo con un grupo de docentes. La segunda fase, la cual se inició en 2022, tiene como objetivo implementar el curso en 40 escuelas (con un máximo de 4 docentes por escuela) y extenderlo a especialistas y funcionarios de las UGEL Cajabamba y Jaén, así como de la Dirección Regional de Educación de Cajamarca. El proyecto se está ejecutando en el departamento de Cajamarca, específicamente en los distritos de Cachachi (Provincia de Cajabamba) y Bellavista (Provincia de Jaén).

GRADE desarrolló este curso con base en un proyecto previo, en el que asesoraron al Ministerio de Educación con el desarrollo de un nuevo currículo con un énfasis en evaluación formativa y estándares de aprendizajes y progresión. Parte de la asesoría consistía en fortalecer las capacidades de los técnicos ministeriales en evaluación curricular formativa. El curso del Proyecto CREER se alinea al currículo desarrollado a nivel nacional, y busca conectar nuevos modelos de evaluación formativa con los capacidades y estándares de aprendizaje establecidos en las mallas curriculares. Es un curso es demandante – dura un poco más de tres meses e incluye dos sesiones por semana, con reuniones sincrónicas adicionales y actividades asincrónicas.

El contenido del curso busca vincular la enseñanza para la diversidad y heterogeneidad (en cuanto a niveles de logro, y también experiencias individuales) con técnicas de evaluación formativa. Los módulos del curso cubren varias metodologías, como el *desarrollo de criterios de evaluación alineados con el currículo y con las realidades del aula o la creación de vínculos entre la evaluación formativa y sumativa*; y buscan promover un desarrollo experimental de capacidades, con actividades prácticas y diálogo activo.

El objetivo del proyecto a largo plazo es crear un modelo de formación replicable para fortalecer el uso de las evaluaciones formativas. Además de expandir sus alianzas para implementar este curso en más contextos peruanos, actualmente, desde GRADE están trabajando para comprimir el modelo y hacerlo más viable y costo-efectivo.

No todas las iniciativas que dependen de los docentes para su correcta aplicación ofrecen capacitaciones o acompañamiento directo, especialmente las de mayor alcance y escala. Evidentemente, el costo de un curso de formación como el de Proyecto CREER es elevado, y no necesariamente viable a escala nacional o regional. Sin embargo, las capacitaciones o cursos de formación dentro de estos modelos se pueden sintetizar y aplicar en contextos distintos, como se ha hecho en Venezuela con el modelo de Fundación Carvajal.

Más allá de la formación, un desafío básico al que se enfrentan todas las iniciativas (especialmente en el contexto de la pandemia), es el de mostrar a los docentes la utilidad y potencial que tienen estos nuevos instrumentos o plataformas para apoyar su pedagogía y fortalecer los aprendizajes de sus estudiantes. Existe la preocupación de que los docentes suelen resistir una nueva evaluación por sus posibles implicancias en cuanto a su seguridad laboral (si los resultados no son óptimos) y porque les supone cargar con una tarea más sin un beneficio claro para sus prácticas (El Diálogo, 2021b).

En las experiencias analizadas, se proponen varios modelos de incentivos y esfuerzos para motivar a los docentes a participar, los cuales resultan especialmente importantes para iniciativas no-gubernamentales (no obligatorias) o recientemente establecidas. La Fundación Carvajal, por ejemplo, solamente comparte los resultados de la prueba con los maestros del aula y el rector de la escuela, e información agregada y no identificable con el ministerio – lo cual alivia la preocupación sobre implicancias de la prueba. Por otro lado, la Fundación busca crear una relación cercana con los docentes para enfatizar el valor de su inversión de tiempo – a través de una capacitación inicial, un espacio de intercambio, y grupos de evaluación cualitativa para que puedan ofrecer retroalimentación al final del proceso. Incluso en los casos de mayor escala donde no es posible implementar mecanismos de intercambio y retroalimentación cualitativa, es clave desarrollar medidas para corregir posibles desconfianzas y empoderar al docente.

De cierta manera, las diferencias en el rol del docente se deben a las distintas prioridades y objetivos de las pruebas. Para iniciativas como las Pausas Evaluativas y la plataforma MEC-CAEd, generalmente existe claramente un objetivo dual – no solamente el de informar esfuerzos pedagógicos, sino el de recolección de información a nivel nacional para la toma de decisiones. El segundo objetivo depende de la aplicación masiva de la prueba y la recolección de datos representativos e inmediatos, lo cual

limita el tiempo y recursos que pueden dedicarse a dar un acompañamiento cercano a los docentes. Sin un objetivo institucional, las iniciativas de menor escala pueden enfocarse en el uso de las pruebas para el beneficio único del docente y/o los alumnos.

Incluso desde los gobiernos, enmarcar las pruebas como un ejercicio de apoyo docente, en primera instancia, puede llegar a ser crucial para promover el interés del personal. Por ejemplo, el carácter formativo de las pruebas de MEJOREDU se centra en el proceso de retroalimentación positiva y directa a los docentes y estudiantes – es decir, no tienen como objetivo integral la recolección de datos para autoridades, sino para el uso docente (Miranda, 2021). A fin de fomentar un mayor y mejor uso de las pruebas y programas remediales asociados, varias instituciones implementadoras enfatizan la importancia de priorizar el objetivo más importante dentro de estas estrategias – su rol en apoyar y empoderar al docente.

D. EJE IV: ALIANZAS Y FINANCIAMIENTO

En la mayor parte de la región, los estados no contaban con la capacidad técnica o financiera necesaria para desplegar un diagnóstico de aprendizajes con la eficiencia y alcance necesarios para apoyar a docentes y alumnos en sus acciones remediales durante la crisis. Por eso, en los últimos meses se ha observado una proliferación de alianzas y colaboraciones entre el sector público y privado, incluyendo a la sociedad civil (El Diálogo, 2021c). Esta clase de colaboraciones ha beneficiado enormemente al estado, pero también a todos los actores del ecosistema educativo que buscaban ampliar el alcance de sus proyectos. Las entidades no gubernamentales cuentan con mayor flexibilidad y agilidad, experiencia directa en terreno y en ciertos casos, mayor capacidad técnica; mientras que el estado tiene acceso directo y el poder de movilizar al sistema educativo público. A futuro, las alianzas pueden seguir siendo una fuente de innovación y mejoramiento para iniciativas educativas, especialmente esfuerzos de evaluación diagnóstica.

En todos los casos analizados, identificamos modelos de alianzas entre los sectores público y privado. Por un lado, los gobiernos solicitaron el apoyo de organizaciones con mayor capacidad técnica para desarrollar nuevos instrumentos o plataformas digitales para apoyar sus objetivos de evaluación diagnóstica. Las colaboraciones se formaron a través de licitaciones de servicio, contratos de colaboración de largo plazo, o solicitudes puntuales, como en el caso de la plataforma de IQ. En el caso de

Brasil o del DIA+, las autoridades educativas contrataron a centros de investigación y empresas para desarrollar los instrumentos, y diseñar y operar las plataformas – en Chile, se colaboró con Colegium y Universidad de los Andes, mientras que en Brasil se trabajó con CAEd.

A parte de la prestación de servicios técnicos y puntuales en la pandemia, las organizaciones de sociedad civil también han buscado ampliar su acceso al sistema durante la pandemia a través de alianzas con el sector

público, con el fin de atender a las brechas que ha causado la crisis en los contextos más vulnerables. MIA, por ejemplo, opera en seis estados mexicanos y colabora cercanamente con las autoridades en cada uno para aplicar sus instrumentos en las escuelas, y con organizaciones como UNICEF para la implementación de campamentos de aprendizaje. El Proyecto CREER depende de sus relaciones con las UGEL para impartir su curso de formación no solamente con docentes, sino también con técnicos en las direcciones regionales.

RECUADRO 8. APALANCANDO LA CAPACIDAD TÉCNICA EN SOCIEDAD CIVIL: NUEVA PRUEBA NACIONAL DIAGNÓSTICA EN LA PLATAFORMA IQ – REPÚBLICA DOMINICANA

La plataforma IQ nació en 2014 de un proyecto de la Fundación Inicia. Desde entonces, la plataforma se ha ofrecido gratuitamente, como un recurso digital abierto, que apoya a estudiantes de secundaria en su preparación para las pruebas nacionales de acceso a la universidad. Hasta 2020, la plataforma ofrecía principalmente una simulación que evaluaba al usuario contra el currículo de la prueba nacional y en base a los resultados, creaba una ruta personalizada con ejercicios seleccionados para reforzar las áreas en las cuales el alumno había fallado.

Al llegar la pandemia, la función de la plataforma se expandió. En primer lugar, IQ empezó a utilizar la plataforma como un repositorio libre y curado, con todo el contenido alineado a las pruebas que habían desarrollado. También se abrieron espacios interactivos virtuales para apoyar a estudiantes y docentes en tiempo real. En varios de estos proyectos, colaboraron con la Dirección de Evaluación de la Calidad del Ministerio de Educación para apoyar sus esfuerzos de ofrecer clases en línea.

En el año escolar 2020 y 2021, la Dirección buscó desplegar un diagnóstico de aprendizajes y desarrollar pruebas diagnósticas en lengua y matemáticas para primaria y secundaria. Su mayor desafío fue asegurar que la prueba llegara a todos los estudiantes y que pudiera recoger datos rápida y efectivamente. En ese marco, la dirección solicita el uso de la plataforma de IQ para colocar estas pruebas a disposición de los estudiantes al final del año escolar. IQ, quien se ha establecido como un recurso conocido y accesible en los últimos años, apoya este esfuerzo con su capacidad técnica.

Gracias a esta colaboración, 30,000 alumnos fueron evaluados en las dos asignaturas en 2021. El objetivo principal del relevamiento era facilitar la toma de decisiones a nivel nacional. Los docentes también tenían acceso a los datos de las pruebas, pero solamente en visión general, sin detalles desagregados.

La post-pandemia y la agenda de recuperación de aprendizajes presenta una oportunidad única para justificar la inversión continua en modelos innovadores y alianzas público-privadas que se han generado durante la crisis. En los tres ejes anteriores se han identificado varias áreas en las cuales las iniciativas gubernamentales podrían beneficiarse de modelos de evaluación formativa desarrollados por las organizaciones de la sociedad civil:

- En primer lugar, las autoridades educativas pueden valorar la importancia de priorizar el currículo y enfocarse en el diagnóstico de aprendizajes fundamentales en el marco de la post-pandemia. En la sociedad civil, existen varios esfuerzos escalables de evaluación no-curricular, cuyos instrumentos

podrían complementar las estrategias actuales de los gobiernos.

- En segundo lugar, en la sociedad civil y el sector privado en general, desde antes de la pandemia se han diseñado modelos de programas remediales prometedores y probadamente efectivos. Más allá de ofrecer guías orientativas, los gobiernos tienen la oportunidad de adaptar y sintetizar partes de estos modelos para apoyar las acciones de los docentes en el aula durante este período.
- Finalmente, con el fin de seguir involucrando a los docentes en iniciativas de evaluación formativa e incentivando su colaboración, todas las iniciativas podrían beneficiarse de modelos de acompañamiento y formación más efectivos.

E. CONSIDERACIONES PARA LA IMPLEMENTACIÓN Y REPLICABILIDAD DE LOS CASOS

Más allá de la efectividad individual de cada una de estas iniciativas, es importante considerar los factores que pueden facilitar o limitar su implementación y replicabilidad según el contexto y objetivos educativos. El costo estimado de las iniciativas es solamente uno de los factores transversales a considerar, determinado por las características de cada modelo ¹⁴. En base al análisis presentado, se identifican tres características clave: (i) capacitación y contratación de personal, (ii) medio de aplicación y (iii) actividades pedagógicas alineadas.

La **contratación de personal técnico** (tutores o aplicadores), o la capacitación de voluntarios y/o docentes, es uno de los costos principales de cualquier iniciativa de evaluación formativa. Más que en los casos de gobierno, los modelos que se implementan desde la sociedad civil requieren del apoyo adicional de personal contratado, ya que generalmente no se despliegan como una obligación para los docentes dentro del sistema. Además, cuando se utilizan instrumentos complejos (como EGRA o EGMA, Colombia) o de creación propia (como los de MIA, México), las organizaciones deben acompañar a los aplicadores o docentes con las capacitaciones necesarias para asegurar un uso apropiado de los mismos. Las iniciativas que también incorporan programas pedagógicos asociados requieren de capacitaciones aún más comprehensivas o de la contratación de técnicos más preparados, lo cual contribuye a un costo total elevado. Por ejemplo, en el modelo actual de la Fundación Carvajal (Colombia), dada la necesidad de contratación y capacitación de tutores, el costo resulta elevado – de 40-45 USD por alumno.

Los gobiernos, por el contrario, tienen a su disposición las redes de docentes en sus sistemas públicos para implementar iniciativas a escala (como el caso de SEA, MEC-CAEd, DIA+, las Pausas Evaluativas o las pruebas de MEJOREDUE), así como varios puntos de entrada para desplegar las capacitaciones u orientaciones necesarias. En general, las capacitaciones para la implementación de estas evaluaciones son más breves, estandarizadas y menos costosas, y con un foco en entender la interfaz de las plataformas que se utilizan. Sin embargo, varias de las iniciativas gubernamentales podrían beneficiarse de los modelos de capacitación empleados en la sociedad civil, los cuáles dan un acompañamiento mucho más directo y pueden ayudar a garantizar el uso efectivo de los instrumentos. Proyecto CREER (Perú), MIA (México) y la Fundación Carvajal

(Colombia) trabajan para simplificar y flexibilizar sus cursos y capacitaciones con el objetivo que sean rentables y replicables para docentes (o tutores) a escala.

El **medio de aplicación** de las evaluaciones es un segundo factor que puede determinar la viabilidad de su implementación según el contexto. Por ejemplo, en países donde la cobertura de conectividad para la educación no es suficientemente alta, las plataformas en línea no serán tan efectivas. En Chile o Uruguay, donde las tasas de cobertura son entre las más altas de la región, plataformas como SEA o DIA ofrecen la posibilidad de integrar las evaluaciones formativas dentro de un sistema automatizado para dar seguimiento completo a las trayectorias de los estudiantes. En el caso de la plataforma MEC-CAEd en Brasil, la plataforma sirve de repositorio y ofrece pruebas descargables que se imprimen para los estudiantes, para después introducir sus respuestas de nuevo al sistema. Este modelo híbrido, similar al de MEJOREDUE o las Pausas Evaluativas, solamente requiere que los docentes, y no los estudiantes, tengan acceso a dispositivos y a la plataforma.

El desarrollo de una plataforma automatizada en sí, además de una estrategia de conectividad educativa universal y las capacitaciones en TIC que la acompañen, requiere de una inversión inicial significativa. Sin embargo, en algunos contextos, las plataformas de evaluación automatizada pueden ser una buena inversión al largo plazo para contribuir a un ecosistema integrado, como en el caso de MEC-CAEd, Chile o Uruguay. También existe la opción de formar alianzas con plataformas ya desarrolladas. Por ejemplo, en República Dominicana, el Ministerio desarrolló la prueba diagnóstica y contrató el uso puntual de una plataforma ya existente, la plataforma IQ.

Finalmente, dependiendo del objetivo de la estrategia de evaluación, la opción de ofrecer **acciones pedagógicas alineadas** puede ser muy valiosa. Generalmente, las iniciativas gubernamentales analizadas cuentan con orientaciones para que los docentes den uso a los resultados, pero no con acciones específicas para responder a las brechas concretas que presenta cada estudiante. Según el contexto, puede ser preferible adoptar una iniciativa o instrumento a escala con orientaciones generales y estandarizadas, y simultáneamente enfocar los esfuerzos en programas de recuperación alineados para grupos particularmente rezagados. Por ejemplo, los modelos de tutorías de la Fundación Carvajal o de MIA, aunque probablemente más costosos, pueden ser muy viables para abordar las brechas en grupos específicos – particularmente en el marco del proceso de recuperación después de la pandemia.

3. Recomendaciones

Uno de los grandes desafíos que enfrentan todos los países de América Latina y el Caribe es cómo relevar información oportuna y a tiempo real sobre el estado, los avances y retrocesos en la adquisición de aprendizajes que permita informar los procesos de mejora educativa. Pero, sobre todo, que permitan retroalimentar las políticas y prácticas pedagógicas para que las escuelas estén desarrollando en los meses y años venideros para recuperar los aprendizajes perdidos durante el período de pandemia y así avanzar hacia escenarios más resilientes y sostenibles de mejora de la calidad educativa.

A partir del estudio realizado, que incluyó conversaciones informadas con distintos actores de los sistemas educativos de la región, fue posible identificar cuatro mensajes centrales de los cuales se derivan distintos cursos de acción. Inicialmente, se enfatizan las acciones que se pueden llevar a cabo en el corto plazo, con el fin de encarar los desafíos inmediatos de la crisis, y al final, se ofrecen estrategias más amplias que deberán planificarse en el mediano y largo plazo. Desde ya, será necesario considerar y matizar los mensajes en función de los distintos puntos de partida, grados de avance, contexto y posibilidades de cada país.

1. Como primer paso, los ministerios de educación deben priorizar la evaluación diagnóstica (formativas o no) para, después de dos años de pandemia, tener una noción certera y actualizada de las pérdidas de aprendizajes de los estudiantes.

Para apoyar las acciones necesarias en el proceso de recuperación, será imperativo priorizar ejercicios de evaluación diagnóstica, a través de dispositivos nacionales de monitoreo o ejercicios formativos, por encima de evaluaciones con otros fines. Las evaluaciones diagnósticas permiten establecer una línea de base para planificar y medir el impacto de las estrategias de recuperación. En la mayoría de casos, será necesario crear o adaptar instrumentos que respondan a los objetivos del diagnóstico, lo cual puede presentar varios desafíos técnicos. Por ejemplo, el uso de nuevas evaluaciones, con enfoques de contenido adaptados, limita la comparabilidad de los resultados con resultados de evaluaciones previas. Por eso, optar por evaluaciones relevantes al periodo actual puede demorar en el retorno a los dispositivos estandarizados, y con ellos,

información más completa sobre la consolidación de objetivos curriculares. Sin embargo, el costo de no establecer una estrategia de diagnóstico prioritaria puede llegar a ser muy alto.

Es necesario no solo contar con estos diagnósticos, sino utilizarlos para crear datos base de referencia, que permitan priorizar las inversiones en recuperación de aprendizajes y utilizar una perspectiva basada en evidencias para la mejora educativa. Este estudio releva modelos innovadores de diagnóstico rápido y a partir de esas experiencias, identifica características clave para el desarrollo e implementación de un diagnóstico efectivo:

- Los diagnósticos formativos deben priorizar una aplicación y análisis rápidos y accesibles. Tanto los instrumentos como los resultados que relevan deben ser fáciles de acceder, entender y operacionalizar por parte de los docentes y líderes escolares – por ejemplo, no pueden incluir lenguaje altamente técnico.
- Los instrumentos de diagnóstico, formativos o no, deben estar enfocados en aprendizajes fundamentales (o cómo mínimo en un currículo priorizado). Este enfoque permite simplificar y agilizar la aplicación del instrumento y el análisis de los resultados. Además, en el contexto de la crisis, el enfoque puede medir más precisamente las brechas y niveles de logro que seguramente existen fuera de los grados (o objetivos curriculares) que corresponden.

Para el desarrollo y despliegue del diagnóstico, las organizaciones de la sociedad civil y los actores fuera del sector público pueden ser aliados clave para los ministerios. Por un lado, numerosas organizaciones tienen experiencia en el desarrollo de evaluaciones diagnósticas formativas (por ejemplo, el trabajo de MIA, que ha desarrollado un arsenal de instrumentos para determinar el nivel de desempeño de los estudiantes en la lectura y matemáticas) y plataformas para su aplicación (especialmente en contextos difíciles – por ejemplo, la Fundación Carvajal en su aplicación de EGMA y EGRA por WhatsApp o la plataforma IQ en República Dominicana). Por otro lado, los actores fuera del gobierno pueden apoyar con el desarrollo de capacitaciones para el personal educativo y guías para el uso e interpretación de los resultados, y con diseño de materiales didácticos remediales para abordar las brechas identificadas.

2. Es posible y necesario **optimizar la información relevada a través de dispositivos nacionales** (e incluso regionales) aplicados en los últimos dos años para informar acciones de recuperación, aunque no tengan fines diagnósticos específicos.

En los países donde ya está en proceso la aplicación de las evaluaciones nacionales, como pruebas de acceso o progresión que han reanudado o que continuaron aplicándose durante la pandemia, los gobiernos deberían pensar desde ahora en cómo aprovechar la información arrojada para informar la toma de decisiones en el marco de recuperación. En varios de los casos, los datos que se releven serán comparables en el tiempo – incluso con resultados obtenidos antes de la pandemia, lo cual puede ser muy valioso para su análisis y uso.

Aún cuando las evaluaciones no sean de carácter diagnóstico, se debe buscar cómo usar los resultados para informar la respuesta educativa a la pandemia. Por ejemplo, las pruebas de admisión a la universidad (ENEM en Brasil, CAPE en el Caribe), por definición, no son formativas y tampoco apoyan acciones remediales, ya que miden los conocimientos y las competencias de los estudiantes cuando están culminando sus estudios secundarios. Sin embargo, un análisis de los datos con una mirada en los potenciales impactos de la pandemia podría revelar información importante. Si los datos apuntan a una caída en el número de participantes, la implicación podría ser que menos estudiantes están considerando estudios en el nivel superior, y acceder en vez al mundo laboral. En tal caso, entre posibles acciones, una expansión de la oferta de educación profesional podría ayudar a garantizar su inserción exitosa en el trabajo.

Nuevamente, para el uso de estos datos, se enfatiza la necesidad de trabajar rápidamente y con eficiencia. Los actores fuera del gobierno pueden ser colaboradores importantes en este esfuerzo. Por ejemplo, invitar a expertos académicos a colaborar en el análisis de los datos puede agilizar el proceso y asegurar el uso de la información para diseñar acciones inmediatas.

3. Las estrategias de evaluación diagnóstica formativas deben **empoderar a los docentes a través de un involucramiento directo y de capacitaciones relevantes** para garantizar la aplicación y uso efectivo de la información relevada.

Para todos los esfuerzos de evaluación, pero particularmente en las evaluaciones formativas, el rol del docente es clave. No es solamente el éxito de las iniciativas el cual depende del rol activo del docente – si los estados no cuentan con el apoyo y la colaboración de los docentes, ningún esfuerzo de diagnóstico podrá ser exitoso. En este sentido, los docentes (y el personal educativo asociado) tienen que ser colaboradores plenos en todas las conversaciones y procesos, desde el inicio. Las autoridades deben ser receptivas ante sus preocupaciones sobre las implicancias de las evaluaciones y buscar soluciones consensuadas y acordadas conjuntamente.

También, para maximizar el uso de las evaluaciones formativas en el aula, los docentes requieren de los conocimientos y capacidades necesarios para su aplicación e interpretación. Aquí, nuevamente, los actores de la sociedad civil pueden ser aliados importantes, puesto que numerosas organizaciones ya tienen experiencia significativa en este área. El diseño de experiencias de capacitación y acompañamiento será crítico para asegurar que los profesores puedan liderar estos procesos de evaluación y recuperación.

Para asegurar que los docentes utilicen las pruebas para mejorar sus acciones, es clave que no desconfíen de su objetivo. Aunque es importante que los datos informen la toma de decisiones en todos los niveles del sistema, su objetivo principal debería ser empoderar al docente para abordar brechas. Si el énfasis no se pone sobre el rol del docente, los gobiernos se arriesgan a perder su participación por miedo a las implicancias del ejercicio.

4. Finalmente, todos los esfuerzos de evaluación en el marco de la recuperación de la pandemia deben contribuir al desarrollo de **un mejor sistema de evaluaciones de aprendizaje a largo plazo**.

Al diseñar e implementar nuevas estrategias diagnósticas, es imperativo ir consolidando los avances para asegurar cambios más duraderos en el tiempo. A largo plazo, todos los esfuerzos de evaluación deberían estar articulados dentro de un sistema común y con las siguientes características:

- Un sistema **integrado**. En el mediano y largo plazo, va a ser imperativo articular la información generada a partir de múltiples dispositivos, con objetivos y perspectivas varias sobre las dimensiones del sistema educativo, con el fin de responder a interrogantes en distintas escalas y niveles educativos. Para eso, va a ser útil aprovechar la tecnología disponible, y los avances ocurridos durante el período de la pandemia, en el diseño de sistemas interoperables. La mayoría de plataformas recién desarrolladas permiten conectar los niveles

del sistema educativo (llevan datos de relevados en el aula a los ministerios y viceversa), pero también centralizan la información de evaluaciones formativas y sumativas, y crean conexiones relevantes.

- Un sistema **comprensivo**. Para responder a todas las necesidades del ecosistema educativo, el sistema de evaluación debe incorporar múltiples mecanismos de medición con objetivos y enfoques distintos. Aunque las evaluaciones diagnósticas formativas ganan relevancia y valor en el contexto actual, el resto de dispositivos nacionales siguen teniendo un rol necesario dentro de un sistema bien articulado.
- Un sistema **sostenible**. Para mantener el sistema en el tiempo, es importante contar con instituciones o sistemas institucionales con suficiente voluntad política y capacidad técnica. Con ese fin, es estratégico invertir en las sinergias y colaboraciones que se han generado durante el periodo de pandemia entre los gobiernos y actores fuera del sector público, incluyendo a la sociedad civil.

BIBLIOGRAFÍA

- Banerji, Rukmini. 2013. "The Birth of ASER." *Learning Curve* 2013: 85–88.
- Coll, César, and Elena Martín. 2006. *Vigencia del debate curricular. Aprendizajes básicos, competencias y estándares*. México: SEP.
- Devtech. 2021. "Diagnóstico de educación básica en Venezuela: Reporte final". En: https://eneed-venezuela.org/wp-content/uploads/VNZ_Education_Diagnostic_spanish.pdf
- El Diálogo. 2021a. "¿Es usted mi maestra?" El regreso gradual a la escuela en ALC, un período para sanar las profundas heridas dejadas por la pandemia y recuperar, de manera urgente, a los estudiantes y al aprendizaje. PREAL Blog. URL: <https://www.thedialogue.org/blogs/2021/11/es-usted-mi-maestra-el-regreso-gradual-a-la-escuela-en-alc-un-periodo-para-sanar-las-profundas-heridas-dejadas-por-la-pandemia-y-recuperar-de-manera-urgente-a-los-estudiant/?lang=es>
- El Diálogo. 2021b. "El uso de la tecnología para innovar la práctica docente: Retos y lecciones aprendidas en la pandemia". URL: <https://www.thedialogue.org/analysis/el-uso-de-la-tecnologia-para-innovar-la-practica-docente-retos-y-lecciones-aprendidas-en-la-pandemia/?lang=es>
- El Diálogo. 2021c. "Cómo fortalecer la colaboración de la sociedad civil con el sector público: Retos y oportunidades para la colaboración estratégica en la post-pandemia". URL: <https://www.thedialogue.org/analysis/como-fortalecer-la-colaboracion-de-la-sociedad-civil-con-el-sector-publico/?lang=es>
- Fundación Carvajal. 2020. Resultados línea de base Aula Global. (Informe). Santiago de Cali: Área de Educación y Cultura de la Fundación Carvajal.
- GRADE, Proyecto CREER, Perú, en: <http://www.grade.org.pe/creer/el-proyecto/>
- Hevia, Felipe J., and Samana Vergara-Lope. 2019. "Educational Accountability or Social Accountability in Education? Similarities, Tensions, and Differences." Working Paper 5. Washington: Medición Independiente de Aprendizajes; Accountability Research Center. <https://accountabilityresearch.org/publication/educational-accountability-or-social-accountability-in-education-similarities-tensions-and-differences/>.
- Hevia, Felipe J., and Samana Vergara-Lope. 2016. "Evaluaciones Educativas Realizadas Por Ciudadanos En México: Validación de La Medición Independiente de Aprendizajes." *Innovación Educativa* 16 (70): 85–110. <http://www.openaccess.innovacion.ipn.mx/index.php/inovacion/article/viewFile/53/38>.
- Hevia, Felipe J., Samana Vergara-Lope, Anabel Velásquez-Durán, and David Calderón. 2022. "Estimation of the Fundamental Learning Loss and Learning Poverty Related to COVID-19 Pandemic in Mexico." *International Journal of Educational Development* 88 (January): 102515. <https://doi.org/10.1016/j.ijedudev.2021.102515>.
- IIEP-UNESCO, 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. UNESCO IIEP Buenos Aires. Oficina para América Latina, Buenos Aires.
- MEJOREDU. 2021. "Evaluaciones Para La Mejora de Los Aprendizajes. La Nueva Generación de Evaluaciones Diagnósticas y Formativas Para Los Estudiantes de Educación Básica." México: Comisión Nacional para la Mejora Continua de la Educación. <https://www.mejoredu.gob.mx/publicaciones/folleto/evaluaciones-para-la-mejora-de-los-aprendizajes>.
- Ministério da Educação. 2021. "Avaliações Diagnósticas e Formativas." https://www.youtube.com/watch?v=G_4-XieLyJl
- Miranda, Francisco. 2021. "Mejora de Los Aprendizajes: La Evaluación Como Estrategia Para La Mejora de Los Aprendizajes." Presentación en foro Consejo Consultivo MejorEdu presented at the Foro Nacional participación ciudadana para la mejora continua de la educación, México. <https://www.gob.mx/mejoredu/prensa/mejoredu-puso-en-marcha-el-foro-nacional-participacion-ciudadana-para-la-mejora-continua-de-la-educacion-283947?idiom=es>.

- Organisation for Economic Co-operation and Development, ed. 2000. *Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy*. Education and Skills. Paris: Organisation for Economic Co-operation and Development.
- Plataforma del Sistema de Evaluación de Aprendizajes (SEA), Uruguay, en: <https://sea.anep.edu.uy/>
- Plataforma DIA+, Chile, en: <https://diamas.cl/>
- Plataforma IQ, República Dominicana, en: www.iq.edu.do
- Plataforma MEC-CAEd, Brasil, en: <https://plataformadeavaliacaoemmonitoramento.caeddigital.net/>
- Pritchett, Lant. 2013. *The Rebirth of Education: Schooling Ain't Learning*. Washington: Center for Global Development.
- Ravela, P., 2006. Para comprender las evaluaciones educativas. Fichas didácticas. Preal, Santiago de Chile.
- Ravela, P., Arregui, P., Valverde, G., Wolfe, R., Ferrer, G., Rizo, F.M., 2008. Las evaluaciones educativas que América Latina necesita. *Revista Iberoamericana de Evaluación Educativa* 1, 19.
- UIS-UNESCO. 2019. "SDG 4 Data Digest. How to Produce and Use the Global and Thematic Education Indicators". UIS/2019/ED/SD/2. Montreal: UNESCO Institute for Statistics. http://uis.unesco.org/sites/default/files/documents/sdg4-data-digest-2019-en_0.pdf.
- UNESCO -IBE. 2015. *El currículo en los debates y en las reformas educativas al horizonte 2030: para una agenda curricular del siglo XXI*. Ginebra: IBE-UNESCO
- UNRISD. 2019. "Acciones Para Enfrentar La Crisis Global de Aprendizajes En México: El Caso de La Medición Independiente de Aprendizajes (MIA)." UNRISD Think Piece Series. Geneva: UNRISD. [http://www.unrisd.org/80256B3C005BE6B5/\(httpNews\)/E7C1C3A-896576072C12583E5004FC4D8?OpenDocument](http://www.unrisd.org/80256B3C005BE6B5/(httpNews)/E7C1C3A-896576072C12583E5004FC4D8?OpenDocument).
- Vaillant, Denise., Rodríguez-Zidán, Eduardo., Questa-Tortero, Mariela. 2022. Pandemia y percepciones docentes acerca de la enseñanza remota de emergencia: El caso de Uruguay." *Revista Electrónica Educare (Educare Electronic Journal)*, 09-4258 Vol. 26(1) ENERO-ABRIL, 2022: 1-21. <https://www.revistas.una.ac.cr/index.php/EDUCARE/article/view/14542/24098>
- Velásquez-Durán, Anabel, Samana Vergara-Lope, and Felipe J. Hevia. 2021. "Improvement Basic Learning Through the Scaling Up of Educational Innovations." In *Ensuring All Children Learn: Lessons from the South on What Works in Equity and Inclusion*, edited by Ishmael I. Munene, 183–200. Lanham, MD: Lexington Books.
- World Bank. 2021. *Actuemos ya para Proteger el Capital Humano de Nuestros Niños: Los Costos y la Respuesta ante el Impacto de la Pandemia de COVID-19 en el Sector Educativo de América Latina y el Caribe*. World Bank, Washington, DC. © World Bank. <https://openknowledge.worldbank.org/handle/10986/35276> License: CC BY 3.0 IGO.

Anexos

ANEXO 1. LISTADO DE PAÍSES Y DISPOSITIVOS DE EVALUACIÓN NACIONALES INCLUIDOS EN EL RELEVAMIENTO

Fuente: Elaboración propia en base a las fuentes de información oficiales de los agentes de evaluación LLECE-UNESCO, CXC, OECD e IEA.

PAÍS	EVALUACIÓN	AGENTE EVALUADOR	FRECUENCIA	UNIVERSO	FINALIDAD	IMPLICANCIAS
ARGENTINA	Aprender	Secretaría de Evaluación del Ministerio de Educación de la Nación	Anual (bienal por nivel)	Censal	Monitoreo	Bajas (uso diagnóstico de los resultados)
BELICE	Primary School Examination (PSE)	Ministerio de Educación de Belice	Anual	Censal	Certificación en el nivel primario	Altas
	Belize Junior Achievement Test (BJAT)	Ministerio de Educación de Belice	Anual	Censal	Monitoreo	Bajas
BOLIVIA	No se desarrollan dispositivos nacionales de evaluación de los aprendizajes					
BRASIL	Sistema de Evaluación de la Educación Básica (SAEB)	Instituto Nacional de Estudios e Investigaciones Educativas (INEP), Ministerio de Educación	Bienal	Censal	Monitoreo	Bajas (uso diagnóstico de los resultados)
	Examen Nacional de Educación Secundaria (ENEM)	Instituto Nacional de Estudios e Investigaciones Educativas (INEP), Ministerio de Educación	Anual	No aplica	Admisión/selección para la educación superior	Altas
CHILE	SIMCE	Agencia de Calidad de la Educación	Anual	Censal	Monitoreo	Altas (uso de rendición de cuentas)
	Prueba de Transición Universitaria (PDT) ex PSU y desde 2023 Prueba de Acceso	DEMRE – Universidad de Chile	Anual	No aplica	Admisión/selección para la educación superior	Altas
COLOMBIA	Pruebas SABER 3°, 5°, 9°	Instituto Colombiano para la Evaluación de la Educación (ICFES)	Anual	Censal. Por pandemia en 2021 se hace muestral	Monitoreo	Bajas (uso diagnóstico de los resultados)
	Pruebas SABER 11°	Instituto Colombiano para la Evaluación de la Educación -ICFES-	Anual	No aplica	Admisión/selección para la educación superior	Altas

PAÍS	EVALUACIÓN	AGENTE EVALUADOR	FRECUENCIA	UNIVERSO	FINALIDAD	IMPLICANCIAS
COSTA RICA	FARO	Dirección de Gestión y Evaluación de la Calidad (DGECC)	Anual	Censal	Certificación en los niveles primario y secundario	Altas
CUBA	No se desarrollan dispositivos nacionales de evaluación de los aprendizajes					
ECUADOR	Ser Estudiante	Instituto Nacional de Evaluación Educativa (INEVAL)	Anual	Muestral	Monitoreo	Bajas (uso diagnóstico de los resultados)
	Examen Transformar (sustituye a Ser Bachiller EAES)	SENESCYT	Anual	No aplica	Admisión/selección para la educación superior	Altas
EL SALVADOR	Pruebas diagnósticas	Ministerio de Educación, Ciencia y Tecnología (MINEDUCYT)	Bienal (última edición en 2014; 2017-2019 prueba logros)	Censal en 2021, 2008 y 2005. El resto muestral	Monitoreo	Bajas (uso diagnóstico de los resultados)
	AVANZO (SUSTITUYE A LA PAES)	Ministerio de Educación, Ciencia y Tecnología (MINEDUCYT)	Anual	Censal	Certificación en el nivel secundario	Altas
GUATEMALA	Prueba para graduandos	Ministerio de Educación de Guatemala (MINEDUC)	Anual	No corresponde	Certificación en el nivel secundario*	Altas*
GUYANA	National Grade Six Assessment (NGSA)	Ministerio de Educación de Guyana y Caribbean Examinations Council (CXC)	Anual	Censal	Certificación en el nivel primario	Altas
	Grade 2 and 4 and 9 National Assessments	Ministerio de Educación de Guyana y Caribbean Examinations Council (CXC)	Anual	Censal	Certificación en el nivel primario	Altas
HONDURAS	Evaluaciones nacionales de desempeño estudiantil	Secretaría de Educación	Anual (varía el año evaluado)	Censal o muestral	Monitoreo	Bajas (uso diagnóstico de los resultados)

PAÍS	EVALUACIÓN	AGENTE EVALUADOR	FRECUENCIA	UNIVERSO	FINALIDAD	IMPLICANCIAS
MÉXICO	Con anterioridad a la pandemia y por motivos no relacionados a ella, desde el 2020 no estaba previsto la aplicación de evaluaciones estandarizadas (PLANEA SEN y Escuelas hasta el 2019)					
PANAMÁ	Crecer	Dirección Nacional de Evaluación Educativa - Ministerio de Educación	Anual desde 2016	Muestral	Monitoreo	Bajas (uso diagnóstico de los resultados)
PARAGUAY	Sistema Nacional de Evaluación del Proceso Educativo - SNEPE	Instituto Nacional de Evaluación Educativa (INEE),	Trienal	Censal	Monitoreo	Bajas (uso diagnóstico de los resultados)
PERÚ	Evaluación Censal de Estudiantes (ECE)	Oficina de Medición de la Calidad de los Aprendizajes (UMC)	Anual	Censal	Monitoreo	Altas (uso de rendición de cuentas)
	Evaluación Muestral de Estudiantes (EM)	Oficina de Medición de la Calidad de los Aprendizajes (UMC)	2001, 2004, 2013, 2018, 2019	Muestral	Monitoreo	Bajas (uso diagnóstico de los resultados)
REPÚBLICA DOMINICANA	Evaluaciones Nacionales	Dirección de Evaluación de la Calidad	Anual	Censal	Certificación en el nivel secundario**	Altas
	Evaluación Diagnóstica Nacional	Dirección de Evaluación de la Calidad	Anual	Censal (desde 2017)	Monitoreo	Bajas (uso diagnóstico de los resultados)
URUGUAY	Aristas	Instituto Nacional de Evaluación Educativa (INEEd)	Trienal	Muestral	Monitoreo	Bajas (uso diagnóstico de los resultados)
VENEZUELA	No se desarrollan dispositivos nacionales de evaluación de los aprendizajes					

Notas: *En 2021 el dispositivo tuvo un alcance muestral y la participación fue opcional.

**También se aplican para el nivel primario en la modalidad de adultos.

ANEXO 2. LISTADO DE DISPOSITIVOS DE EVALUACIÓN REGIONALES E INTERNACIONALES INCLUIDOS EN EL RELEVAMIENTO.

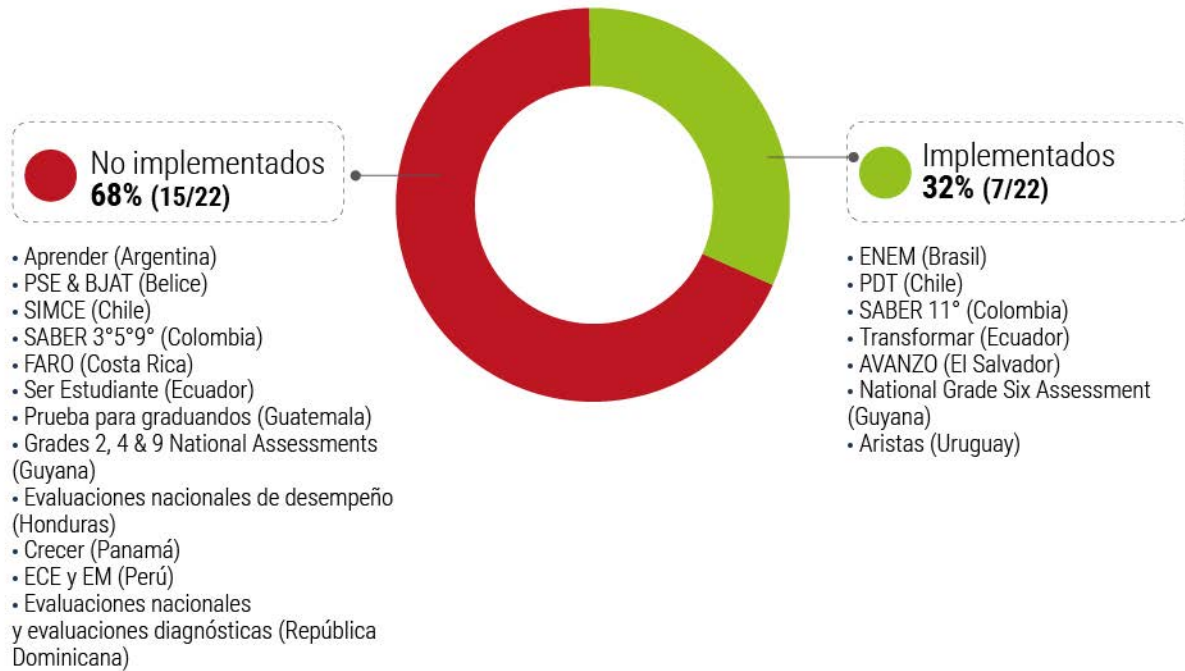
Fuente: Elaboración propia en base a las fuentes de información oficiales de los Ministerios de Educación.

EVALUACIÓN	AGENTE EVALUADOR	ALCANCE	FRECUENCIA	PAÍSES DE LA REGIÓN PARTICIPANTES EN LA ÚLTIMA EDICIÓN
PISA	OCDE	Internacional	Trienal	Argentina, Brasil, Chile, Colombia, Costa Rica, México, Panamá, Perú, República Dominicana y Uruguay (última aplicación 2018)
PISA-D	OCDE	Internacional	Única aplicación (entre el 2014 y 2016)	Ecuador, Guatemala, Honduras, Panamá y Paraguay
ICCS	IEA	Internacional	2009, 2016	Chile, Colombia, México, Perú y República Dominicana (última aplicación 2016)
PIRLS	IEA	Internacional	Quinquenal	Brasil y Chile (última aplicación 2021)
TIMSS	IEA	Internacional	Cuatrienal	Chile (última aplicación 2019)
ERCE	LLECE-UNESCO	Regional	PERCE 1997 SERCE 2006 TERCE 2013 ERCE 2019	Argentina, Brasil, Bolivia, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Honduras, México, Nicaragua, Panamá, Paraguay, Perú, República Dominicana y Uruguay (última aplicación 2019)
CPEA	Caribbean Examinations Council (CXC)	Regional	Anual	Antigua y Barbuda, Barbados, Belice, Dominica, Granada, Guyana, Jamaica, San Vicente y las Granadinas, Surinam, Saba, San Cristóbal y Nieves, Santa Lucía y Trinidad y Tobago*
CCSLC	Caribbean Examinations Council (CXC)	Regional	Anual	Antigua y Barbuda, Barbados, Belice, Dominica, Granada, Guyana, Jamaica, San Vicente y las Granadinas, Surinam, Saba, San Cristóbal y Nieves, Santa Lucía y Trinidad y Tobago*
CSEC	Caribbean Examinations Council (CXC)	Regional	Semestral	Antigua y Barbuda, Barbados, Belice, Dominica, Granada, Guyana, Jamaica, San Vicente y las Granadinas, Surinam, Saba, San Cristóbal y Nieves, Santa Lucía y Trinidad y Tobago*
CAPE	Caribbean Examinations Council (CXC)	Regional	Semestral	Antigua y Barbuda, Barbados, Belice, Dominica, Granada, Guyana, Jamaica, San Vicente y las Granadinas, Surinam, Saba, San Cristóbal y Nieves, Santa Lucía y Trinidad y Tobago*

Notas: * Se excluyen Anguila (UK), las Islas Caimán (UK), Monserrat (UK), las Islas Turcas y Caicos (UK), las Islas Vírgenes Británicas (UK), Saba (Países Bajos), Surinam (Países Bajos) y San Martín (Países Bajos).

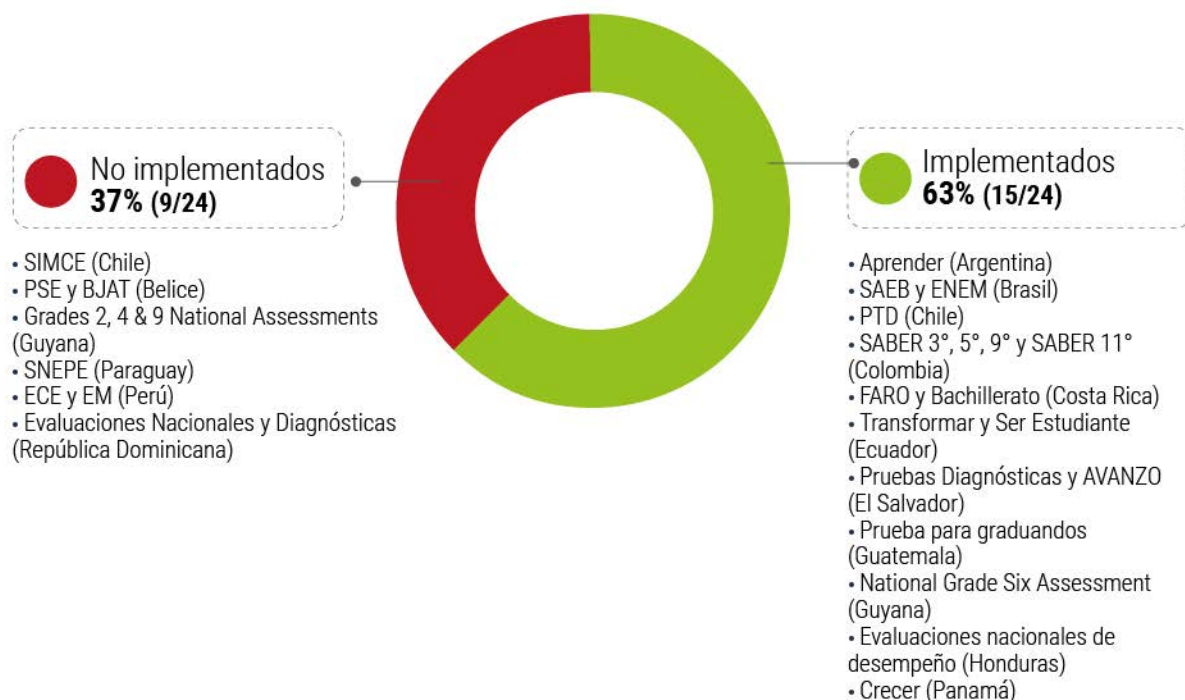
ANEXO 3. DISTRIBUCIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN SEGÚN SU APLICACIÓN EN 2020.

Fuente: Elaboración propia en base a las fuentes de información oficiales de los Ministerios de Educación



ANEXO 4. DISTRIBUCIÓN DE LOS DISPOSITIVOS NACIONALES DE EVALUACIÓN SEGÚN SU APLICACIÓN EN 2021.

Fuente: Elaboración propia en base a las fuentes de información oficiales de los Ministerios de Educación



NOTAS

1. Se excluyen las modalidades especial, de jóvenes y adultos y tecnicaturas o especializaciones en distintas disciplinas (artes, ciencias, entre otras).
2. La distinción entre evaluaciones de altas implicancias, o de alto impacto, y evaluaciones de bajas implicancias o de bajo impacto aparece en la literatura con distintas clasificaciones, y está asociada al propósito de las evaluaciones y a las características de las consecuencias “duras” o “blandas” (Ravela, 2006). Las evaluaciones con consecuencias “duras”, son aquellas cuyos resultados afectan directamente a los estudiantes, sus trayectorias educativas o las instituciones si, por ejemplo, tuvieran incentivos asociados a los puntajes obtenidos. Las evaluaciones con consecuencias “blandas”, se caracterizan por relevar información puede tener fines diagnósticos o de monitoreo de los avances o progresos de los aprendizajes de los estudiantes (se excluyen los dispositivos cuyos resultados pueden afectar las trayectorias escolares de los estudiantes o que pueden tener incentivos asignados para el personal educativo o para las instituciones).
3. Aprender (Argentina), PSE y BJAT (Belice), SAEB y ENEM (Brasil), SIMCE y PDT (Chile), SABER 3°5°9° y SABER 11° (Colombia), FARO (Costa Rica), Ser Estudiante y Transformar (Ecuador), Prueba para graduandos (Guatemala), Evaluaciones nacionales de desempeño estudiantil (Honduras), Crecer (Panamá), SNEPE (Paraguay), ECE y EM (Perú), NGSa y Grade 2,4 & 9 Assessments (Guyana), Evaluaciones Nacionales y Evaluación Diagnóstica Nacional (República Dominicana), AVANZO y Pruebas diagnósticas (El Salvador), Aristas (Uruguay).
4. Argentina, Belice, Bolivia, Brasil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Guatemala, Guyana, Honduras, México, Panamá, Paraguay, Perú, República Dominicana, Uruguay y Venezuela.
5. En 2020, tres (3) de los dispositivos nacionales de evaluación no correspondía que fueran aplicados. Estos son: el Sistema de Evaluación de la Educación Básica (SAEB) en Brasil y las Pruebas Diagnósticas en El Salvador, cuya periodicidad es bienal, y el Sistema Nacional de Evaluación del Proceso Educativo (SNEPE) en Paraguay, cuya periodicidad es trienal.
6. En 2021, la evaluación Aristas en Uruguay no se implementó debido a que su frecuencia es trienal y había sido aplicada en el 2020.
7. Se trata de Belice, Perú y República Dominicana. En el caso de Paraguay, en 2020 no correspondía que se aplicara la evaluación SNEPE, por lo que fue excluido del grupo. De todas formas, en los tres casos se han desarrollado estrategias alternativas a los dispositivos nacionales de evaluación previstos.
8. La plataforma y las pruebas fueron desarrolladas por Colegium y Universidad de los Andes.
9. Involucra una coordinación entre la Dirección Sectorial de Planificación Educativa, la División de Investigación, Evaluación y Estadística, y el Departamento de Evaluación de Aprendizajes.
10. No existe información pública sobre el número de estudiantes evaluados.
11. *Base Nacional Comum Curricular* (Base Curricular Nacional Común)
12. En el caso de MEJOREDU, no se relevó suficiente información para verificar si el sistema devuelve los datos recopilados, una vez analizados, a los docentes y/o escuelas.
13. *Base Curricular Nacional Común (Base Nacional Comum Curricular)*
14. Debido a la duración del periodo de investigación, y la metodología basada en revisión de información pública, en la mayoría de los casos no se localizó información validada sobre el costo de las iniciativas destacadas.



Grupo Banco Mundial
1818 H St NW
Washington, DC 20433
Correo electrónico: edutech@worldbank.org
worldbank.org



thedialogue.org

Diálogo Interamericano
1155 15th Street NW, Suite 800
Washington, DC 20005
Tel: +1 202-822-9002
education@thedialogue.org
thedialogue.org/education