



# ENSEÑANZA BASADA EN LA EVALUACIÓN: Nivel de sistemas

## Introducción

Lo que se mide se puede mejorar. Durante mucho tiempo, el mundo se centró en medir el acceso a la educación. Como resultado, hemos visto importantes mejoras en el número de niños en las aulas. Sin embargo, la medición de los resultados del aprendizaje se ha quedado atrás. En 2017, 100 países no evaluaban sistemáticamente los resultados del aprendizaje de los estudiantes. Donde existe evidencia, la misma es aleccionadora: según las Naciones Unidas, 617 millones de niños de primaria y primer ciclo de secundaria (es decir, el 55 % de todos los niños escolarizados) no dominan mínimamente la lectura o las matemáticas. Este problema es especialmente grave en el África subsahariana, donde más del 87 % y el 84 % de los estudiantes no eran competentes en lectura o matemáticas en 2015, respectivamente.<sup>1</sup>

El Banco Mundial ha argumentado que para hacer frente a esta crisis, “es necesario—aunque no suficiente—medir el aprendizaje”. **Para que la evaluación sirva de base a la enseñanza y mejore los resultados del aprendizaje, debe ir seguida de la acción.** La medición permite a los gobiernos identificar quién está aprendiendo y quién no, de modo que puedan dirigir el apoyo a las escuelas y poblaciones más necesitadas y garantizar la equidad en el apoyo al aprendizaje. También les permite identificar lo que funciona para aumentar los resultados del aprendizaje y cómo podría mejorarse la enseñanza.

La enseñanza basada en la evaluación puede influir significativamente en la calidad de la enseñanza y en la mejora del aprendizaje, pero la evidencia es algo limitada. Los estudios han demostrado que existe una relación modesta entre las evaluaciones de alta calidad en diferentes niveles del sistema y mejores resultados de aprendizaje. Las evaluaciones del aprendizaje a gran escala pueden servir para responsabilizar al sistema. Se ha demostrado que los países con exámenes de egreso de alta calidad obtienen mejores resultados de aprendizaje que los que no los tienen, y las evaluaciones formativas en el aula se han relacionado con mejores resultados de aprendizaje.

Sin embargo, solo 17 de los 48 países del África subsahariana han completado recientemente evaluaciones a gran escala de alta calidad. Varios países han puesto en marcha diversas evaluaciones, incluidas evaluaciones nacionales del aprendizaje, y algunos donantes han intentado apoyar a los gobiernos en el desarrollo de estructuras de evaluación a nivel del sistema, pero pocos países han sido capaces de integrar las evaluaciones a nivel del sistema en un programa más amplio de aprendizaje de competencias básicas en lectura, escritura y matemáticas (FLN).

## Propósito y estructura de la guía

Esta guía describe cómo los responsables políticos, donantes y socios de la implementación pueden diseñar o perfeccionar las evaluaciones en el sistema educativo y conectar esos resultados con la enseñanza diaria.

### CUADRO 1. Definiciones

**Evaluación en el sistema:** Medición de los resultados de la enseñanza y el aprendizaje en los distintos niveles de un sistema educativo

**Evaluaciones a nivel del sistema:** Evaluaciones destinadas a valorar la efectividad de un sistema educativo

**Competencias básicas en lectura, escritura y matemáticas:** La capacidad de leer con sentido y realizar cálculos matemáticos básicos en 3o. grado

**Evaluaciones de aprendizaje a gran escala:** Pruebas estandarizadas nacionales o transnacionales que proporcionan una instantánea de los logros de aprendizaje

**Estándares de contenido:** Expectativas de lo que debe enseñarse a los estudiantes de un determinado grado

**Estándares de desempeño:** Descripciones de lo que los estudiantes deben ser capaces de hacer

**Exámenes de alto riesgo:** Exámenes que determinan si un estudiante pasará al siguiente nivel

**Parámetro:** Calificación específica en una evaluación que demuestra que se ha cumplido un estándar

**Objetivo:** Una meta relativa al porcentaje de niños que alcanzarán el parámetro



Estas actividades permitirán al sistema obtener datos para orientar la enseñanza, apoyar la adaptación del programa, garantizar responsabilidades serias y no punitivas, y vincular los resultados de las evaluaciones a gran escala con lo que los maestros enseñan cada día. El propósito de la enseñanza basada en la evaluación es identificar dónde existen brechas de aprendizaje y dirigir los recursos y mejorar la calidad de la enseñanza para abordar esas brechas. Esto incluye las brechas entre los resultados del aprendizaje y el desempeño real de los estudiantes, así como las brechas entre grupos de niños debidas a desigualdades sistémicas (como brechas de género, brechas entre grupos étnicos, etc.).

Esta guía está organizada de la siguiente manera: En primer lugar, explica cómo analizar el contexto actual de la evaluación en su país. Luego describe los diversos niveles de evaluación que existen en muchos sistemas, así como los tipos de evaluaciones que pueden ser útiles para estos diferentes niveles. A continuación, explora las consideraciones de costo vinculadas a estas evaluaciones. Por último, ofrece una serie de recomendaciones para aumentar el retorno de las inversiones en evaluación en términos de mejora de los resultados del aprendizaje.



## ANÁLISIS DEL SISTEMA DE EVALUACIÓN ACTUAL

Antes de embarcarse en una campaña para mejorar el uso y la sostenibilidad de las evaluaciones, es importante llevar a cabo una revisión de qué evaluaciones se utilizan actualmente en el contexto y cómo se relacionan con la enseñanza. Esta revisión debe incluir un análisis de lo siguiente:

**Qué evaluaciones existen actualmente:** Esto incluye examinar tanto “lo que debería ser” como “lo que es”. Para identificar “lo que debería ser”, comience por realizar una revisión documental de las leyes y políticas gubernamentales con respecto a las evaluaciones e inspecciones, si existen. Intente averiguar en qué niveles se exigen las evaluaciones y cómo estas evaluaciones existentes están vinculadas a la mejora de la enseñanza. Para identificar “lo que es”, esta revisión debe incluir un examen de los informes técnicos de las evaluaciones más recientes en el sistema, los informes de inspección del gobierno, los resultados de los datos del Enfoque sistémico para lograr mejores resultados en la educación (Systems Approach for Better Education Results, SABER) del Banco Mundial, y los datos del sistema de información de gestión educativa (education management information system, EMIS) del país. Tras la revisión documental, realice entrevistas con el Ministerio de Educación, el cuerpo de inspectores (si existe) y cualquier otro organismo gubernamental que participe en la realización de evaluaciones e inspecciones, con el fin de determinar la frecuencia de las evaluaciones sobre el terreno, si se trata de evaluaciones por muestreo o censales, cómo se financian y quién las realiza.

**Cómo se utilizan las evaluaciones:** Resulta útil comprender cómo se difunden y utilizan actualmente los resultados de las evaluaciones, especialmente si se utilizan para identificar dónde existen brechas de aprendizaje en un esfuerzo por abordar esas brechas. A este respecto, debe examinar varios niveles de evaluación:

- **Nivel internacional** – p. ej., para informar a los donantes o sobre los Objetivos de Desarrollo Sostenible
- **Nivel nacional** – p. ej., para orientar las políticas y planes de estudios, dirigir recursos o hacer avanzar a los estudiantes
- **Nivel subnacional** – p. ej., para una orientación más detallada de los recursos y el apoyo, como formación, mentoría y visitas de inspección
- **Nivel escolar y del aula** – p. ej., para fundamentar cambios en la pedagogía, priorizar el tiempo lectivo u orientar la enseñanza diferenciada y los programas de aceleración

La recopilación de esta información requerirá entrevistas con las organizaciones gubernamentales responsables de la difusión y el uso de los resultados—para identificar a quién distribuyen los resultados, cómo y cuándo. Estos esfuerzos de recopilación de información no deben limitarse a una encuesta en papel; hay que realizar entrevistas o debates de grupo con funcionarios del gobierno y maestros. **Asegúrese de preguntar si han visto los resultados, qué significan para ellos, cómo los utilizaron y qué podría facilitar su comprensión y uso.**

Examine la alienación entre la evaluación, el plan de estudios y la enseñanza (como se describe con más detalle en el punto 3 de la sección sobre Mejora de la evaluación a nivel del sistema, más adelante).

**En qué se basa el comportamiento de los maestros: También es importante comprender qué impulsa actualmente el comportamiento de los maestros y las decisiones didácticas, y en particular cómo estos elementos se interrelacionan con el sistema de evaluación.** Es conveniente saber cómo deciden los maestros qué incluir en sus planes de lección diarios, qué métodos de enseñanza y evaluación utilizan y cómo afectan los exámenes a sus decisiones pedagógicas. Por ejemplo, ¿qué determina que los maestros sean percibidos como exitosos? ¿Sienten los maestros la presión de “enseñar para la prueba”, y se dirigen esas pruebas a garantizar la comprensión conceptual por parte de los alumnos de las competencias FLN clave? ¿Cómo se apoya a los maestros en las técnicas clave de enseñanza y evaluación? Esta información se puede recabar de los maestros mediante entrevistas, debates de grupo o encuestas.

## COMPRENDER LOS NIVELES DE EVALUACIÓN

Por lo general, los países cuentan con alguna estructura de evaluación que existe en varios niveles del sistema educativo, desde el internacional al regional, nacional, subnacional, comunitario y escolar, como se muestra en la Figura 1. Para garantizar un sistema de evaluación coherente, es fundamental que las evaluaciones a estos diferentes niveles estén alineadas y se basen unas en otras. A continuación analizamos con más detalle los niveles y tipos de evaluaciones.



Esta guía se centra en las evaluaciones a nivel escolar y superior. La guía independiente [Enseñanza basada en la evaluación: Guía a nivel del aula](#) se centra en el nivel del aula y ofrece orientación sobre cómo desarrollar medidas de evaluación sencillas y de bajo riesgo que simplifiquen la tarea de comprender los niveles de aprendizaje, sin aumentar la carga de los maestros.

## DESCRIPCIÓN GENERAL DE LAS EVALUACIONES DE APRENDIZAJE A GRAN ESCALA

Existen varios tipos de evaluaciones de aprendizaje a gran escala (large-scale learning assessments, LSLA), que se utilizan en los diferentes niveles del sistema de evaluación (véase la Tabla 1). Considere cuáles de estas LSLA podrían ser apropiadas en el contexto de su país y cómo pueden funcionar conjuntamente para reducir costos, al tiempo que se maximiza su utilización para la toma de decisiones didácticas.

## EVALUACIONES INTERNACIONALES Y REGIONALES

*Los países deben tener en cuenta las evaluaciones internacionales y regionales, ya que ofrecen oportunidades para aumentar la comparabilidad de los resultados con otros países, incluida la utilización de los años de escolarización ajustados al aprendizaje (learning-adjusted years of schooling, LAYS).* Estas evaluaciones tienden a ser exámenes a papel y lápiz de alta calidad, confiables, transparentes y de bajo riesgo, que permiten a los países comparar sus resultados de aprendizaje a lo largo del tiempo y con los de otros países. Sin embargo, dada la necesidad de que sean relevantes en todos los contextos de los países, aparecen más tarde en el ciclo educativo (generalmente en 4o. grado o superior) y es posible que no se alineen bien con los planes de estudios de un país específico. También pueden ser sustancialmente más costosas que otras evaluaciones, y sus resultados pueden tardar años. Sin embargo, las evaluaciones internacionales y regionales funcionan bien a la hora de desarrollar la capacidad de evaluación en el país y pueden poner de manifiesto la necesidad de reformar el plan de estudios o identificar necesidades generales. Por ejemplo, en Uganda, el Ministerio de Educación creó una unidad de género después de que los resultados del SAQMEC mostraran impactos diferenciales por género, y en Namibia, los funcionarios desarrollaron una política para compartir buenas prácticas y recursos entre grupos de escuelas, después de que el SAQMEC revelara deficiencias en este sentido. Al igual que con otros tipos de evaluaciones, **asegúrese de que los resultados de las evaluaciones internacionales se compartan con actores de los distintos niveles del sistema educativo, con recomendaciones clave para la acción.** Con demasiada frecuencia, los maestros no tienen la oportunidad de revisar su pedagogía en base a los hallazgos de estas evaluaciones. Los programas FLN deben revisar detenidamente los hallazgos de estas evaluaciones para responder a los desafíos identificados en ellas.

## EXÁMENES NACIONALES

*Los países deben considerar la creación de estructuras en sus exámenes nacionales que puedan apoyar de manera efectiva la mejora de la enseñanza.* Los exámenes nacionales tienden a ser censales, a papel y lápiz, y de alto riesgo para los estudiantes y las escuelas, ya que a menudo determinan la asignación de los estudiantes o sirven como indicadores de la calidad de las escuelas, lo que da lugar a la clasificación de escuelas en muchos países. Al haber mucho en juego, se prestan para hacer trampa, incluso expulsando del sistema a los alumnos de bajo desempeño para inflar los resultados,<sup>76</sup> y pueden actuar reduciendo el enfoque de la enseñanza en el aula. Para combatir estos desafíos, **los exámenes nacionales deberían abarcar la totalidad del plan de estudios, incluidas, si es posible, las destrezas de orden superior, como la resolución de problemas, el razonamiento y la reflexión; utilizarse para incentivar (basándose en la mejora o en modelos que controlen otros factores que afectan a los resultados del aprendizaje) en lugar de castigar; y no utilizarse para clasificar públicamente a las escuelas si esas clasificaciones inducen a la trampa o a medidas punitivas.** Es importante garantizar una difusión amplia y selectiva de los resultados, pero de forma que incite a la acción y no a la búsqueda de culpables.<sup>77</sup> Por ejemplo, el Consejo Nacional de Exámenes de Kenia distribuye resultados que muestran punto por punto ejemplos de respuestas correctas e incorrectas y errores comunes. Las escuelas y los maestros pueden utilizar esta información para orientar la enseñanza. Es importante que los programas FLN se alineen con lo que se evalúa en estos exámenes nacionales, dado lo poderosos que son. En muchos casos, las destrezas abordadas en FLN son componentes básicos de las destrezas que se evalúan en los exámenes—en este caso, crear conciencia entre las partes interesadas sobre este vínculo puede ser útil y puede permitir una implementación más efectiva del programa FLN.

FIGURA 1. Niveles de evaluación

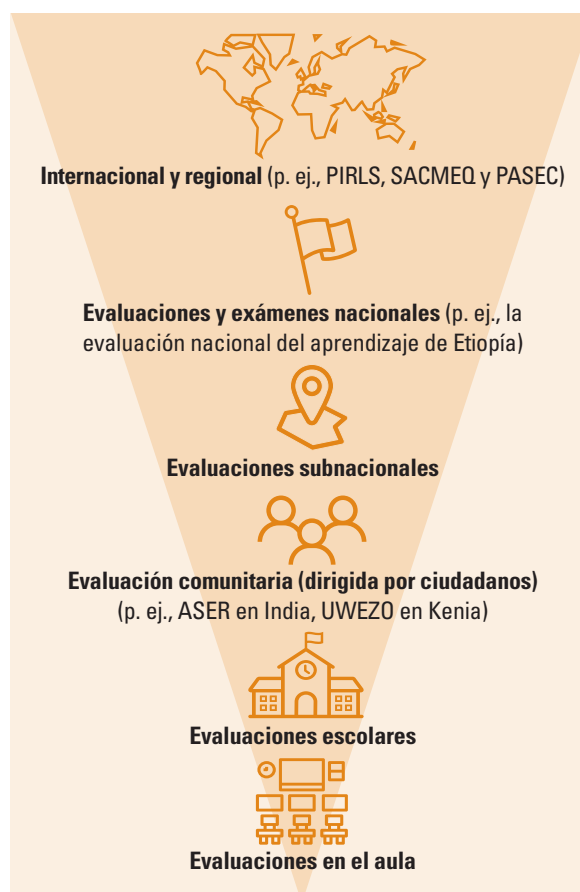




TABLA 1. Tipos de evaluaciones de aprendizaje a gran escala

TIPO DE EVALUACIÓN	DESCRIPCIÓN GENERAL	EJEMPLOS	COBERTURA GEOGRÁFICA	POBLACIÓN OBJETIVO	ASIGNATURA(S)
Evaluaciones internacionales	Evaluaciones estandarizadas, administradas en grupo, basadas en la escuela o por muestreo y destinadas a permitir comparaciones del desempeño de los estudiantes entre países y a lo largo del tiempo <sup>5</sup>	PISA <sup>6</sup>	En todo el mundo, principalmente países de la OCDE <sup>7</sup>	Niños de 15 años <sup>8</sup>	Lectura, matemáticas y ciencias <sup>9</sup>
		PISA-D <sup>10</sup>	Países de bajos y medianos ingresos (LMIC) <sup>11</sup>	Niños de 15 años en 7o. grado o superior y niños de 14 a 16 años no escolarizados o inscritos en 6o. grado o inferior <sup>12</sup>	Lectura, matemáticas y ciencias <sup>13</sup>
		TIMSS <sup>14</sup>	En todo el mundo, principalmente países de la OCDE <sup>15</sup>	4o. y 8o. grados <sup>16</sup>	Matemáticas y ciencias <sup>17</sup>
		PIRLS <sup>18</sup>	En todo el mundo, principalmente países de la OCDE <sup>19</sup>	4o. grado <sup>20</sup>	Lectura <sup>21</sup>
		LaNA <sup>22</sup>	Países LMIC <sup>23</sup>	4o., 5o. o 6o. grado <sup>24</sup>	FLN <sup>25</sup>
Evaluaciones regionales	Evaluaciones estandarizadas, administradas en grupo, basadas en la escuela y por muestreo, en las que participan varios países de una región <sup>26</sup>	PASEC <sup>27</sup>	Países africanos de habla francesa <sup>28</sup>	2o. y 6o. grados <sup>29</sup>	Lectura y matemáticas <sup>30</sup>
		LLECE <sup>31</sup>	América Latina <sup>32</sup>	3o. y 6o. grados <sup>33</sup>	Lectura, escritura, matemáticas, ciencias, factores nacionales que afectan al aprendizaje (diversidad cultural, violencia, etc.) <sup>34</sup>
		SACMEQ <sup>35</sup>	Países del sur y este de África <sup>36</sup>	6o. grado <sup>37</sup>	Lectura y matemáticas <sup>38</sup>
Exámenes nacionales	Exámenes de alto riesgo administrados en grupo y utilizados para tomar decisiones de certificación y promoción <sup>39</sup>	Muchos <sup>40</sup>	Nacional o subnacional <sup>41</sup>	Final de la escuela primaria; final del primer ciclo de secundaria; exámenes de ingreso a la universidad <sup>42,43</sup>	Varían según el país, pero la mayoría incluye lectura y matemáticas como mínimo <sup>44</sup>
Evaluaciones nacionales y subnacionales	Evaluaciones variadas, pero generalmente administradas en grupo, basadas en la escuela y por muestreo, destinadas a evaluar la calidad de todo el sistema <sup>45,46</sup>	NARN en Nepal; NLA en Etiopía <sup>47</sup>	Nacional o subnacional, respectivamente <sup>48</sup>	Generalmente 4o. grado y superior <sup>49</sup>	Varían según el país, pero la mayoría incluye lectura y matemáticas <sup>50</sup>
Evaluaciones de la lectura en los primeros grados y las matemáticas en los primeros grados	Evaluaciones basadas en la escuela y por muestreo, administradas individualmente, que miden las FLN y pensadas como diagnóstico del sistema <sup>51</sup>	EGRA <sup>52</sup>	LMIC <sup>53</sup>	1o. a 3o. grado (y a veces 4o.) <sup>54</sup>	Lectura <sup>55</sup>
		EGMA <sup>56</sup>	LMIC <sup>57</sup>		Matemáticas <sup>58</sup>
Evaluaciones comunitarias y en el hogar	Evaluaciones basadas en el hogar y por muestreo, administradas individualmente por voluntarios, que miden las FLN y están destinadas a conformar la percepción de la comunidad <sup>59</sup>	ASER <sup>60</sup>	LMIC <sup>61</sup>	Niños de 5 a 16 años <sup>62</sup>	Lectura y matemáticas <sup>63</sup>
		ICAN <sup>64</sup>	13 países de todo el mundo <sup>65</sup>	Niños de 5 a 16 años <sup>66</sup>	Matemáticas <sup>67</sup>
		MICS <sup>68</sup>	118 países de todo el mundo <sup>69</sup>	Niños de 5 a 17 años <sup>70</sup>	Lectura y matemáticas <sup>71</sup>
		UWEZO <sup>72</sup>	Kenia, Tanzania y Uganda <sup>73</sup>	Niños de 6 a 16 años <sup>74</sup>	Lectura y matemáticas básicas <sup>75</sup>

## EVALUACIONES NACIONALES Y SUBNACIONALES

Los países deben considerar la posibilidad de realizar evaluaciones nacionales o subnacionales para identificar las zonas del país que requieren apoyo adicional, así como las destrezas específicas que los maestros necesitan mejorar en la enseñanza. Se trata de evaluaciones de bajo riesgo, por muestreo y normalmente a papel y lápiz, diseñadas para alinearse con el plan de estudios. Se pueden administrar a escala nacional o en unidades administrativas de nivel medio dentro del sistema educativo y se utilizan para examinar la efectividad del sistema o los impactos de las políticas educativas nacionales o subnacionales. Los hallazgos suelen tomarse muy en serio porque, por lo general, pertenecen y son implementados por el gobierno. Sin embargo, las evaluaciones nacionales tienden a realizarse hacia el final de la escuela primaria o en secundaria, lo que a menudo llega demasiado tarde para identificar las brechas en los resultados del aprendizaje de FLN e intervenir en consecuencia. Los alumnos que vienen rezagados en 1o. y 2o. grados tienden a quedarse atrás, y la brecha entre estos alumnos y el resto de su clase se amplía.<sup>78</sup> Al implementar una evaluación nacional o subnacional, asegúrese de que incluya la evaluación de los primeros grados, que esté adecuadamente alineada con el sistema educativo (véase la Figura 6 en la página 10) y que incluya parámetros y objetivos. Considere la posibilidad de utilizar metodologías que permitan comparar los resultados a lo largo del tiempo y entre contextos subnacionales. Véase el Cuadro 4 en la página 11 sobre parámetros y vinculación de las evaluaciones.



## EVALUACIONES DE LA LECTURA EN LOS PRIMEROS GRADOS Y LAS MATEMÁTICAS EN LOS PRIMEROS GRADOS (EARLY-GRADE READING Y EARLY-GRADE MATH ASSESSMENTS, EGRA Y EGMA)

Los países deben considerar el uso de evaluaciones de la lectura y las matemáticas en los primeros grados—como EGRA y EGMA, o evaluaciones similares—para identificar los niveles de destreza particulares de los niños en FLN y los desafíos pedagógicos del sistema. Se trata de evaluaciones de FLN individuales, basadas en la escuela y por muestreo, que suelen administrarse a estudiantes de 1o. a 3o. grado. Se pueden implementar a cualquier nivel—desde el nacional hasta el escolar—y utilizan herramientas de código abierto que se adaptan al contexto local y al idioma de la evaluación. El Banco Mundial, USAID y otras instituciones identificaron en 2006 una brecha en las evaluaciones, señalando que las evaluaciones y los exámenes nacionales no solían realizarse hasta 4o. grado—demasiado tarde para identificar las principales brechas de aprendizaje, como se mencionó anteriormente. Además, los resultados de estas evaluaciones a menudo eran deficientes y resultaba difícil saber si se debía a que los alumnos no podían leer y comprender las instrucciones de la evaluación o a que no poseían los conocimientos evaluados. Por ello, en 2006, el Banco Mundial, USAID y otros financiaron la creación de una evaluación sencilla y efectiva de los resultados de lectura de los estudiantes de los primeros grados—EGRA. En la Figura 2 se muestra un ejemplo del módulo de comprensión lectora de EGRA. Tras la demanda inicial de EGRA, USAID financió en 2008 la creación de EGMA para su uso en la evaluación de los primeros resultados en matemáticas.

EGRA, EGMA y otras evaluaciones similares se diseñaron para proporcionar un diagnóstico a nivel del sistema de las brechas en los resultados del aprendizaje. Varios países, entre ellos Angola y Ghana, han adoptado estas herramientas como sus evaluaciones nacionales. Angola creó una unidad de evaluación en el Ministerio de Educación y dedicó una partida presupuestaria de USD 1 millón<sup>80</sup> para estas actividades de evaluación. El ministerio utiliza los resultados de la evaluación para orientar la formación docente continua.

También se han realizado esfuerzos para desarrollar evaluaciones administradas en grupo que se centran en muchas de las mismas destrezas abordadas a través de EGRA. Por ejemplo, la más tradicional Evaluación de lectura y escritura administrada en grupo (Group Administered Literacy Assessment, GALA) se diseñó para aprovechar las lecciones aprendidas de EGRA y evaluar una serie de habilidades de lectura en los primeros grados, a la vez que se reducía el tiempo de administración y se simplificaba la calificación.<sup>81</sup> GALA se ha utilizado en Egipto, Ghana, Etiopía y Tanzania para ahorrar costos, mientras que en Nepal se desarrolló una herramienta similar basada en GALA, la Evaluación en el aula de la lectura en los primeros grados.<sup>82</sup>

EGRA y EGMA son más útiles cuando se integran en los sistemas gubernamentales para maximizar su utilización, lo que puede ser más fácil de lograr con versiones administradas en grupo. Estas herramientas han sido criticadas por sus elevados costos por estudiante.

### EVALUACIONES COMUNITARIAS

Los países deben fomentar la implementación de evaluaciones comunitarias para tener una comprensión desglosada de los resultados del aprendizaje. Las evaluaciones comunitarias son realizadas oralmente en los hogares por voluntarios de la comunidad. Estas evaluaciones, que suelen ser rápidas y fáciles de administrar, reflejan los resultados de FLN en niños de un amplio rango de edades, tanto escolarizados como no escolarizados. En las Figuras 3 y 4 a continuación se incluyen ejemplos de elementos de las evaluaciones comunitarias. Si bien las evaluaciones comunitarias han demostrado tener éxito en generar una amplia conciencia sobre los malos resultados del aprendizaje en muchos países y también sirven como control de los sistemas de evaluación gubernamentales, rara vez han resultado en cambios inmediatos en los sistemas educativos en su conjunto. Sin embargo, en el estado de Bihar, India, se utilizaron para identificar comunidades con grandes necesidades, en las que el gobierno estatal luego implementó programas de enseñanza correctiva y creó bibliotecas, con efectos positivos.<sup>86</sup> Al implementar una evaluación comunitaria, asegúrese de que haya defensores en los niveles del sistema donde se utilizarán los resultados y de que esos defensores tengan recomendaciones prácticas para mejorar realmente el aprendizaje.

**FIGURA 2. Ejemplo del módulo de comprensión lectora de EGRA<sup>79</sup>**

**Subtarea 2b. Comprensión lectora**

Al cabo de 60 segundos o si el niño completó la lectura del pasaje en menos de 60 segundos, dígame al niño que le va a hacer algunas preguntas sobre el cuento pero que primero quiere darles más tiempo para que vuelvan a leer el cuento. Dele al estudiante hasta 3 minutos para leer el cuento y luego pregúntele si terminó de leer. Si el estudiante dice que sí, haga las preguntas después de decir:

Ahora te haré algunas preguntas acerca del cuento que acabas de leer. Trata de responder las preguntas lo mejor que puedas.

Kati ηηenda kubuuzayyo ebibuuzo bitonotono ku ebyo by'osomye. Gezaako okuddamu obulungi nga bw'osobola.

- Dele al niño 15 segundos como máximo para comenzar a responder la pregunta, marque la respuesta y prosiga con la siguiente pregunta.
- Lea las preguntas para cada línea hasta el corchete que indica hasta dónde leyó el niño.

Sam vive en una granja. Tiene cabras. A Sam le gustan las cabras. Son simpáticas. ¡Comen de todo! Él les da de comer después de la escuela. Sam coloca su tarea cerca de las cabras. Juega al fútbol. Luego lo llama su hermana. ¡Su tarea desapareció! Una de las cabras está comiendo algo blanco. Ahora sabe dónde está su tarea.

	Correcto	Incorrecto	Sin respuesta
¿Dónde vive Sam? [en una granja]			
¿Qué tiene Sam? [cabras]			
¿A qué jugó Sam? [fútbol]			
¿Quién llama a Sam? [su hermana]			
¿Qué pasó con la tarea de Sam? [se la comió una cabra]			



## CUADRO 2. Actualización de EGRA

En los últimos años, los usuarios de EGRA han realizado algunos ajustes para comprender mejor las capacidades de los estudiantes. Por ejemplo:

- Proporcionar más tiempo para el pasaje de lectura oral o un pasaje adicional. Esto aumenta el número de niños que son evaluados en todas las preguntas de comprensión correspondientes.
- Agregar un módulo de lenguaje expresivo, aún en desarrollo por USAID en 2021. El módulo ayuda a los usuarios a comprender las capacidades de los alumnos que reciben una calificación de cero en los módulos de decodificación y comprensión lectora.
- Ampliar las medidas de comprensión lectora. Esto incluye la lectura silenciosa, la lectura “sin tiempo”, el laberinto, rellenar los espacios (cloze), mirar atrás y una variedad de otras formas de aumentar la confiabilidad y validez de las medidas de comprensión de EGRA.
- Desarrollar y poner a prueba otras medidas de ortografía y lenguaje oral.<sup>83</sup>

**FIGURA 3. Ejemplos de elementos de la evaluación en matemáticas de UWEZO<sup>84</sup>**

**Ejemplo de prueba de matemáticas**

Contar:  
¿Cuántos números hay en cada conjunto?  
(Debería intentar cualquiera de los 5, por lo menos 4 deben ser correctos)

Reconocimiento de números 10-99:  
(Debería intentar cualquiera de los 5, por lo menos 4 deben ser correctos)

15	47	25	23
30	94	36	51

Reconocimiento de números 100-999:  
(Debería intentar cualquiera de los 5, por lo menos 4 deben ser correctos)

104	129	200	476
374	234	581	943

**FIGURA 4. Módulo de evaluación de la lectura de ASER<sup>85</sup>**

En el jardín había un gran árbol. Estaba solo y se sentía solitario. Un día llegó un pájaro y se posó sobre él. El pájaro tenía una semilla en el pico. Dejó caer la semilla cerca del árbol. Allí creció una pequeña planta. Pronto, creció otro árbol. El árbol grande estaba feliz.

A Rani le gusta su escuela. Su clase es en una sala grande. Rani tiene una bolsa y un libro. También tiene una lapicera.

e	d	w
s	c	
g	h	z
i	q	

mano	estrella
autobús	
gato	libro
día	algunos
	viejo
cantar	valiente

## EVALUACIONES ESCOLARES

Las evaluaciones escolares consisten en evaluaciones a escuelas individuales, exámenes locales e inspecciones formales.

- **Las evaluaciones a escuelas individuales** se llevan a cabo en toda la escuela para orientar los planes de mejora escolar. Los directores de escuela suelen compartir los resultados con los maestros, los padres y las comunidades de gestión escolar, quienes luego participan en la planificación de la mejora escolar.
- Los **exámenes locales** son exámenes establecidos por un pequeño número de escuelas para permitir una revisión y comparación de los resultados del aprendizaje.
- Las **inspecciones** consisten en visitas de funcionarios de la educación, normalmente de nivel subnacional, para evaluar la calidad o el desempeño de las escuelas y los maestros mediante la observación de las aulas, entrevistas con los maestros y la revisión de documentos y datos clave.<sup>87</sup> Las inspecciones también pueden conocerse como supervisión. Es importante diferenciar las inspecciones, que tienen por objeto evaluar la calidad, de la mentoría, que no es un tipo de evaluación y está destinada a brindar apoyo a los maestros.

Las evaluaciones escolares podrían reflejarse en el EMIS de un país para permitir a los funcionarios locales hacer un seguimiento de las escuelas con el fin de garantizar la implementación de los planes de mejora y permitir a los funcionarios locales y nacionales dirigir el apoyo a las escuelas según sea necesario.

## Usos de las evaluaciones escolares

Aunque históricamente las inspecciones no incluían evaluaciones de los resultados del aprendizaje de los estudiantes, cada vez son más los países que incluyen pruebas a pequeñas muestras de estudiantes en sus visitas de inspección para vincular más claramente los aportes a nivel escolar (como el acceso a los materiales, el liderazgo escolar y la enseñanza) con los resultados del aprendizaje de los estudiantes. Algunos países también han difundido la retroalimentación de evaluaciones de aprendizaje a gran escala a través de inspectores o “coaches”, para que sirvan de base a los ajustes de los elementos del programa a nivel escolar. Los resultados de los exámenes locales pueden utilizarse del mismo modo y a menudo son una buena oportunidad para compartir y discutir los resultados del aprendizaje con los padres. En algunos países, organizaciones locales como UWEZO y Pratham también han desempeñado un papel en el seguimiento y la difusión de los resultados del aprendizaje con las escuelas y las comunidades para apoyar la mejora escolar. Algunos ejemplos de cómo se han utilizado los resultados de las evaluaciones de los estudiantes a nivel escolar son:



- **Involucrar a los padres y a la comunidad en las evaluaciones escolares y la planificación de mejoras:** La Campaña para la educación femenina, que funciona en Ghana, Malawi, Tanzania, Zambia y Zimbabue, apoya a sus exalumnas para que regresen a las comunidades rurales y sirvan como Guías del alumno para las niñas. Las Guías del alumno utilizan tabletas para recoger datos sobre la retención, el progreso y los resultados del aprendizaje de los alumnos. Luego comparten esta información con las escuelas y las comunidades para que sirva de base a la planificación de acciones. A cambio de su compromiso, las exalumnas acceden a préstamos sin intereses para desarrollar sus propias empresas o reciben una calificación profesional docente que les permite acceder rápidamente a los institutos de formación docente.<sup>88</sup>
- **Incorporación de evaluaciones puntuales de los resultados del aprendizaje de los estudiantes en las visitas de inspección periódicas:** Los gobiernos de Ghana, Jordania, Kenia, Ruanda y Uganda han incorporado evaluaciones de una pequeña muestra de estudiantes en sus visitas de inspección, con el fin de orientar mejor las áreas de apoyo escolar y tutoría de maestros. Por ejemplo, la actividad Soma Umenye de USAID colaboró con el gobierno de Ruanda para incorporar a las inspecciones escolares evaluaciones de fluidez y comprensión lectora de 1o. a 3o. grado. El equipo de Soma Umenye formó a los funcionarios de educación de distrito sobre cómo realizar las evaluaciones durante las visitas a las escuelas. Los funcionarios realizan las evaluaciones utilizando pasajes de las lecciones del libro de texto de esa semana o de la semana anterior, con cinco estudiantes por clase seleccionados al azar. Luego, los resultados de las evaluaciones se utilizaron para orientar las conversaciones de mentoría con los maestros y las discusiones con los directores de escuela. Por ejemplo, si una evaluación mostraba que los estudiantes decodificaban pero no comprendían los pasajes de lectura, el apoyo podía dirigirse a ayudar a los maestros a mejorar la calidad de sus estrategias de comprensión o a dedicar más tiempo a la práctica de la lectura dentro del horario. Esta actividad de supervisión ha demostrado ser práctica y rentable para los gobiernos.<sup>90</sup>
- **Compartir con las escuelas los resultados de los exámenes locales y las evaluaciones de aprendizaje a gran escala:** Las inspecciones escolares pueden aprovecharse para compartir los resultados de los exámenes locales y de las evaluaciones de aprendizaje a gran escala con los maestros y directores de escuela, y para ayudarles a utilizar los resultados para introducir mejoras a nivel escolar.<sup>91</sup> Por ejemplo, el gobierno de Malawi, con el apoyo de UNICEF y de la actividad Yesani Ophunzira Activity (YESA) de USAID, aprovechó las estructuras existentes del EMIS para incluir lo siguiente:
  - Datos de observación en el aula facilitados por funcionarios del Ministerio de Educación a nivel zonal
  - Datos de las inspecciones escolares realizadas por funcionarios de educación de distrito (la unidad administrativa por encima del nivel zonal)
  - Datos sobre los resultados de lectura de los estudiantes, procedentes de la Evaluación nacional de lectura del país

### CUADRO 3. EGRA administrada en grupo en Nepal

Con el apoyo del Early-Grade Reading Program (Programa de lectura para los primeros grados) de USAID, el Ministerio de Educación, Ciencia y Tecnología de Nepal desarrolló una evaluación de la lectura en los primeros grados administrada en grupo, a la que denominó CB-EGRA. Esta evaluación fue administrada por maestros de todas las escuelas de los distritos participantes en el Programa nacional de lectura en los primeros grados del gobierno. Los resultados se compartieron en las reuniones de la asociación de padres y maestros y de los comités de educación de los poblados, en las que padres y miembros de la comunidad analizaron los resultados y exploraron las razones del bajo desempeño en comprensión lectora. Estas iniciativas llevaron a todas las partes interesadas, incluidos gobiernos locales, comités de gestión escolar, asociaciones de padres y maestros, padres y comunidades, a centrarse en formas de apoyar el desempeño académico de los niños.<sup>89</sup>

#### Esta integración de los resultados de lectura de los estudiantes es un poderoso complemento de las actividades típicas del EMIS.

Los asesores de educación primaria y los inspectores escolares introducen los datos obtenidos en sus visitas a las escuelas directamente en el EMIS mediante una aplicación móvil, accesible sin conexión, que se sincroniza con la base de datos una vez que los funcionarios tienen acceso a Internet. Esta información luego se pone a disposición de maestros, directores de escuela,

“coaches”, inspectores y funcionarios de todos los niveles del sistema educativo a través de un panel de control móvil.

**Aún más poderosa es la capacidad de utilizar los datos para servir de base al diseño e implementación de la formación y tutoría de los maestros a nivel escolar, y a la orientación de los recursos y el apoyo a nivel de zona, distrito y nacional.**

### Mejores prácticas para las evaluaciones escolares

Para garantizar que las inspecciones y los exámenes locales sean efectivos:

- **Identifique su propósito**, reconociendo que las evaluaciones de los estudiantes a nivel escolar son más útiles cuando conducen a un apoyo específico para los maestros y directores de escuela.
- **Dirija el apoyo:** Inspeccione todas las escuelas y ofrezca apoyo continuo a todos los maestros, pero realice inspecciones y preste apoyo con mayor frecuencia a los maestros y las escuelas con dificultades (en las que los datos de las evaluaciones sugieren que los estudiantes se están quedando atrás).<sup>92</sup>
- **Asegúrese de que las visitas de inspección resulten en una retroalimentación práctica**, con un nivel de detalle suficiente para orientar las adaptaciones de las escuelas y los maestros.
- **Responsabilice a las escuelas:** Si bien las inspecciones y los exámenes locales no deben ser de alto riesgo (p. ej., dar lugar a sanciones para los estudiantes, los maestros o las escuelas), las escuelas deben hacerse responsables del cumplimiento de recomendaciones claras. Los “coaches” deben ayudar a las escuelas a seguir las recomendaciones, y los inspectores deben hacer un seguimiento con las escuelas para asegurarse de que se implementen las recomendaciones.



## EVALUACIÓN EN EL AULA

Las evaluaciones utilizadas en el aula son evaluaciones formativas y sumativas, administradas por el maestro, de todos los alumnos que los maestros pueden utilizar para orientar la enseñanza diferenciada y los cambios en su enfoque pedagógico general, o en el enfoque del contenido. Estas evaluaciones se analizan con más detalle en la guía [Enseñanza basada en la evaluación: Nivel del aula](#) y en un artículo publicado por la [Global Reading Network \(Red global de lectura\) sobre la evaluación formativa](#). El objetivo es garantizar que los maestros tengan y puedan utilizar herramientas y métodos para realizar un seguimiento del progreso y de los resultados del aprendizaje en sus aulas. Esto debe lograrse sin crear una carga adicional significativa para los maestros y, al mismo tiempo, con un conjunto claro de sugerencias sobre cómo mejorar la enseñanza en base a la información obtenida a través de la evaluación.

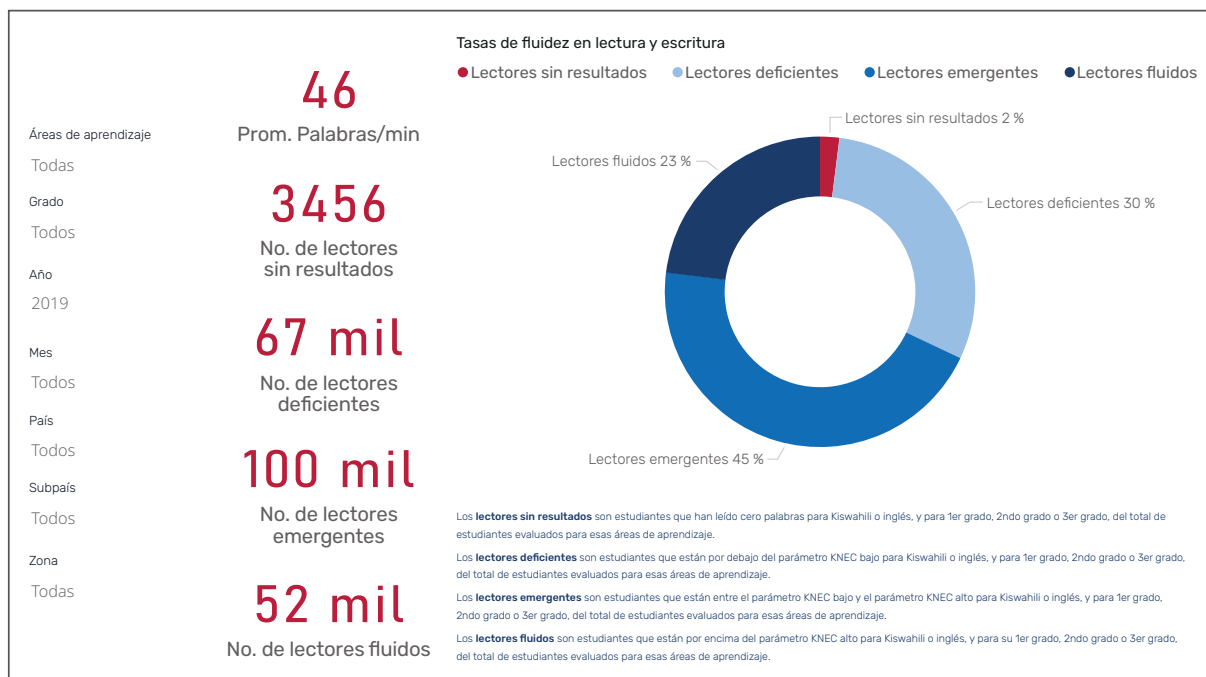
## Costo de las evaluaciones

Aunque los costos de las inspecciones suelen correr a cargo de los gobiernos locales, la escasez de recursos públicos hace que a menudo no se realicen con la frecuencia necesaria para maximizar su impacto. En cambio, los costos de las LSLA suelen correr a cargo de los donantes. Por lo tanto, al considerar los costos, asegúrese de que los beneficios potenciales de la evaluación superen los costos y que la evaluación sea sostenible más allá de la duración del programa FLN.<sup>93</sup> Planifique la transición de los costos al gobierno a lo largo del tiempo.

Los costos varían considerablemente en función del contexto, por lo que es importante examinar los costos correspondientes de los distintos tipos de evaluaciones en el sistema, así como la forma en que se utilizan actualmente los resultados de cada evaluación. Algunas contrapartidas a considerar cuando se intenta reducir costos son:

- **La amplitud de las evaluaciones:** Las evaluaciones que se utilizarán para supervisar la calidad del sistema educativo y dirigir el apoyo a los grandes grupos no tienen por qué ser tan largas o exhaustivas (p. ej., podrían evaluar solo unas pocas destrezas clave en lugar de todo el plan de estudios) como las evaluaciones centradas en dirigir el apoyo a maestros o estudiantes individuales, que deberían cubrir toda la amplitud del plan de estudios para ayudar a diagnosticar brechas específicas en el aprendizaje.
- **Tamaño de la muestra:** El uso de muestras más pequeñas, incluidos los métodos de muestreo para garantizar la calidad de los lotes (lot quality assurance sampling, LQAS)<sup>94</sup>, puede ser útil para identificar tendencias generales en los resultados del aprendizaje que puedan sugerir debilidades del sistema. Por ejemplo, si una evaluación LQAS realizada por un inspector a cinco estudiantes en una aula determina que todos los estudiantes de la muestra son capaces de multiplicar números correctamente, se puede inferir que la mayoría de los estudiantes probablemente sean capaces de multiplicar números correctamente y que el maestro no necesita apoyo para enseñar a multiplicar. En el otro extremo del espectro, las evaluaciones que se vayan a utilizar para diagnosticar brechas de aprendizaje o hacer avanzar a los estudiantes al siguiente nivel requieren medidas de resultados censales.
- **El uso de funcionarios del gobierno local o encuestadores formados a nivel central para recopilar datos:** Recurrir a encuestadores o funcionarios del gobierno locales para llevar a cabo las evaluaciones puede reducir los costos de desplazamiento. Si se hace así, asegúrese de no asignar funcionarios a escuelas que ya estén supervisando, para evitar posibles conflictos de interés que puedan sesgar los resultados.
- **La complejidad de la información requerida:** Aumentar el tiempo dedicado a la recopilación de datos individuales obviamente aumentará los costos. Parte del tiempo invertido puede ser esencial, como medir la comprensión y la fluidez de lectura oral de los estudiantes mediante EGRA. Pero considere el valor marginal agregado de cosas tales como añadir más estudiantes a la muestra o más elementos al instrumento. Es posible que se necesiten más estudiantes solo cuando la evaluación se va a utilizar para promover los propósitos de la enseñanza diferenciada, o cuando el pequeño tamaño de la muestra no genera tendencias perceptibles. Pueden ser necesarios más elementos cuando los resultados sugieren una brecha en una destreza clave, pero no hay suficiente información sobre el motivo (p. ej., en muchos países del África subsahariana, los alumnos tienen dificultades con la comprensión lectora). Es posible que esto se deba a problemas de comprensión del idioma, vocabulario u otra habilidad precursora; añadir elementos que midan estos conocimientos podría ayudar a los funcionarios de la educación a comprender mejor las debilidades del sistema.
- **Paneles rápidos y económicos:** El desarrollo de un sistema de evaluación a gran escala ofrece la oportunidad de compartir datos de forma rápida y económica para apoyar las responsabilidades y la toma de decisiones. Por ejemplo, en Kenia, el Ministerio de Educación está ampliando la utilización de un panel que se introdujo en el marco del programa Tusome (véase la Figura 5), que incluye datos de evaluación de los estudiantes desglosados por parámetros kenianos y datos de las visitas a las aulas. A este panel pueden acceder los funcionarios de la educación a nivel nacional, de condado y local. Esta difusión virtual de datos tiene el beneficio de incurrir solo en costos marginales. Además, los estudios sugieren que la herramienta del panel de Tusome ha estimulado una cultura de uso de datos dentro del gobierno keniano en general.<sup>95</sup>




**FIGURA 5. Ejemplo del panel de datos del gobierno de Kenia<sup>96</sup>**


## MEJORA DE LA EVALUACIÓN A NIVEL DEL SISTEMA

A partir de la comprensión de la importancia y los diferentes tipos de evaluaciones disponibles para su uso en su país, considere las formas en que el sistema de evaluación en su contexto puede mejorar:

### 1 INTEGRAR LA EVALUACIÓN EN EL SISTEMA EDUCATIVO

Los donantes y los socios suelen diseñar sistemas de evaluación que no están integrados en el sistema educativo desde el principio, lo que afecta negativamente a la aceptación del gobierno, al uso de los resultados y a la sostenibilidad. Para evitar este error, los donantes y los socios, en colaboración con el gobierno, deben trabajar para mejorar las evaluaciones existentes o, cuando no existan evaluaciones, legislar la creación, dirección y financiación de programas y agencias de evaluación, incluida la identificación de qué agencia realizará las evaluaciones y cuál utilizará los resultados. Por ejemplo, cuando el gobierno de la India, junto con sus socios, identificó la necesidad de una evaluación escolar censal para diagnosticar las brechas de aprendizaje, legisló la creación de dicha evaluación para 3o., 5o. y 8o. grados en su Política nacional de educación 2020.<sup>97</sup> La fundación Central Square Foundation trabajó entonces con la Junta central de educación secundaria (Central Board of Secondary Education, CBSE) del gobierno para desarrollar la evaluación y garantizar la alineación de los planes de estudios y el Marco de Competencia Global (véase a continuación el Cuadro 4 sobre el establecimiento de parámetros). La CBSE implementará la evaluación, y los resultados se utilizarán a nivel escolar para apoyar la enseñanza diferenciada. Los datos también pueden utilizarse a nivel estatal y nacional para dirigir el apoyo a las escuelas, así como difundirse públicamente (de forma anónima).

A pesar de lo prometedor de la nueva evaluación en la India, en otros contextos a menudo lo más efectivo que se puede hacer no es desplegar nuevas evaluaciones, sino trabajar con el gobierno para introducir ajustes en la difusión de los resultados de las evaluaciones existentes y conectar esto con la implementación y la enseñanza de los programas. Las consideraciones en cuanto a la financiación también deben incluir la reserva de fondos para cualquier intervención que resulte del uso de los hallazgos de la evaluación. Cuando trabaje con el gobierno, garantice un proceso inclusivo que implique la participación de los principales responsables políticos y profesionales (incluidos los maestros) en cada etapa del diseño de la evaluación, la recopilación de datos, el análisis y la difusión de los resultados. Incluir a funcionarios del gobierno en el proceso tiende a reducir costos cuando sustituyen al apoyo externo contratado. Por ejemplo, en Malawi, la actividad YESA<sup>98</sup> de USAID apoya la Evaluación nacional de la lectura del Ministerio de Educación, Ciencia y Tecnología y utiliza a directores distritales de educación, asesores de educación primaria y otros funcionarios del gobierno para recoger los datos (en lugar de encuestadores contratados), y actualmente se está formando a personal del ministerio para analizar y difundir los resultados.

### 2 SER REALISTAS EN CUANTO A COSTOS Y SOSTENIBILIDAD

Las limitaciones financieras pueden dificultar la sostenibilidad de las evaluaciones. Asegúrese de tener en cuenta las repercusiones en los costos mencionadas anteriormente en esta guía y asegúrese de que las decisiones se tomen teniendo claro si el gobierno será capaz de sostener esos costos a lo largo del tiempo. Considere la transición gradual de las evaluaciones financiadas por donantes al gobierno a lo largo de la duración de un programa FLN.



### 3 GARANTIZAR LA ALINEACIÓN DE LAS EVALUACIONES DENTRO DEL SISTEMA EDUCATIVO

Al diseñar o revisar un sistema educativo en coordinación con las principales partes interesadas, es importante garantizar la alineación de las evaluaciones con los componentes clave del sistema. Este elemento clave también se analiza en la guía Enseñanza basada en la evaluación: Nivel del aula. Tal como se ilustra en la Figura 6, debe haber coherencia entre las normas educativas, los parámetros y las evaluaciones, y todos ellos deben estar estrechamente vinculados a los planes de estudios, la formación docente y los materiales. Herramientas como las Encuestas sobre el plan de estudios aprobado<sup>99</sup>—que examinan la relación entre los planes de estudios, la enseñanza y la evaluación—pueden ayudar a determinar qué tan bien se alinea un sistema de evaluación con el sistema educativo y qué brechas deben abordarse. Aunque no siempre sea posible realizar un análisis formal de las Encuestas sobre el plan de estudios aprobado, es fundamental que sus esfuerzos de evaluación examinen cómo una evaluación determinada se alinea con las normas educativas, los planes de estudios y las prácticas pedagógicas existentes. La alineación garantiza que los maestros se mantengan enfocados en el contenido crítico y que los resultados de las evaluaciones reflejen el aprendizaje real en el aula, lo que es fundamental para garantizar una enseñanza basada en la evaluación. El análisis de alineación del sistema se puede realizar mediante una revisión de los planes de estudios y los documentos normativos; entrevistas y discusiones con funcionarios de la educación, administradores y maestros; y observaciones en el aula.

Considere también la alineación de la evaluación del sistema y los planes de estudios con los estándares de competencia global mínima, como se describe en el Cuadro 4, si existe el deseo de utilizar la evaluación para informar sobre los avances hacia el Objetivo de Desarrollo Sostenible 4.1.1, que se centra en la calidad de la educación.

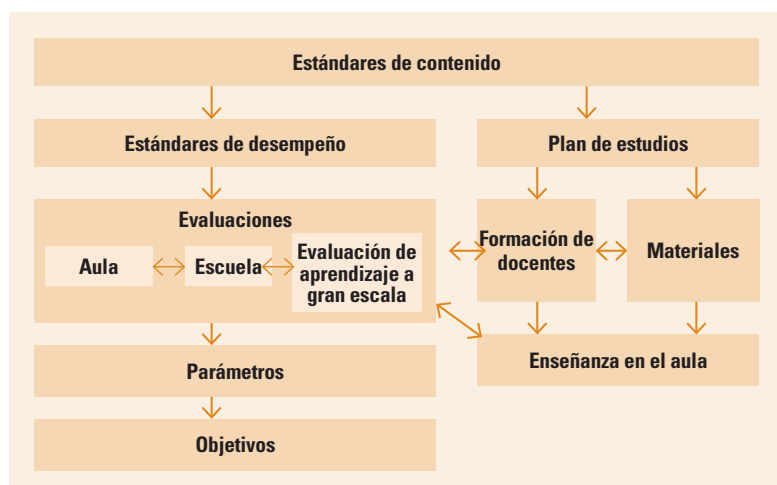
### 4 ACLARAR EL PROPÓSITO DE LAS EVALUACIONES

Otro error clave que los donantes, los gobiernos y los socios suelen cometer es realizar evaluaciones por el solo hecho de evaluar, en lugar de hacerlo para fundamentar la orientación del apoyo y mejorar la enseñanza. En los comienzos del proceso de diseño, trabaje con el gobierno, los donantes, los socios y otras partes interesadas clave para identificar el propósito de las evaluaciones. Al hacerlo, puede ser útil tener presentes las tres funciones básicas del sistema educativo, tal como las identifican Luis Crouch y Joe DeStefano.<sup>100</sup>

- **Establecer expectativas para los resultados del aprendizaje:** Es importante que los sistemas educativos tengan estándares de contenido y de desempeño bien definidos, así como parámetros y objetivos. Sin metas acordadas, los donantes, los gobiernos, los socios y los maestros no saben hacia dónde van ni cómo asignar los recursos para llegar allí. Si el sistema educativo no dispone de parámetros y objetivos que le ayuden a orientarse, colabore con las principales partes interesadas para planificar talleres de fijación de parámetros y objetivos, y garantice una amplia difusión de las metas entre los maestros, las comunidades y otras partes interesadas clave.

Los parámetros deben identificar la calificación o el nivel de desempeño que un estudiante debe alcanzar para ser considerado competente en una determinada destreza (p. ej., al final de 2o. grado, los estudiantes deben ser capaces de alcanzar una calificación de 50 sobre 75 en la Evaluación nacional de lectura). Hay varias formas de establecer parámetros. Por ejemplo, el Instituto de Estadística de la UNESCO y USAID han desarrollado una vinculación de políticas para medir los resultados globales del aprendizaje, descrita en el Cuadro 4 (más arriba). Esto complementa el trabajo existente de la UNESCO para vincular estadísticamente las evaluaciones del aprendizaje con el fin de aumentar la comparabilidad de los resultados entre países y entre diferentes evaluaciones.

**FIGURA 6. Alinear la evaluación dentro del sistema educativo**





Los objetivos deben proporcionar a los países metas intermedias y realistas sobre el porcentaje de una población dada que debe alcanzar el parámetro en un período de tiempo dado (p. ej., a finales de 2021, el 35 % de los estudiantes de 2o. grado deben alcanzar el parámetro). La mejor forma de fijar los objetivos es considerar primero el desempeño actual de los estudiantes y estimar un ritmo de cambio que sea realista pero que impulse al sistema hacia la mejora. El ritmo actual de mejora en los países es demasiado lento para garantizar el logro del objetivo global de que todos los niños alcancen una competencia mínima en lectura y matemáticas para 2030. Es importante que los países busquen formas de acelerar la mejora utilizando una enseñanza basada en la evaluación y que esta meta se refleje en los objetivos de los países—dedicando recursos suficientes para garantizar que la meta sea alcanzable.

**Los enfoques del Diseño universal de evaluación (Universal design for assessment, UDA) pueden ayudar a los ministerios a alcanzar sus parámetros garantizando que las evaluaciones utilizadas sean inclusivas y accesibles para todos (véase el Cuadro 5 sobre UDA).**

Las evaluaciones deben diseñarse teniendo en cuenta el marco del UDA y en estrecha coordinación con las principales partes interesadas.

**Supervisar y responsabilizar a las escuelas por el cumplimiento de esas expectativas:** El gobierno debe supervisar el desempeño general del sistema educativo con respecto a los estándares, los parámetros y los objetivos. La supervisión a nivel del sistema puede realizarse a través de evaluaciones periódicas censales o por muestreo de los resultados de aprendizaje de los estudiantes,<sup>103</sup> que pueden lograrse a través de cualquier LSLA o LQAS. Dicho esto, los exámenes nacionales, cuando están bien diseñados, son las mejores herramientas para responsabilizar a estudiantes y maestros, ya que tienden a utilizar muestras censales. Al mismo tiempo, las evaluaciones comunitarias (como las utilizadas por ASER y UWEZO) tienen la ventaja de proporcionar retroalimentación instantánea a los padres y la comunidad sobre los resultados del aprendizaje. Sin embargo, como se describió anteriormente, la supervisión solo es útil si incita a la acción—en forma de recursos y apoyo a las escuelas que no cumplen con las expectativas.

**Intervenir para apoyar a los estudiantes y las escuelas que tienen dificultades y garantizar la equidad para todos los estudiantes:** Las evaluaciones nacionales, subnacionales, no estandarizadas y dirigidas por ciudadanos—cuando utilizan metodologías de muestreo que permiten la representatividad de los grupos de interés y recogen información demográfica relevante—pueden utilizarse para orientar el apoyo a los grupos más necesitados. Para identificar las escuelas más necesitadas, se necesita una muestra grande. También es importante garantizar que las propias evaluaciones sean inclusivas, ya que, de lo contrario, podrían pasarse por alto grandes grupos de estudiantes y escuelas que necesitan apoyo (véase el Cuadro 5). Los métodos de Enseñanza al nivel adecuado (Teaching at the Right Level, TaRL) también ayudan a garantizar la equidad para todos los estudiantes. El enfoque TaRL utiliza la evaluación del aprendizaje en el sistema para identificar a los estudiantes a fin de agruparlos por su nivel de desempeño y luego adapta la enseñanza al nivel de cada grupo respectivo. Por lo general, la implementación de TaRL exige una importante formación y apoyo continuo a los maestros, pero se ha utilizado con éxito en la India y en diez países africanos.<sup>104</sup>

#### CUADRO 4. Establecimiento de parámetros mediante la vinculación de políticas para medir los resultados globales del aprendizaje

La vinculación de políticas es un proceso que vincula evaluaciones con estándares de contenido y desempeño y establece parámetros para los estándares mínimos de competencia. La metodología involucra a maestros titulares y especialistas en planes de estudios para que actúen como panelistas en talleres en los que emiten juicios fundamentados sobre el desempeño probable de los estudiantes con un nivel mínimo de competencia en una evaluación dada. Las partes interesadas del gobierno han informado de una mejor apropiación de los parámetros establecidos mediante métodos de vinculación de políticas.

Una vez establecidos los parámetros mediante la vinculación de políticas, pueden fijarse nuevos parámetros para futuras evaluaciones utilizando métodos de vinculación de políticas o de vinculación estadística (véase a continuación el cuadro sobre el desarrollo de instrumentos equivalentes).

La vinculación de políticas puede utilizarse junto con el [Marco de Competencia Global](#), que define un conjunto de estándares de contenido y desempeño globales comunes adoptados por el Instituto de Estadística de la UNESCO, para establecer parámetros que vinculen las evaluaciones con el Objetivo de Desarrollo Sostenible 4.1.1, que se centra en la calidad de la educación.<sup>101</sup> Los países también pueden utilizar métodos estadísticos como LAYS para vincular sus evaluaciones nacionales o subnacionales con una evaluación regional o internacional que ya esté vinculada al ODS 4.1.1.



## 5 CENTRAR LAS EVALUACIONES EN LO QUE LOS MAESTROS Y LOS RESPONSABLES POLÍTICOS NECESITAN SABER

**Enfoque el diseño de la evaluación en recoger solo la información necesaria para mejorar la enseñanza y, por lo tanto, los resultados del aprendizaje.** Algunas ideas sobre cómo simplificar las evaluaciones, aumentar su efectividad y reducir costos, garantizando al mismo tiempo la calidad, son las siguientes:

- **Medir las destrezas importantes:** Asegúrese de que las destrezas abordadas en la evaluación se alineen con los estándares de contenido y desempeño. Si un sistema educativo aspira a desarrollar destrezas de orden superior (p. ej., comprensión, pensamiento crítico y resolución de problemas), las evaluaciones en el sistema deben incluir al menos alguna medida de esas destrezas, sobre todo porque las evaluaciones a menudo determinan lo que se enseña en clase. Si bien es fácil diseñar elementos que requieran automaticidad o memoria (como nombrar los sonidos de las letras o resolver un simple problema de suma), es más difícil diseñar elementos que requieran comprensión, pensamiento crítico e integración de destrezas, razón por la cual muchas LSLA no miden las destrezas de orden superior en absoluto o no las miden de manera efectiva. Las destrezas de orden superior suelen evaluarse mejor con elementos de respuesta abierta que con elementos de respuesta cerrada (p. ej., opciones de verdadero-falso y opción múltiple).
- **Garantizar que las herramientas midan las destrezas previstas:** En las evaluaciones de FLN, considere el uso de elementos orales (como los utilizados en EGRA, EGMA y evaluaciones comunitarias) para que la capacidad de lectura no se confunda con las habilidades matemáticas, y la capacidad de decodificación no se confunda con la comprensión. Por ejemplo, al responder a los elementos de una prueba escrita, los niños pueden equivocarse en un elemento porque no pudieron leer las instrucciones, siendo que de otro modo hubieran sido capaces de responder correctamente. Considere, por ejemplo, un problema escrito de matemáticas que requiere que el alumno lo lea y luego lo resuelva. Si el niño resuelve el problema incorrectamente, es imposible saber si se debe a una brecha en sus habilidades de lectura o matemáticas. De modo similar, las preguntas no deben exigir información previa innecesaria que pueda no reflejar el conocimiento cultural típico y la enseñanza en el aula (p. ej., la necesidad de saber que la nieve es blanca en un contexto en el que nunca nieva o de estar familiarizado con un alimento o animal no común en el contexto).
- **Establecer un nivel adecuado al grado y la capacidad de los estudiantes:** Muy pocos estudiantes deberían responder bien a todos los elementos o mal a todos los elementos de una evaluación, ya que esto limita la utilidad de la información de la evaluación. Para evitar este problema, busque la ayuda de los maestros para desarrollar o perfeccionar las pruebas, y realice pruebas piloto de los instrumentos para ver si la evaluación tiene efectos de techo o piso. Por ejemplo, si las EGRA y EGMA que se están utilizando en una prueba piloto arrojan una gran cantidad de calificaciones cero, puede ser necesario añadir más subsecciones básicas y considerar la posibilidad de suprimir las subsecciones más avanzadas—o bien aplicar una regla para detenerse que pida a los niños que no pasen a la siguiente sección, si tienen calificación cero en secciones anteriores.

### CUADRO 5. Diseño universal de evaluación<sup>102</sup>

Al diseñar o perfeccionar las evaluaciones, es fundamental garantizar la inclusión. Las evaluaciones deben diseñarse para que sean universales—accesibles para todos—desde el principio. Esto implica:

- **Eliminar las barreras a la evaluación**—Por ejemplo, elimine las barreras cognitivas, sensoriales, emocionales y físicas (como pedir a los alumnos que señalen con el dedo, sin dar opción a quienes no pueden hacerlo).
- **Crear elementos accesibles desde el principio**—Por ejemplo, no haga referencia a colores, que no son accesibles para alumnos ciegos, o a música que no es accesible para alumnos sordos.
- **Diseñar teniendo en cuenta las adaptaciones**—Considere las discapacidades físicas, intelectuales y de aprendizaje.
- **Utilizar instrucciones y un lenguaje claros y sencillos**

#### Adaptaciones del UDA para recibir información

- Hacer que alguien lea el texto en voz alta o permitir grabaciones de audio.
- Incluir menos elementos en una página o utilizar letra grande o braille.
- Entregar a los alumnos una lista escrita de instrucciones.

#### Adaptaciones del UDA para expresar información

- Permitir respuestas orales o el uso de dispositivos de comunicación alternativos y aumentativos o lenguaje de señas.
- Permitir el uso de calculadora o tablas matemáticas.
- Permitir respuestas escritas en braille.

#### Adaptaciones del UDA para la motivación

- Dar tiempo adicional para completar la prueba.
- Permitir que las pruebas se realicen en varias sesiones, dejando tiempo para los descansos.
- Permitir que las pruebas se realicen en los momentos del día en que un alumno determinado esté más concentrado.
- Permitir que las pruebas se realicen en entornos alternativos con menos alumnos y distracciones.





- **Incluir el número adecuado de elementos para el propósito:** Las evaluaciones largas cansan a los estudiantes y pueden disminuir su desempeño. Utilice solo el número de elementos necesarios para el propósito. Las evaluaciones utilizadas para promocionar a los estudiantes o diferenciarlos en grupos de capacidades pueden tener más elementos que las evaluaciones centradas en supervisar la efectividad del sistema educativo en su conjunto. La prueba piloto ayudará a determinar si es necesario modificar el número de elementos.
- **Incluir información contextual que realmente se vaya a utilizar:** Los responsables políticos necesitan más datos que solo los resultados del aprendizaje (véase el Cuadro 7). Sin embargo, muchos sistemas de evaluación recogen información contextual que nunca es utilizada por los responsables políticos u otras partes interesadas. Por ello, para reducir costos, es importante estudiar detenidamente qué datos adicionales recoger con una evaluación. Alguna información que ha resultado útil incluye (1) los idiomas que hablan los niños y el maestro (véase la [guía Elecciones prácticas de idioma](#) para un mayor análisis de este y otros temas relacionados con el idioma de enseñanza); (2) si el maestro ha recibido formación; (3) si los niños tienen el libro de texto o el cuaderno de ejercicios; (4) si el maestro tiene y utiliza una guía del maestro; (5) si los maestros implementan en el aula el programa FLN y las prácticas pedagógicas presentadas durante la formación.
- **Mantener la comparabilidad en el tiempo:** Si está previsto que los instrumentos de evaluación cambien de un año a otro, pero aun así piensa utilizarlos para comparar resultados y realizar un seguimiento de los resultados del aprendizaje a lo largo del tiempo, deberá crear evaluaciones comparables. Véase el Cuadro 6 anterior para obtener consejos al respecto.

#### CUADRO 6. Desarrollo de instrumentos equivalentes<sup>105</sup>

Para garantizar la confianza de las partes interesadas en los resultados de la evaluación del aprendizaje y en la comparación de los resultados a lo largo del tiempo, debe tener en cuenta lo siguiente:

**Uno o varios instrumentos:** Los instrumentos solo deben cambiarse de un año a otro si se teme por filtraciones de las pruebas. De lo contrario, puede utilizarse la misma evaluación a lo largo del tiempo, suponiendo que siga siendo contextualmente relevante. Si se necesitan varias versiones de la evaluación, asegúrese de que tengan el mismo nivel de dificultad.

**Garantizar una dificultad similar:** Los elementos que se centran en nombrar letras o números pueden seguir siendo los mismos, pero en distinto orden. Los pasajes de lectura pueden modificarse simplemente cambiando los nombres, temas de las historias, acciones y adjetivos por equivalentes de nivel del grado. Se pueden utilizar parámetros o criterios para garantizar que los problemas de matemáticas puedan ajustarse sin cambiar el nivel de dificultad.

**Equiparación:** A pesar de haber creado instrumentos lo más similares posible, es inevitable que haya pequeñas diferencias de dificultad. Estas diferencias pueden abordarse haciendo que una muestra de estudiantes representativa del país realice ambas evaluaciones o incluyendo elementos de anclaje comunes, y luego utilizando esa información para hacer que los resultados sean equivalentes.

## 6 DECIDIR CUÁNDO Y CON QUÉ FRECUENCIA RECOGER DATOS DE LAS EVALUACIONES

La frecuencia y el calendario de las evaluaciones deben basarse en su propósito y en la financiación disponible. Las evaluaciones escolares, destinadas a orientar y dirigir el apoyo, deben realizarse varias veces al año. Las evaluaciones centradas en supervisar los cambios en el sistema solo son necesarias cada dos años. Las evaluaciones destinadas a comparar los resultados entre países pueden realizarse incluso con menos frecuencia, a intervalos de cuatro o cinco años. Las evaluaciones centradas en el seguimiento de los resultados y su comparación con los parámetros deben realizarse en el mismo período del año, preferentemente al final del año escolar.<sup>107</sup>

## 7 ENCONTRAR UN EQUILIBRIO ENTRE RESPONSABILIDAD Y APRENDIZAJE

La evaluación educativa debe lograr un delicado equilibrio: si no hay responsabilidades asociadas a los resultados de la evaluación, las evaluaciones harán poco por el cambio; en cambio, si hay demasiado en juego, las evaluaciones pueden dar lugar a “enseñanza para la prueba”, a la filtración de las pruebas o a que los maestros pidan a los estudiantes con bajo desempeño que se queden en casa el día de la prueba.<sup>108</sup> **La evaluación debe promover la mejora, ya sea a través de mecanismos de aprendizaje o responsabilización; debe utilizarse para ayudar a los programas a adaptarse.** Para mejorar la responsabilización por los resultados y reducir los incentivos para hacer trampa, considere la posibilidad de implementar las siguientes medidas:

- **Para centrarse en el aprendizaje, difunda los resultados a nivel de la escuela y los maestros,** abordando a maestros y funcionarios de la educación con información que puedan utilizar para reformar las prácticas, en lugar de hacer énfasis en los informes de clasificación de las escuelas. Ghana y Kenia, por ejemplo, han reducido el uso de este tipo de informes, logrando así una reducción efectiva de muchos de los efectos negativos de sus exámenes nacionales. Ghana introdujo tablas de clasificación (o liga escolar) en 2004,



pero eso provocó un fuerte aumento de las malas prácticas. Cuando eliminó el uso de estas tablas, se redujeron las trampas. La limitación de la publicación de resultados de exámenes a nivel de distrito al final del nivel básico de escolarización se ha asociado con un impacto positivo en algunos distritos.<sup>109</sup> Debe haber un cuidadoso equilibrio entre la responsabilidad por los resultados y la creación de incentivos para el fraude. En algunos países, tener acceso a datos disponibles públicamente aumenta las responsabilidades, pero en otros puede inducir comportamientos negativos. Es importante comprender si los resultados disponibles ayudan o perjudican.

- **Brinde apoyo a escuelas y maestros y luego hágalos responsables de las medidas que toman** en base a los resultados de la evaluación, recurriendo a la mentoría, otros tipos de apoyo continuo e inspectores.
- **Utilice incentivos para la mejora en lugar de sanciones** (y tenga en cuenta factores como el origen socioeconómico de los estudiantes, los recursos de la escuela, etc., que pueden estar afectando a las calificaciones, utilizando la evaluación para identificar formas de satisfacer las necesidades de las poblaciones que se ven desproporcionadamente afectadas por estos desafíos; la idea es garantizar que el sistema satisfaga las necesidades incluso de las escuelas y poblaciones más desfavorecidas). Por ejemplo, la entrega de certificados y el reconocimiento cuando mejoran las calificaciones, o el aporte de recursos adicionales para ayudar a las escuelas con dificultades, pueden ser efectivos para impulsar la mejora.

## 8 GARANTIZAR UNA DIFUSIÓN SELECTIVA

La difusión efectiva es uno de los aspectos más importantes de un sistema de evaluación de alta calidad. Sin embargo, la difusión de los resultados de las evaluaciones, en particular de las LSLA, no suele ser efectiva. Algunas sugerencias para garantizar una difusión selectiva son las siguientes:

- **Administración escolar y maestros:** Para que las evaluaciones a gran escala puedan maximizar sus beneficios, los administradores escolares y los maestros necesitan información específica sobre los resultados de los alumnos, qué elementos tienden a responder bien o mal los estudiantes, cuál fue la respuesta incorrecta más común, y las formas sugeridas para abordarlos, como se muestra en el Cuadro 9.<sup>111</sup> El apoyo continuo a los maestros en la implementación de esas acciones es primordial. Además, los datos deben llegar a tiempo a las escuelas y a los maestros, no mucho después (en el plazo de unos meses) de que haya finalizado una evaluación. Esto puede lograrse mediante el uso de paneles en vivo para difundir los resultados de las evaluaciones sistemáticas. La iniciativa RAMP de Jordania, por ejemplo, utilizó con éxito herramientas gubernamentales de recopilación de datos para proporcionar a las escuelas datos específicos de la escuela sobre el desempeño de los estudiantes, que complementó con reuniones comunitarias a nivel de escuela para discutir los resultados y sugerir estrategias para responder a los desafíos en la enseñanza planteados por maestros y padres por igual.<sup>112</sup>
- **Funcionarios de educación descentralizada:** Al igual que los demás grupos de interesados, los funcionarios de educación a nivel regional o de distrito necesitan datos presentados en formatos concisos, fáciles de usar y oportunos. También pueden necesitar orientación sobre el uso de los datos para dirigir el apoyo y los recursos a las escuelas y los maestros.

## CUADRO 7. Desglose de los resultados de la evaluación<sup>106</sup>

Los datos de la evaluación en el sistema deben desglosarse por factores como:

- Sexo
- Etnia
- Región geográfica
- Escuelas urbanas y rurales
- Alumnos con discapacidades (esta información debe basarse en un riguroso proceso de selección)
- Otras características clave específicas del contexto

Esta información debe utilizarse para dirigir el apoyo a las escuelas y a los estudiantes en función de los resultados de las evaluaciones por población, así como para considerar si el plan de estudios y los enfoques pedagógicos son inclusivos para todos los grupos.

## CUADRO 8. Difusión de resultados escolares en India<sup>110</sup>

En el estado de Maharashtra en India, todas las escuelas completan una evaluación escolar que incluye evaluaciones de los alumnos y otros componentes de evaluación contextual. Luego, las escuelas reúnen a directores de escuela, maestros, estudiantes, padres y miembros del comité de gestión escolar para facilitar la planificación de la mejora escolar. Los comités de gestión escolar elaboran informes de evaluación (incluida una lista de áreas de mejora) a través de un panel digital centralizado, que es consolidado y utilizado por los funcionarios a nivel de grupo, bloque, distrito, estado y nacional para identificar las necesidades específicas de cada escuela y las áreas comunes de intervención que podrían mejorar el desempeño escolar. Los inspectores escolares hacen un seguimiento de los planes de mejora para supervisar el progreso durante sus visitas.



- La Actividad de aprendizaje de USAID en Ghana, por ejemplo, desarrolló paneles en vivo que los funcionarios de educación distritales y regionales revisaban dos veces por trimestre y luego utilizaban para orientar los planes de acción distritales para mejorar el apoyo a las escuelas. Los paneles se crearon utilizando el software Power Bi e incluían registros escolares introducidos por directores de escuela y maestros, como índices de ausentismo, información de los supervisores de circuito procedente de visitas a las escuelas y observaciones en el aula, y datos de las evaluaciones de PACE. El resultado eran unas simples pantallas que mostraban los distritos y escuelas que necesitaban más apoyo (rojo), algo de apoyo (amarillo) o un apoyo limitado (verde), como se muestra en la Figura 7 anterior. El equipo que creó el panel recomienda colaborar estrechamente con los funcionarios de educación locales para adaptar la difusión y completar varias rondas de pruebas y retroalimentación de los usuarios.<sup>113</sup> Las conversaciones con los funcionarios de la educación que utilizan el panel de la Figura 7 se centrarían en animar a los maestros a continuar con las mejores prácticas identificadas en el panel, pero aumentando el número de sesiones de mentoría por maestro y las reuniones escolares que se realizan.

### CUADRO 9. Ejemplo de retroalimentación detallada sobre el desempeño de los estudiantes en la evaluación

Elemento 1:  $45 - 5 \times 4 =$

a. 160      c. 25

b. 20      d. 16

**Respuesta correcta**

c. 25 (dada por el 15% de los estudiantes)

**Respuesta incorrecta más frecuente**

a. 160 (60%)

**Motivo del error**

Falta de dominio del concepto de orden de las operaciones

**Recomendación**

Repasar y practicar el orden de las operaciones con los estudiantes

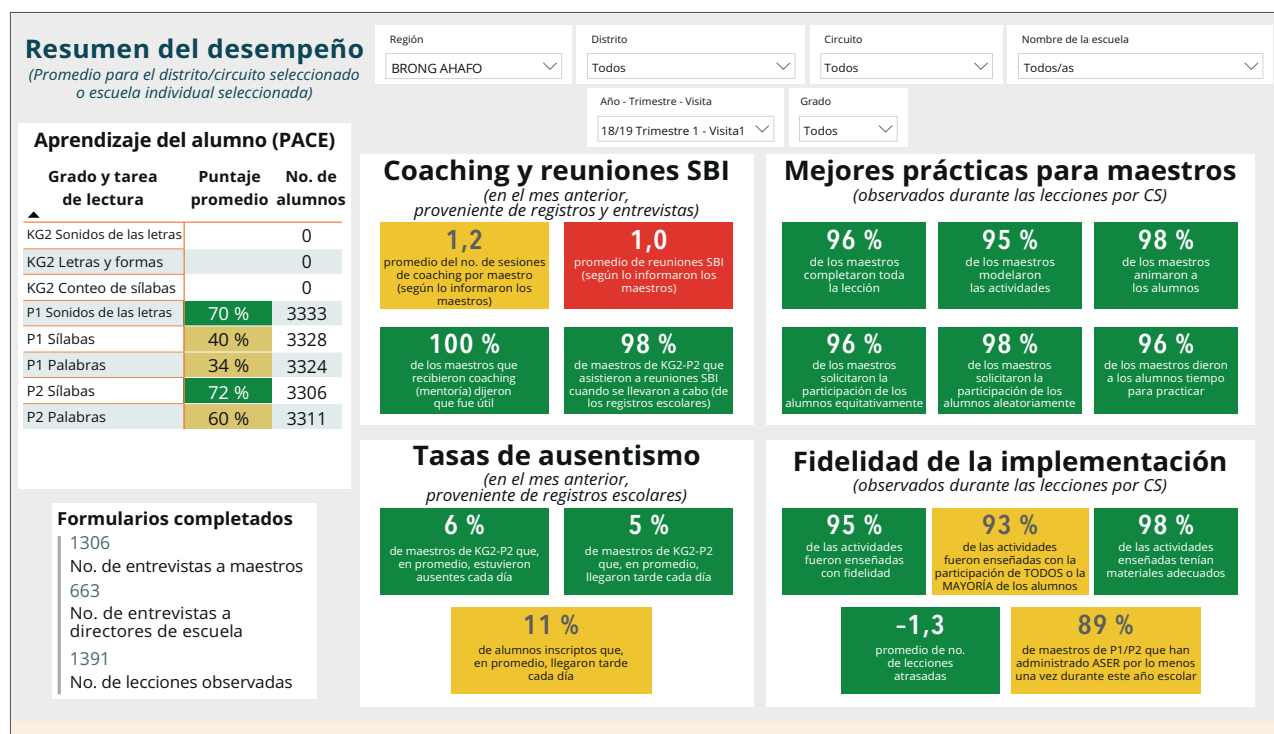
- Responsables políticos nacionales:** Además de utilizar los datos para orientar su apoyo a las escuelas, los responsables políticos nacionales y los funcionarios de la educación también pueden utilizar los datos de las evaluaciones como base para los cambios en los planes de estudios o las políticas educativas, o para establecer parámetros y objetivos. De hecho, en varios países se ha cuestionado la idoneidad de los planes de estudios a raíz de los hallazgos de las evaluaciones nacionales. Por ejemplo, un examen nacional en Bután llevó al Ministerio de Educación a reformar el plan de estudios de matemáticas del país tras determinar que los estudiantes estaban experimentando una sobrecarga curricular; y el gobierno de Etiopía introdujo cambios en sus planes de estudios tras recibir retroalimentación de maestros, estudiantes y padres sobre las razones de los malos resultados en las pruebas.<sup>114</sup>
- Los padres y la comunidad:** Los padres y las comunidades están bien posicionados para hacer oír su voz cuando se trata de abogar por mejoras escolares; por lo tanto, es importante identificar formas de compartir los resultados con ellos. En Nepal y la India, las reuniones comunitarias de los pueblos y las campañas de información se han utilizado como canales para difundir los resultados de las evaluaciones. Dichas campañas han recurrido a la divulgación en la comunidad para aumentar la conciencia sobre los roles y responsabilidades de los comités de supervisión y gestión escolar y para difundir información de forma repetida, por ejemplo mediante festivales, pósteres, pinturas, calendarios para llevar a casa y folletos de evaluación del aprendizaje. La participación de los padres y la comunidad generada a través de estas campañas ha impulsado mejoras en los resultados del aprendizaje. Al difundir información a las comunidades, es probable que un solo punto de contacto sea insuficiente, y las acciones clave que la comunidad puede llevar a cabo deben ser parte integrante del mensaje.<sup>115</sup>
- Medios de comunicación:** Los medios de comunicación pueden contribuir a que el gobierno y las partes interesadas se tomen en serio los resultados de las evaluaciones, presionando al sistema para que introduzca cambios. Dicho esto, hay que dar prioridad a garantizar que los resultados lleguen primero a los usuarios clave (con detalles sobre las acciones específicas que pueden tomar), ya que difundir los resultados solo a los medios de comunicación no generará cambios. Los resultados deben compartirse en tantos medios como sea posible (p. ej., televisión, radio, redes sociales, WhatsApp) para llegar a un público amplio. Los medios de comunicación también pueden ser útiles para celebrar y promover los éxitos con el fin de incentivar a otras escuelas a mejorar, mediante el reconocimiento de las escuelas que más han mejorado por región y de otros hitos. Este también es un uso posible de los resultados para los funcionarios de educación locales y nacionales.



## Conclusión

La evaluación permite a las escuelas, los gobiernos y otras partes interesadas comprender qué tan bien están aprendiendo los estudiantes e identificar dónde se necesita mejorar. Un sistema de evaluación que esté bien alineado con los parámetros y estándares curriculares puede proporcionar información sobre lo que funciona para aumentar los resultados del aprendizaje, cómo podría mejorarse la enseñanza y cómo puede orientarse mejor el apoyo. Comprender los diferentes tipos de evaluación disponibles y el modo en que deben funcionar juntos en un sistema coherente, y tener en cuenta las recomendaciones establecidas en esta guía, le permitirá avanzar hacia un sistema de evaluación que contribuya mejor a la mejora de los resultados del aprendizaje.

**FIGURA 7. Panel de la Actividad de aprendizaje en Ghana**







## RECURSOS

### Análisis del sistema de evaluación

- Datos SABER del Banco Mundial: <https://saber.worldbank.org/index.cfm?indx=5&sub=2> (véase la sección sobre Evaluación de los estudiantes)
- Análisis de los sistemas nacionales de evaluación del aprendizaje (Analysis of National Learning Assessment Systems, ANLAS) de la Asociación mundial para la educación (Global Partnership for Education, GPE): <https://www.globalpartnership.org/sites/default/files/docs/2020-11-GPE-ANLAS-manual.pdf>

### Establecimiento de parámetros

- Herramientas de vinculación de políticas para medir los resultados globales del aprendizaje: <https://www.edu-links.org/resources/toolkit-setting-internationally-linked-benchmarks-early-grade-reading-and-math>

### Alineación del sistema educativo

- Encuestas sobre el plan de estudios aprobado: <https://curriculumanalysis.org/products-SEC.asp#>

### Implementación de una evaluación nacional del aprendizaje

- *Guía rápida núm. 3 de la UNESCO Implementación de una evaluación nacional del aprendizaje:* <http://uis.unesco.org/sites/default/files/documents/quick-guide-3-implementing-national-learning-assessment.pdf>

### Implementación de un examen nacional

- *Exámenes públicos del Banco Mundial analizados:* <https://openknowledge.worldbank.org/bitstream/handle/10986/32352/9781464814181.pdf?sequence=2&isAllowed=y>

### Implementación de una EGRA/EGMA

- Kit de herramientas de EGRA y kit de herramientas de EGRA 2.0: [https://pdf.usaid.gov/pdf\\_docs/PA00M4TN.pdf](https://pdf.usaid.gov/pdf_docs/PA00M4TN.pdf) y [https://www.globalreadingnetwork.net/sites/default/files/media/file/EGMA\\_Toolkit\\_FINAL\\_0.pdf](https://www.globalreadingnetwork.net/sites/default/files/media/file/EGMA_Toolkit_FINAL_0.pdf)



### EXPERIENCIA TÉCNICA NECESARIA

- Desarrollo y análisis de evaluaciones, para apoyar el desarrollo, la equiparación, el análisis de las evaluaciones y el establecimiento de parámetros.
- Enseñanza de la lectura o las matemáticas, con especial atención a los contenidos y la enseñanza de los primeros grados, a fin de apoyar el desarrollo de las evaluaciones para abordar adecuadamente las destrezas clave.



## NOTAS FINALES

- 1 UNESCO Institute for Statistics, "More Than One-Half of Children and Adolescents Are Not Learning Worldwide," Fact Sheet No. 46 (September 2017); United Nations, The Sustainable Development Goals Report 2019 (New York: United Nations, 2019).
- 2 Marguerite Clark, "Measuring Learning: How Effective Student Assessment Systems Can Help Achieve Learning for All," Education Notes 67346 (World Bank, 2012); OECD, PISA 2018 Results (Volume VI): Are Students Ready to Thrive in an Interconnected World? (Paris: OECD, 2020).
- 3 Girindre Beeharry, "Accelerating Foundational Literacy and Numeracy through Focus, Measurement, Support and Accountability," World Bank Blogs (October 19, 2020).
- 4 Phil Elks, "The Impact of Assessment Results on Education Policy and Practice in East Africa," DFID Think Piece (January 2016).
- 5 John Cresswell, Ursula Schwantter, and Charlotte Waters, A Review of International Large-Scale Assessments in Education: Assessing Component Skills and Collecting Contextual Data (Paris: OECD, 2015).
- 6 Ieva Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," IIER-UNESCO Working Paper (April 2019); Ina V. S. Mullis, TIMSS 2019 Assessment Frameworks: Introduction (Amsterdam: IEA, 2017)
- 7 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019; OECD, PISA 2018 Results, 2020.
- 8 Ibid.
- 9 Ibid.
- 10 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019.
- 11 Ibid.
- 12 Ibid.
- 13 Ibid.
- 14 Ibid.; OECD, PISA 2018 Results, 2020; Clark, "Measuring Learning," 2012.
- 15 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019; Mullis, TIMSS 2019 Assessment Frameworks, 2017.
- 16 Ibid.
- 17 Ibid.
- 18 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019; OECD, PISA 2018 Results, 2020; Clark, "Measuring Learning," 2012.
- 19 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019.
- 20 Ibid.
- 21 Ibid.
- 22 Kate Anderson, "Strengthening Learning Assessment Systems," Knowledge and Innovation Exchange Discussion Paper (July 2019).
- 23 Ibid.
- 24 Ibid.
- 25 Ibid.
- 26 Australian Council for Educational Research, "The Latin-American Laboratory for Assessment of the Quality of Education: Measuring and Comparing Educational Quality in Latin America," Assessment GEMS Series, No. 3 (March 2014).
- 27 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019; OECD, PISA 2018 Results, 2020; Clark, "Measuring Learning," 2012.
- 28 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019.
- 29 Ibid.
- 30 Ibid.
- 31 Australian Council for Educational Research, "The Latin-American Laboratory for Assessment of the Quality of Education," 2014.
- 32 Ibid.
- 33 Ibid.
- 34 Ibid.
- 35 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019; OECD, PISA 2018 Results, 2020; Clark, "Measuring Learning," 2012.
- 36 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019.
- 37 Ibid.
- 38 Ibid.
- 39 Vincent Greaney and Thomas Kellaghan, Assessing National Achievement Levels in Education (Washington, DC: World Bank, 2008).
- 40 Anderson, "Strengthening Learning Assessment Systems," 2019.
- 41 Ibid.
- 42 Ibid.
- 43 Una revisión de 2018 de la educación en el África subsahariana recomendó la eliminación de los exámenes de egreso de primaria, dado que son una barrera para el acceso a las escuelas secundarias.
- 44 Anderson, "Strengthening Learning Assessment Systems," 2019.
- 45 Thomas Kellaghan, Vincent Greaney, and T. Scott Murray, Using the Results of a National Assessment of Educational Achievement (Washington, DC: World Bank, 2009).
- 46 Las evaluaciones subnacionales son una variante de las evaluaciones nacionales en las que la muestra se limita a una región, distrito, provincia o estado dentro de un país. Estas evaluaciones suelen llevarse a cabo en sistemas grandes y descentralizados con diferentes políticas educativas o idiomas de evaluación. Las evaluaciones no suelen estar vinculadas entre las distintas zonas de un país, lo que dificulta las comparaciones directas del desempeño de los estudiantes entre regiones (Greaney and Kellaghan, Assessing National Achievement Levels, 2008).
- 47 Kellaghan, Greaney, and Murray, Using the Results of a National Assessment of Educational Achievement, 2009.
- 48 Ibid.
- 49 Ibid.
- 50 Ibid.
- 51 Amber Gove, Early Grade Reading Assessment (EGRA) Toolkit, Second Edition (Washington, DC: USAID, 2015).
- 52 Ibid.; Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019; Thomas Kellaghan and Vincent Greaney, Public Examinations Examined (Washington, DC: World Bank, 2020).
- 53 Ibid.
- 54 Ibid.
- 55 Gove, Early Grade Reading Assessment (EGRA) Toolkit, 2015; Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019.
- 56 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019.
- 57 Ibid.
- 58 Ibid.
- 59 Kellaghan, Greaney, and Murray, Using the Results of a National Assessment of Educational Achievement, 2009; ASER Centre, Annual Status of Education Report (Rural) 2018 (New Delhi: ASER Centre, 2019)



- 60 ASER Centre, Annual Status of Education Report (Rural) 2018, 2019.
- 61 Ibid.
- 62 Ibid.
- 63 Ibid.
- 64 PAL Network, ICAN: International Common Assessment of Numeracy; Background, Features and Large-Scale Implementation (Nairobi: People's Action for Learning Network, 2020).
- 65 Ibid.
- 66 Ibid.
- 67 Ibid.
- 68 UNICEF, "Acerca de MICS" (2019), <https://mics.unicef.org/about>.
- 69 Ibid.
- 70 Ibid.
- 71 Ibid.
- 72 Raudonyte, "Use of Learning Assessment Data in Education Policy-Making," 2019.
- 73 Ibid.
- 74 Ibid.
- 75 Ibid.
- 76 Jacobus Cilliers, Isaac M. Mbiti, and Andrew Zeitlin, "Can Public Rankings Improve School Performance? Evidence from a Nationwide Reform in Tanzania," *Journal of Human Resources* (March 2020).
- 77 Daniel Wagner, Marlaime Lockheed, Ina Mulis, et al., "The Debate on Learning Assessments in Developing Countries," *Compare: A Journal of Comparative and International Education* 42, no. 3 (2012): 509–545.
- 78 ACDP Indonesia, "The Critical Importance of Early Grade Reading and Assessment," working paper (December 2014).
- 79 Adapted from Uganda Literacy Achievement and Retention Activity, EGRA and Supplementary Data Collection Instruments, <https://shared.rti.org/content/ugandalara-egra-and-supplementary-data-collection-instruments>.
- 80 RTI International, Early Grade Mathematics Assessment (EGMA) Toolkit (March 2014), [https://ierc-publicfiles.s3.amazonaws.com/public/re-sources/EGMA%20Toolkit\\_March2014.pdf](https://ierc-publicfiles.s3.amazonaws.com/public/re-sources/EGMA%20Toolkit_March2014.pdf).
- 81 RTI International, An Addition to the Toolbox for Measuring Literacy Skills of the Youngest Students: The Group Administered Literacy Assessment (GALA), [https://ierc-publicfiles.s3.amazonaws.com/public/resources/1071\\_GALA\\_brief\\_02-16-17\\_r6.pdf](https://ierc-publicfiles.s3.amazonaws.com/public/resources/1071_GALA_brief_02-16-17_r6.pdf).
- 82 RTI International, USAID's Early Grade Reading Program (EGRP) in Nepal: Final Progress Report, Years 1–5: March 2015–October 2020 (Washington, DC: USAID, 2020), [https://pdf.usaid.gov/pdf\\_docs/PA00X6M7.pdf](https://pdf.usaid.gov/pdf_docs/PA00X6M7.pdf).
- 83 RTI International, Measurement and Research Support to Education Strategy Goal 1 (Washington, DC: USAID, 2016), <https://shared.rti.org/content/development-and-pilot-testing-additional-subtasks-early-grade-reading-assessment-egra-20>.
- 84 Uwezo Uganda, Are Our Children Learning? Annual Learning Assessment Project (Kampala: Uwezo Uganda, 2014).
- 85 Aser Centre, "Reading Assessment Module," [http://img.asercentre.org/docs/Aser%20survey/Tools\\_Testing/Reading/english.jpg](http://img.asercentre.org/docs/Aser%20survey/Tools_Testing/Reading/english.jpg).
- 86 Molly Jamieson Eberhardt, Bringing Learning to Light: The Role of Citizen-led Assessments in Shifting the Education Agenda (Washington, DC: Results for Development Institute, 2015).
- 87 Brian Wilcox, Making School Inspection Visits More Effective: The English Experience (Paris: UNESCO, 2000).
- 88 CAMFED, "Our Data, Our Destiny: Young Women Deliver Community Accountability to Girls" (May 18, 2016), <https://camfed.org/latest-news/young-women-deliver-data-accountability/> <https://camfed.org/latest-news/young-women-deliver-data-accountability/>.
- 89 RTI International, USAID's Early Grade Reading Program (EGRP) in Nepal, 2020.
- 90 Jennifer Bowers, interview with Soma Omenye staff (2021).
- 91 Melissa Chiappetta, interview with Jeremy Koch, Chief of Party for UNICEF/USAID Yesani Unuzira Activity (January 2021).
- 92 Wilcox, "Making School Inspection Visits More Effective," 2000.
- 93 Henry Braun, Anil Kanjee, Eric Bettinger, and Michael Kremer, Improving Education through Assessment, Innovation, and Evaluation (Cambridge, MA: American Academy of Arts and Sciences, 2006).
- 94 LQAS es una metodología de muestreo que utiliza tamaños de muestra pequeños para clasificar las áreas locales según cumplan o no cumplan los estándares mínimos de desempeño. Véase Herramientas para el enfoque de supervisión de la educación para obtener más información: <https://shared.rti.org/content/local-education-monitoring-approach-lemma-toolkit>.
- 95 Elks, "The Impact of Assessment Results," 2016; Raudonyte, "Use of Learning Assessment Data," 2019.
- 96 Evelyn Jepkemei, "The Power of Instructional Support: Using Existing Systems to Change Teacher Behavior in Kenya," presentation for the USAID Tusome Early Grade Reading Activity.
- 97 Interview with Central Square Foundation, 2021.
- 98 Interview with Jeremy Kock, Abt Associates (2021).
- 99 Julius Atuhurra and Michelle Kaffenberger, "System (In)Coherence: Quantifying the Alignment of Primary Education Curriculum Standards, Examinations, and Instruction in Two East African Countries," RISE Working Paper 20/057 (December 2020).
- 100 Luis Crouch and Joe DeStefano, "A Practical Approach to In-Country Systems Research" (June 2015).
- 101 Management Systems International, Policy Linking for Measuring Global Learning Outcomes Toolkit: Linking Assessments to the Global Proficiency Framework (USAID, UNESCO, Global Alliance to Monitor Learning, 2020).
- 102 Anne Hayes, Ann Turnbull, and Norma Moran, Universal Design for Learning to Help All Children Read: Promoting Literacy for Learners with Disabilities (Washington, DC: USAID, 2018).
- 103 Luis Crouch and Joe DeStefano, "A Practical Approach to In-Country Systems Research" (June 2015).
- 104 Teaching at the Right Level, "TaRL in Action," <https://www.teachingattherightlevel.org/tarl-in-action>.
- 105 Gove, Early Grade Reading Assessment (EGRA) Toolkit, 2015.
- 106 Ibid.
- 107 María José Ramírez, Quick Guide No. 3: Implementing a National Learning Assessment (Quebec: UNESCO Institute for Statistics, 2017).
- 108 Braun et al., Improving Education, 2006.
- 109 Kellaghan, Greaney, and Murray, Using the Results of a National Assessment of Educational Achievement, 2009.
- 110 Ankita Bhatkande, "Schools in Maharashtra Asked to Do Self-Assessment under Centre's Initiative," *Hindustan Times* (November 24, 2019).
- 111 Kellaghan, Greaney, and Murray, Using the Results of a National Assessment of Educational Achievement, 2009.
- 112 RTI International, USAID's Early Grade Reading and Mathematics Initiative (RAMP) in Jordan: Year 3 Progress Report; October 2016–September 2017 (Washington, DC: USAID, 2017), [https://pdf.usaid.gov/pdf\\_docs/PA00TJ7P.pdf](https://pdf.usaid.gov/pdf_docs/PA00TJ7P.pdf).
- 113 Yvonne Cao, "Interactive Dashboards for Adaptive Management: Enabling Differentiated School Support across 100 School Districts in Ghana" (2019).
- 114 Kellaghan, Greaney, and Murray, Using the Results of a National Assessment of Educational Achievement, 2009; T. S. Powdye, "The Bhutanese Education Assessment Experience: Some Reflections," *Prospects* 35, no. 1 (2005): 45–57.
- 115 Eberhardt, Bringing Learning to Light, 2015; RTI International, USAID's Early Grade Reading Program (EGRP) in Nepal, 2020.



Este documento está bajo una Licencia Internacional de Creative Commons Attribution 4.0.

<https://creativecommons.org/licenses/by/4.0/>