

**DATOS DE  
EVALUACIONES DE  
APRENDIZAJE Y  
ENCUESTAS DE  
HABILIDADES**

DESAFÍOS Y SOLUCIONES A FUTURO

FEBRERO 2024

2024 CONFERENCIA SOBRE  
**DATOS Y ESTADÍSTICAS  
DE EDUCACIÓN**

## INTRODUCCIÓN

Hay varias décadas de desarrollo y debate sobre la medición del progreso en los resultados y habilidades del aprendizaje. Este documento analiza y evalúa el estado y las brechas en los marcos y metodologías que ayudan a medir y monitorear los indicadores sobre resultados de aprendizaje y habilidades relacionados con las metas 4.1, 4.4, 4.6 y 4.7 de los ODS (**Cuadro 1**). Su objetivo es identificar áreas que requieren mayor atención y mejora. Más específicamente, el documento pretende hacer una contribución significativa a los debates e iniciativas en curso en este campo, con el objetivo final de establecer una comunidad internacional de práctica que pueda abordar colectivamente los desafíos futuros.

**Cuadro 1. Indicadores del ODS 4 relacionados con los resultados del aprendizaje y las habilidades**

Indicador		Dominio	Definiciones requeridas
4.1.1	Proporción de niños y adolescentes a) en grado 2 o 3, b) al final de la educación primaria y c) al final de la educación secundaria baja que han alcanzado niveles mínimos de competencia en i) lectura y ii) matemáticas, desglosada por sexo	Lectura y matemáticas	Nivel mínimo de competencia Mínimo de calidad procesal
4.2.1	Proporción de niños de 24 a 59 meses cuyo desarrollo es adecuado en cuanto a la salud, el aprendizaje y el bienestar psicosocial, desglosada por sexo	Aprendizaje, salud socioemocional.	Preparación para la escolarización
4.4.2	Porcentaje de jóvenes y adultos que han alcanzado al menos un nivel mínimo de competencia en alfabetización digital	Habilidades de alfabetización digital	Habilidades digitales relevantes para el empleo, el empleo decente y el emprendimiento
4.6.1	Proporción de la población en un grupo de edad determinado que ha alcanzado al menos un nivel fijo de competencia funcional en a) alfabetización y b) aritmética, desglosada por sexo	Lectura y matemáticas	Nivel fijo de matemática funcional y alfabetización.
4.7.4	Porcentaje de alumnos por grupo de edad (o nivel de educación) que demuestran una comprensión adecuada de las cuestiones relacionadas con la ciudadanía mundial y la sostenibilidad	Ciudadanía global y sostenibilidad	Comprensión adecuada de la ciudadanía global y la sostenibilidad
4.7.5	Porcentaje de alumnos en el último grado de la educación secundaria baja que demuestran dominio del conocimiento de ciencias de la Tierra y ciencias ambientales	Ciencias ambientales y geociencias.	Competencia en ciencias ambientales y conocimientos de geociencias.

## EVALUACIÓN DEL ESTADO DE MEDICIÓN DEL INDICADOR 4.1.1 DE LOS ODS

El indicador 4.1.1 de los ODS es la proporción de niños y jóvenes: (a) en los grados 2 o 3; (b) al final de la educación primaria; y (c) al final de la educación secundaria baja que han alcanzado niveles mínimos de competencia (MPL *por sus siglas en inglés*) en (i) lectura y (ii) matemáticas, por sexo. Se refiere a tres niveles de escolarización (primaria baja, primaria alta y secundaria baja) y dos materias (lectura y matemáticas).

El formato de presentación de informes del indicador tiene como objetivo comunicar dos datos:

- El porcentaje de estudiantes que alcanzan al menos un nivel mínimo de competencia (MPL) para los dominios relevantes (matemáticas y lectura) para cada punto de medición (grados 2/3; final de primaria y final de secundaria baja).
- Las condiciones bajo las cuales el porcentaje de niños que alcanzan o superan el nivel mínimo de competencia puede considerarse comparable al porcentaje informado en otro país.

El indicador requiere los siguientes insumos:

- Dominios: lectura y matemáticas.
- Nivel mínimo de competencia (MPL): Conocimientos básicos de referencia en los ámbitos a una edad/grado determinados.
- Muestra: debe ser representativa a nivel nacional.
- Procedimientos: deben cumplir con estándares mínimos de calidad.

### Desafíos

Hay algunas cuestiones críticas al informar sobre el indicador 4.1.1.

#### *Comparabilidad de grados y niveles educativos*

Debido a que la educación primaria tiene una duración diferente en diferentes países, términos como "fin de la primaria" pueden significar cosas diferentes en diferentes lugares, lo que

complica las comparaciones entre países y programas de evaluación. Sin embargo, el 89% de los países terminan su ciclo primario en los grados 5 a 7, por lo que el problema podría ser menor.

#### *Comparabilidad de los resultados de la evaluación en el espacio y el tiempo*

Si bien la comparabilidad de las estadísticas entre países influye en la comparabilidad a lo largo del tiempo, esta última no implica lo primero.

- La comparación entre países mediante evaluaciones transnacionales ayuda a la comparabilidad entre países, en un momento dado. Si cada programa de evaluación produce estadísticas que sean comparables en el tiempo, entonces las estadísticas serán comparables en el tiempo y en los países.
- Los programas nacionales de evaluación no son comparables entre sí por diseño, pero aun así pueden proporcionar datos de tendencias relativamente confiables si la calidad de las mediciones es lo suficientemente buena.

#### *Oportunidad e impacto político de las estadísticas*

Las evaluaciones producen estadísticas nacionales, y a menudo subnacionales, que pueden influir positivamente en la formulación y aplicación de políticas. Para que estos impactos positivos se sientan, las estadísticas no sólo deben ser precisas sino que también deben ser vistas como creíbles, y el tiempo de respuesta entre la evaluación y la presentación de resultados debe ser lo más corto posible.

#### *Calidad procesal*

Operaciones y procedimientos sólidos y consistentes son una parte esencial de cualquier evaluación a gran escala, para maximizar la calidad de los datos y minimizar el impacto de la variación de los procedimientos en los resultados. Se pueden encontrar ejemplos de estándares procesales en todas las evaluaciones internacionales a gran escala, y en muchas evaluaciones a gran escala a nivel regional, donde el objetivo es establecer coherencia procesal entre países. Muchas evaluaciones nacionales también establecen directrices de procedimiento claras para respaldar la coherencia en su puesta en práctica.

La implementación de la evaluación enfrenta muchas decisiones metodológicas, incluidos formatos de prueba y decisiones de muestreo. No hay necesidad de procedimientos y formatos

idénticos en todas las evaluaciones. Sin embargo, se necesita un conjunto mínimo de procedimientos (alineamiento en los procedimientos) para que la integridad de los datos esté protegida y los resultados sean confiables y razonablemente comparables para cualquier país determinado a lo largo del tiempo, así como entre países en cualquier momento dado.

#### *Costos financieros de las evaluaciones para los países*

Las evaluaciones son relativamente costosas. Sin embargo, incluso para los países en desarrollo, el costo de evaluar sistemáticamente los resultados es extremadamente bajo en relación con el costo general de proporcionar educación y el costo de no medir los resultados del aprendizaje y las habilidades de la población.<sup>1</sup>

#### *Baja cobertura de evaluaciones transnacionales en países de ingresos bajos y medianos bajos*

El indicador 4.1.1 de los ODS se informa utilizando varios estudios transnacionales que son internacionales ([PIRLS](#), [TIMSS](#)) o regionales ([PILNA](#), [SEA-PLM](#), [PASEC](#), [LLECE](#), [SACMEQ](#)) (**Cuadro 2**). Estas herramientas no han sido diseñadas para informar sobre los ODS, pero, en 2018, la Alianza Mundial para el Monitoreo del Aprendizaje ([GAML por sus siglas en inglés](#)) y el Grupo de Cooperación Técnica sobre los indicadores del ODS 4 ([TCG por sus siglas en inglés](#)) acordaron que estas evaluaciones podrían usarse para informar el aprendizaje en función de sus niveles de competencia que mejor se 'mapea' con el MPL global.

---

<sup>1</sup>Para obtener más información sobre costos, consulte UIS (2023). *Reporting learning outcomes in basic education: Country's options for indicator 4.1.1*: [https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Countrys-reporting-option\\_Zambia\\_2023.05.15\\_FINAL.pdf](https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Countrys-reporting-option_Zambia_2023.05.15_FINAL.pdf).

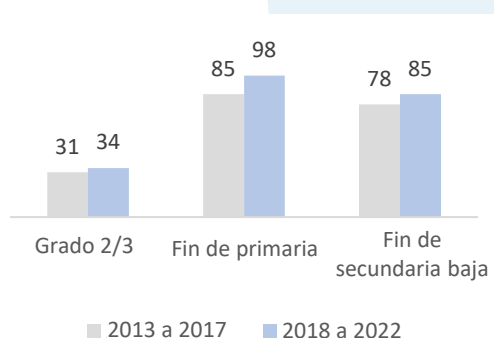
**Cuadro 2. Programas de evaluación por grado o edad y uso para informar sobre el indicador 4.1.1 de los ODS**

Grado/edad	Programa de evaluación internacional
ODS 4.1.1a: Primeros grados	
2	<a href="#">PASEC</a>
3	<a href="#">ERCE</a> , <a href="#">AMPLa</a>
ODS 4.1.1b: Fin de las primarias	
4	<a href="#">PILNA</a> , <a href="#">LaNA</a> , <a href="#">PIRLS</a> , <a href="#">TIMSS</a>
5	<a href="#">SEA-PLM</a>
6	<a href="#">LaNA</a> , <a href="#">PASEC</a> , <a href="#">PILNA</a> , <a href="#">SACMEQ</a> , <a href="#">ERCE</a> , <a href="#">AMPLb</a>
ODS 4.1.1c: Fin de la educación secundaria baja	
8	<a href="#">TIMSS</a>
Edad: 15 años	<a href="#">PISA</a>

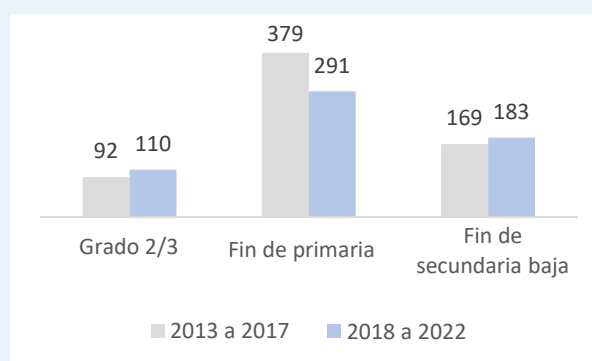
Sin embargo, la producción de resultados de aprendizaje comparables no avanza lo suficientemente rápido y de manera equitativa. Independientemente del criterio de cobertura (el número de países o la población), la cobertura es mucho mayor al final de la primaria y de la secundaria baja que para los grados 2 o 3 (**Figura 1**).

**Figura 1. Cobertura de las evaluaciones de aprendizaje, por nivel de educación**

a. Número de países



b. Población en edad escolar en millones



#### *Evaluaciones nacionales: alineación y calidad procesal*

Si bien los datos de muchas evaluaciones nacionales del aprendizaje están disponibles, cada país establece sus propios estándares, lo que lleva a definiciones inconsistentes de los niveles de desempeño. Por lo tanto, el análisis de los resultados se limita a una sola prueba, metodología y escala. Si bien las metodologías tienden a converger entre las evaluaciones internacionales y regionales, todavía es difícil establecer un nivel de referencia común para las evaluaciones

nacionales y el conjunto de procedimientos utilizados para el muestreo, la gestión de datos y la presentación de informes que también difieren.

El segundo aspecto que deben cumplir los programas de las evaluaciones nacionales es la calidad de los procedimientos. La información sobre los procesos que intervienen en los programas nacionales debe estar disponible públicamente y con suficiente nivel de detalle. El UIS ha proporcionado orientación y herramientas de autoevaluación relacionadas con este aspecto en el documento: [Aligning and reporting on indicator 4.1.1: UIS annotated workflow](#).

### Estándares

Ante estos desafíos, la UIS se ha centrado en definir (i) el nivel mínimo de competencia, alineándolo con un concepto de competencia que sea independiente de un marco de evaluación particular, elementos o pruebas específicos, para permitir la presentación de informes y (ii) un conjunto de estrategias de vinculación al marco de competencia.

#### *Nivel mínimo de competencia*

El nivel mínimo de competencia (MPL) es el punto de referencia del conocimiento básico en un dominio (por ejemplo: matemáticas o lectura) en una edad/grado determinado medido a través de evaluaciones de aprendizaje. El MPL es un punto de referencia para informar sobre las competencias mínimas en cada nivel de escolaridad, sin requerir una prueba para resolver la comparabilidad.

El descriptor del nivel de competencia de MPL es el estándar clave para cada grado y dominio que permite el uso de múltiples evaluaciones para informar el indicador (**Cuadro 3**)<sup>2</sup>. En 2018, se llegó a un acuerdo con los programas de evaluación transnacionales sobre cuál de sus descriptores estaba mejor alineado con el descriptor MPL, aunque es necesaria una mayor

---

<sup>2</sup>Se acordó informar de acuerdo con la definición textual del nivel mínimo de competencia para cada dominio y nivel en las evaluaciones transnacionales. Esto se estableció mediante la realización de un análisis de los descriptores de nivel de desempeño de estas evaluaciones en lectura y matemáticas.

validación, mediante un ejercicio de establecimiento de estándares, ya que los programas de evaluación no han sido diseñados para medir ODS 4.

**Cuadro 3. Niveles mínimos de competencia en lectura y matemáticas para el indicador 4.1.1**

Nivel educativo	Descriptor	
	Lectura	Matemáticas
Grado 2	Los estudiantes leen y comprenden la mayoría de las palabras escritas, particularmente las que le resultan familiares, y extraen información explícita de las oraciones.	Los estudiantes demuestran habilidades en sentido numérico y computación, reconocimiento de formas y orientación espacial.
Grado 3	Los estudiantes leen en voz alta palabras escritas con precisión y fluidez. Entienden el significado general de frases y textos breves.	
Grados 4 a 6	Los estudiantes interpretan y dan algunas explicaciones sobre las ideas principales y secundarias en diferentes tipos de textos. Establecen conexiones entre las ideas principales de un texto y sus experiencias personales, así como conocimientos generales.	Los estudiantes demuestran habilidades en sentido numérico y cálculo, medición básica, lectura, interpretación y construcción de gráficos, orientación espacial y patrones numéricos.
Grados 8 a 9	Los estudiantes establecen conexiones entre las ideas principales de diferentes tipos de texto y las intenciones del autor. Reflexionan y sacan conclusiones a partir del texto.	Los estudiantes demuestran habilidades en computación, problemas de aplicación, comparación de tablas y gráficos y uso de representaciones algebraicas.

*Fuente: Instituto de Estadística de la UNESCO, 2021, p. 4.*

Como parte de este proceso fue necesario:

- Aproximar los niveles de grado de interés para los informes
- Utilizar el programa de evaluación de los niveles de competencia existentes.
- Identificar el descriptor del nivel de competencia que esté mejor alineado con los niveles mínimos de competencia.
- Utilizar ese nivel para informar hasta que finalice el ejercicio de establecimiento de estándares.
- Desarrollar capacidad técnica nacional pero no abordar directamente el desarrollo de evaluaciones nacionales.



### *Marco de competencia global*

El marco de competencia global (GPF por sus siglas en inglés) es una referencia global útil, que define una referencia común para los niveles de competencia en lectura y matemáticas que se espera que los estudiantes demuestren al final de cada nivel de grado, desde el grado 1 al 9. Los cuatro niveles descritos en el GPF – 'Inferior a cumple parcialmente', 'Cumple parcialmente', 'Cumple' y 'Supera la competencia mínima global' – forman una escala común de logro bajo a alto. El GPF y sus niveles de competencia relacionados brindan orientación sobre el conjunto mínimo de habilidades que los estudiantes deben adquirir en el camino hacia el dominio de la lectura y las matemáticas.

### *Vincular los programas de evaluación al nivel mínimo de competencia*

Vincular una evaluación nacional, regional o internacional a la definición global de MPL requiere una metodología para identificar los mismos conceptos y definiciones en programas de evaluación creados para propósitos completamente diferentes para permitir cierto grado de comparabilidad e inferencias justas sobre los países comparados.

El proceso de hacer comparables diferentes evaluaciones (moderación) puede ser estadístico o no estadístico. Tres procesos son el proyecto Rosetta, la Metodología de Vinculación de Políticas y las Evaluaciones para Niveles Mínimos de Competencia (AMPL *por sus siglas en inglés*).<sup>3</sup>

---

<sup>3</sup>Para obtener más información sobre la vinculación de estrategias, sus costos, beneficios, estado de ejecución, hitos ejecutados y pendientes y cronogramas, consulte UIS (2023) *Reporting learning outcomes in basic education: Country's options for indicator 4.1.1* [gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Countrys-reporting-option\\_Zambia\\_2023.05.15\\_FINAL.pdf](https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Countrys-reporting-option_Zambia_2023.05.15_FINAL.pdf).

### Piedra Rosetta

El proyecto **Piedra Rosetta**, liderado por la Asociación Internacional para la Evaluación del Rendimiento Educativo (IEA), intentó armonizar datos de diferentes evaluaciones. Debe su nombre al famoso descubrimiento arqueológico que permitió la traducción entre diferentes lenguas escritas y está diseñado para relacionar las evaluaciones de aprendizaje regionales con las internacionales. El objetivo es proporcionar a los países que participaron en evaluaciones regionales (o nacionales), pero no en evaluaciones internacionales, información sobre la proporción de estudiantes que lograron MPL en lectura y matemáticas. En un primer esfuerzo por implementar este enfoque y establecer tablas de concordancia, las evaluaciones regionales Estudio Regional Comparativo y Explicativo (ERCE) y Programme d'analyse des systèmes éducatifs de la CONFEMEN ([PASEC](#)) estaban vinculados al Estudio Internacional de Tendencias en Matemáticas y Ciencias ([TIMSS](#)) para matemáticas y al Estudio Internacional de Progreso en Capacidad Lectora ([PIRLS](#)) para lectura.

### Vinculación de políticas

Otro enfoque para armonizar las evaluaciones es la metodología **Policy Linking**, que es un método no estadístico que utiliza el juicio para alinear y hacer coincidir elementos de la evaluación nacional con el GPF. Este proceso establece puntos de referencia globales comparables internacionalmente basados en los descriptores de cada punto de referencia del GPF. Tres tareas principales (alineación, coincidencia y establecimiento de puntos de referencia) se completarán en un taller de 5 a 6 días con 15 a 20 panelistas (docentes), expertos en currículo y evaluación de cada grado/materia, para identificar y establecer, si es posible, los puntos de referencia requeridos para la presentación de informes internacionales sobre el indicador 4.1.1 de los ODS.

Para producir puntos de referencia confiables para la presentación de informes internacionales, el *Policy Linking Toolkit* especifica cinco criterios, entre ellos: que un número suficiente de elementos nacionales estén alineados con el GPF; que las muestras son representativas a nivel nacional; y las evaluaciones nacionales se administran de acuerdo con estándares mínimos de calidad. La metodología de vinculación de políticas se propuso en 2017, se puso a prueba en 2019, se revisó en 2020 y se puso a prueba nuevamente en 2021-22. Posteriormente, el *toolkit* se revisó en 2023 y ahora se encuentra en una fase piloto.

### Evaluaciones para niveles mínimos de competencia

También se propusieron las Evaluaciones de Niveles Mínimos de Competencia (AMPL) como una forma de obtener datos que puedan usarse para medir y monitorear el indicador 4.1.1 de los ODS. Los AMPL son herramientas sólidas destinadas a medir el logro de MPL en lectura y matemáticas en un nivel determinado del ciclo educativo. Esto permite a los países producir datos de resultados de aprendizaje comparables internacionalmente para informar sobre el indicador global 4.1.1 de los ODS.

AMPL-a, que mide el dominio en los primeros grados, está en desarrollo y se pondrá a prueba y se administrará en 2023, tanto en inglés como en francés. AMPL-b, que mide el dominio al final de la primaria, se desarrolló en 2021 tanto en inglés como en francés y se implementó en seis países africanos como parte del proyecto de Monitoreo de los Impactos en los Resultados del Aprendizaje: Burkina Faso, Burundi, Costa de Marfil, Kenia, Senegal y Zambia. AMPL-b se administró como módulo independiente en Sierra Leona en 2022 y su implementación está prevista en Jordania y Pakistán.

### *Alternativas de los países para informar*

Para guiar la elección de la medición del aprendizaje y garantizar que los datos de la evaluación sean consistentes con los objetivos estratégicos a largo plazo de una toma de decisiones efectiva, el UIS, la UNESCO, el Banco Mundial y UNICEF han desarrollado un conjunto de principios conocidos como el Pacto para los Datos sobre el Aprendizaje. Estos principios son importantes no sólo para diseñar evaluaciones o decidir qué evaluación usar "de la estantería", sino también para desarrollar sistemas de evaluación nacionales. Esos principios se basan en los existentes; permitir la flexibilidad para asegurar la alineación con las necesidades de los países (no son "una solución única para todos"); fomentar la apropiación nacional mediante un enfoque impulsado por la demanda; y garantizar que los datos sean relevantes para la toma de decisiones.

Un sistema de evaluación nacional debería poder generar informes, gestionar mejoras en todos los niveles de la educación, guiar la toma de decisiones y vincular las evaluaciones a nivel del sistema con las evaluaciones formativas y las prácticas en el aula. Para garantizar que las evaluaciones puedan monitorear con precisión el progreso para la toma de decisiones, los datos

deben ser comparables internacionalmente. Cada país debería tener una evaluación diseñada para la comparabilidad internacional o que pueda utilizarse para lograr la comparabilidad internacional – un compromiso en el proceso de los ODS. Las opciones de los países para informar son varias (**Cuadro 4**), pero la elección debe guiarse por qué evaluaciones son adecuadas para el propósito y más rentables, teniendo en cuenta la situación inicial del país y el objetivo de tener comparabilidad a lo largo del tiempo y representatividad de los resultados a nivel nacional.

Cabe señalar dos casos especiales. Se podrían utilizar evaluaciones nacionales para informar, sujeto al uso de vinculaciones estadísticas que podrían implementarse utilizando módulos calibrados como AMPL. Herramientas como el MPL y el GPF sirven para comprender y comparar estándares globales, mientras que la vinculación de políticas puede involucrar a las partes interesadas nacionales para analizar la evaluación frente a esos estándares. El UIS considera a los países que quieren informar globalmente sobre el indicador 4.1.1. El UIS agrega un módulo calibrado a la evaluación nacional, como AMPL, y lo complementa con vinculación de políticas con fines de desarrollo de capacidades, dado que la metodología aún se encuentra en una fase piloto.

Un segundo caso está relacionado con los primeros grados, o 4.1.1a, donde herramientas como la Evaluación de Lectura/Matemáticas para los primeros grados (EGRA/EGMA), las evaluaciones dirigidas por ciudadanos de la PAL Network y el Módulo de Competencias Fundamentales de UNICEF de su encuesta de hogares MICS podrían ser utilizados. relevante y ser utilizado para la presentación de informes. Aunque estas evaluaciones se han aplicado globalmente, actualmente no pueden usarse para informes globales, principalmente porque no estaban destinadas a generar datos comparables. Sin embargo, tienen el potencial de usarse para informes globales y el UIS está estudiando cómo aprovechar al máximo dichas evaluaciones.<sup>4</sup>

---

<sup>4</sup>Consulte <https://world-education-blog.org/es/>

**Cuadro 4. Alternativas para que los países informen sobre el indicador 4.1.1 de los ODS**

	4.1. 1.a	4.1. 1.b	4.1.1.c	Cobertura
Evaluaciones nacionales: vinculación estadística a través de módulos calibrados				
AMPL	•	•		
Módulo PISA			•	
Participación en evaluaciones internacionales				
PILNA		•		Islas del pacifico
PASEC	•	•	•	África (francófona) principalmente
SACMEQ		•		África (sur y este)
SEA-PLM		•		Sudeste Asiático
LLECE	•	•		América Latina
TIMSS	•	•	•	Global
PIRLES	•	•		Global
PISA			•	Global

## EVALUACIÓN DEL ESTADO DE MEDICIÓN DE LOS INDICADORES 4.6.1, 4.7.4 Y 4.7.5

Se han logrado avances significativos en el establecimiento de marcos conceptuales, metodológicos y de presentación de informes para los indicadores 4.6.1, 4.7.4 y 4.7.5. Los marcos proporcionan un enfoque estructurado y ofrecen directrices y principios para la recopilación y el análisis de datos. Sin embargo, a pesar de los avances en la definición de marcos desde 2015, un problema sustancial es la baja cobertura de datos. Abordar este desafío requiere un esfuerzo concertado para establecer definiciones y métricas comunes, asegurando un enfoque estandarizado en todas las evaluaciones (**Cuadro 5**).

**Cuadro 5. Disponibilidad de evaluaciones para medir los indicadores 4.6.1, 4.7.4 y 4.7.5 de los ODS**

Indicadores	Metadatos	Fuentes de datos	Última administración	Duración del ciclo	Cobertura de países (N)	Cobertura de población (%)
4.6.1	<a href="#">Sí</a>	Programa para la Evaluación Internacional de Competencias de Adultos (PIAAC)	2017	10 años	37	21
4.7.4	<a href="#">Sí</a>	Estudio Internacional de Cívica y Ciudadanía (ICCS)	2016	6/7 años	23	10
4.7.5	<a href="#">Sí</a>	TIMSS, PISA	2019/2022	4/5 años	38	16

Los países de ingresos bajos y medios a menudo carecen de los recursos y la infraestructura necesarios para desarrollar e implementar herramientas para medir estos indicadores de manera efectiva. Para abordar esta cuestión se requiere no sólo el desarrollo de herramientas de medición apropiadas, sino también iniciativas específicas de creación de capacidad en países de ingresos bajos y medianos para garantizar que estos indicadores se midan de manera integral y precisa.

El UIS ha propuesto una solución para aumentar la presentación de informes sobre el indicador global 4.6.1 de los ODS. Aprovechando su experiencia con el Programa de Evaluación y Monitoreo de la Alfabetización (LAMP), ha desarrollado una versión reducida de la encuesta para reducir la carga operativa, financiera y técnica. En particular, esta mini encuesta LAMP reduciría, entre otros ajustes, el número de dominios de habilidades evaluados, adaptaría los procedimientos de muestreo y ofrecería la posibilidad de introducir la encuesta como un módulo de una encuesta de hogares existente, en lugar de como una encuesta independiente.<sup>5</sup>

---

<sup>5</sup>Véase <https://uis.unesco.org/sites/default/files/documents/mini-lamp-monitoring-progress-sdg4.6.1-2018-en.pdf>.

Actualmente no hay otra alternativa para aumentar la presentación de informes sobre los indicadores 4.7.4 y 4.7.5 de los ODS, aparte de que más países se unan a las evaluaciones transnacionales que los miden.

## CUESTIONARIOS DE CONTEXTO

Más allá de armonizar datos sobre resultados de aprendizaje y habilidades, armonizar cuestionarios de contexto de diferentes evaluaciones a gran escala es de vital importancia para permitir análisis comparativos sólidos de tendencias, patrones y determinantes de la desigualdad educativa entre países y a lo largo del tiempo.

Si bien el objetivo principal de las evaluaciones transnacionales es estimar los resultados del aprendizaje, también recopilan un rico conjunto de información básica:

- De los estudiantes, recopilan información sobre sus experiencias escolares, sus actitudes hacia las materias que se enseñan y las características de sus padres y hogares, además de información demográfica básica, como edad y sexo (Cuadro 6).
- De los docentes, recopilan información sobre sus actitudes hacia la enseñanza, sus opiniones sobre los recursos didácticos, su formación académica y su desarrollo profesional continuo.
- De las escuelas recopilan información sobre infraestructura, ubicación y opiniones de los directores de las escuelas sobre la disponibilidad de recursos en la escuela y cómo interactúan con los padres. Existe cierta variación en la objetividad de los datos recopilados.

**Cuadro 6. Cuestionarios típicos y datos recopilados en evaluaciones transnacionales**

Cuestionario	Preguntas típicas
Alumno	Información demográfica (sexo, edad) Antecedentes familiares y socioeconómicos Experiencias relacionadas con la escuela (incluida la exposición al acoso) Experiencias relacionadas con el aprendizaje (actividades en el aula) Autopercepciones, intereses y aspiraciones relacionadas con diferentes temas. Uso y dominio de las tecnologías de la información y la comunicación (TIC)
Docente	Información demográfica y de antecedentes (sexo, edad, años de docencia, materias impartidas) Calificaciones y capacitación Tipos de prácticas docentes utilizadas y desafíos enfrentados
Escuela (director)	Características de la escuela Opiniones sobre disponibilidad y adecuación de recursos Gestión y gobernanza Interacción con padres y comunidades escolares.

Algunos de estos cuestionarios proporcionan información sobre indicadores seleccionados del ODS 4. Estos cuestionarios también determinan para qué subpoblaciones se pueden estimar estos indicadores en función de características como la ubicación urbana o rural de las escuelas, el nivel socioeconómico de los estudiantes (en relación con otros estudiantes, no con la población) y el sexo de los estudiantes y docentes, todos los que pueden contribuir a los debates sobre la equidad.<sup>6</sup>

Existen importantes desafíos para desarrollar asesoramiento sobre políticas, debido a las diferencias en las definiciones de los cuestionarios en varias dimensiones, como las clasificaciones rurales/urbanas, el nivel socioeconómico (o riqueza), el período de referencia, etc. Las diferencias en las definiciones dificultan la comparabilidad de los resultados de la evaluación y la interpretación de los datos. Por ejemplo, lo que constituye "rural" en un país puede ser diferente en otro. De manera similar, las distintas definiciones de estatus socioeconómico

<sup>6</sup>Para obtener más información, consulte UIS (2022). *Monitoring of the Sustainable Development Goals using large-scale international assessments*  
<https://tcg.uis.unesco.org/wp-content/uploads/sites/4/2022/04/Monitoring-of-the-SDGs-Using-Large-Scale-International-Assessments-April-2022.pdf>.



pueden afectar el análisis de las disparidades en los resultados educativos entre diferentes estratos sociales o económicos. Cerrar estas brechas de definición requiere colaboración internacional y el desarrollo de marcos estandarizados que garanticen la uniformidad en las definiciones, permitiendo evaluaciones precisas y facilitando comparaciones significativas entre regiones y contextos socioeconómicos.

## AGENDA A FUTURO

Se proponen las siguientes soluciones para los problemas identificados en este documento.

### **Manual de armonización y presentación de informes de evaluaciones**

Con los avances logrados en los últimos años, es hora de compilar y actualizar periódicamente un manual con toda la información sobre los criterios de elegibilidad para presentar informes. Este manual servirá como guía técnica para los programas de evaluación del aprendizaje y presentará y actualizará el menú de evaluaciones adecuadas para informes. Los aspectos por incluir son, entre otros:

- Alineación con los estándares y currículo (las evaluaciones deben medir adecuadamente el plan de estudios y las habilidades previstos);
- Propiedades psicométricas (evidencias de confiabilidad, validez, dificultad adecuada y discriminación, entre otras).
- Representatividad (las muestras deben reflejar las poblaciones objetivo).
- Comparabilidad de las administraciones (los procedimientos de administración consistentes y estandarizados afectan la comparabilidad)
- Transparencia de los procesos (el diseño de la evaluación, el muestreo y el análisis deben estar bien documentados)
- Suficiencia para el proceso de vinculación (suficientes elementos equivalentes/niveles de competencia para permitir la vinculación)

- Consulta a las partes interesadas (proceso de consulta durante todo el diseño y la implementación)
- Viabilidad de la implementación de las pruebas de aprendizaje (costos, cronogramas, creación de capacidad y esfuerzo razonable para los países)

### **Sistema de acreditación**

Es el momento adecuado para introducir un sistema de acreditación claro y transparente. Los proveedores de evaluaciones, incluidas las organizaciones gubernamentales, podrán solicitar que las evaluaciones sean examinadas para determinar su idoneidad para informar sobre el indicador 4.1.1 de los ODS. Según el manual, una lista de verificación contendrá los estándares y criterios de elegibilidad que los solicitantes deben cumplir.

### **Incrementar la inversión y el enfoque para financiar datos de aprendizaje.**

La cobertura de datos sigue siendo desigual, especialmente para los países en desarrollo. Es fundamental realizar más inversiones para ampliar los datos de evaluación del aprendizaje a nivel mundial con un enfoque holístico que busque construir infraestructura para la producción de datos a nivel nacional, transfiriendo conocimientos y habilidades, pero también poder adquisitivo y toma de decisiones.

### **Promover la estandarización de los cuestionarios de contexto en las evaluaciones de aprendizaje.**

Los cuestionarios de contexto son un siguiente paso importante para respaldar la comparabilidad. Un acuerdo sobre definiciones clave y un acuerdo sobre elementos y formatos estándar para capturar las características individuales de los estudiantes, maestros y escuelas permitiría una mejor comparabilidad de las dimensiones de equidad y de los impulsores del aprendizaje.

### **Ampliar el enfoque mini-LAMP para medir la alfabetización de adultos para aumentar la cobertura del indicador 4.6.1 de los ODS**

La baja cobertura del indicador global 4.6.1 de los ODS significa que será eliminado de la lista de indicadores globales durante la Revisión de 2025 por parte del Grupo Interinstitucional y de Expertos sobre indicadores de los ODS. Sin embargo, las habilidades de los adultos seguirán estando en la agenda educativa mundial y una solución rentable sigue siendo una prioridad. Se invita a los Estados miembros a poner a prueba la herramienta mini-LAMP, que puede agregarse a las encuestas de hogares existentes.

### **Metodologías innovadoras sobre indicadores de baja cobertura**

Los ODS 4.7.4 y 4.7.5 merecen la exploración de enfoques que reflejen el enfoque AMPL de 4.1.1 al habilitar un módulo que mida los estándares mínimos acordados aprovechando las evaluaciones existentes, como el Estudio Internacional de Cívica y Ciudadanía de la IEA (ICCS).