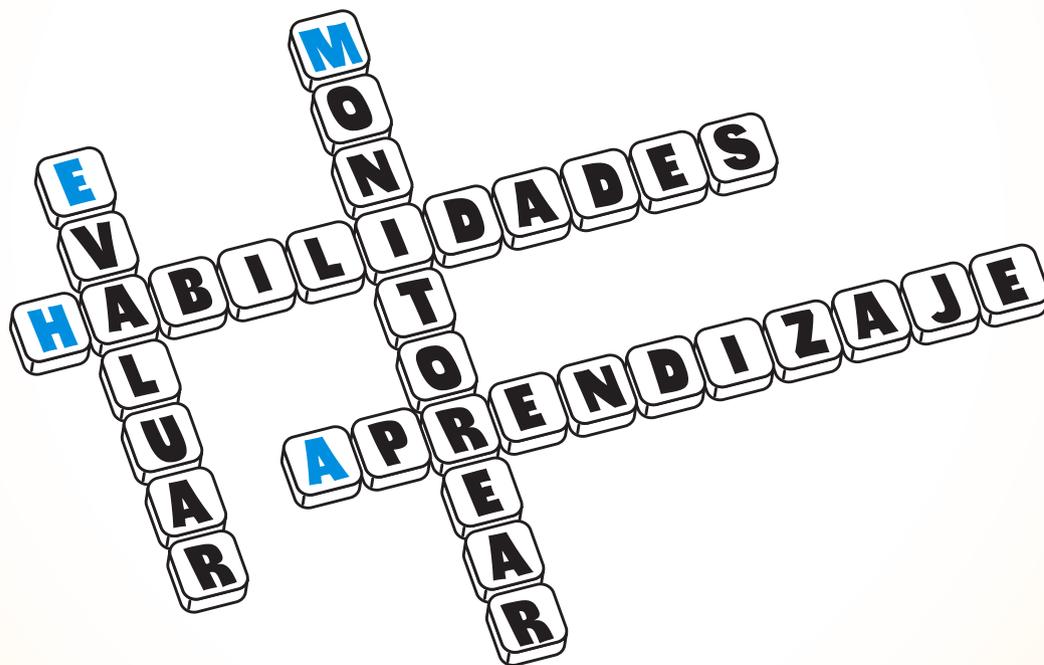


Guía Rápida No.3

Implementación de una Evaluación Nacional del Aprendizaje



Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura



INSTITUTO
DE ESTADÍSTICA
DE LA UNESCO



OBJETIVOS
DE DESARROLLO
SOSTENIBLE



ALIANZA
MUNDIAL para
LA EDUCACIÓN





Tabla de contenidos

| | |
|--|----|
| Introducción..... | 4 |
| 1. ¿Qué decisiones se deben tomar previo al lanzamiento de la evaluación?..... | 7 |
| 2. ¿Cómo se mide el aprendizaje? | 11 |
| 3. ¿Cómo se seleccionan las escuelas y los estudiantes? | 15 |
| 4. ¿Cómo se administra una evaluación? | 17 |
| 5. ¿Cómo se crea una base de datos? | 19 |
| 6. ¿Cómo se calculan los resultados? | 21 |
| 7. ¿Cómo se comunican los resultados? | 22 |
| 8. ¿Qué tipo de personal e instalaciones se requieren? | 24 |
| 9. ¿Cómo se elabora un presupuesto? | 26 |
| 10. ¿Cuáles son los principales problemas y desafíos?..... | 28 |
| Dónde encontrar más información. | 31 |

Publicada en 2019 por:

Instituto de Estadística de la UNESCO
P.O. Box 6128, Succursale Centre-Ville
Montreal, Quebec H3C 3J7 Canadá

Tel: +1 514 343-6880
Email: uis.publications@unesco.org
<http://uis.unesco.org>

Documento preparado por: Dra. María José Ramírez, Consultora, mariajose.ramirez001@gmail.com
ISBN 978-92-9189-238-9
Ref: UIS/2017/LO/TD/16

© UNESCO-UIS 2019

Introducción



Aprender es una condición esencial que permitirá a las personas alcanzar su pleno potencial y a los países desarrollarse social y económicamente. Por consiguiente, es importante que los países sepan cuánto están aprendiendo sus estudiantes.

Para averiguarlo, es indispensable implementar sólidas evaluaciones del aprendizaje. Los países administran evaluaciones nacionales con el objetivo de monitorear y apoyar el aprendizaje para todos. Esto se logra implementando pruebas de evaluación diseñadas para medir lo que los estudiantes saben y lo que pueden hacer con este conocimiento. Posteriormente, los resultados de estas evaluaciones se utilizarán como el fundamento de políticas y prácticas educativas.

El propósito de esta Guía Rápida es apoyar a los países en la implementación de sus evaluaciones nacionales del aprendizaje proporcionando respuestas a las siguientes preguntas.

- ¿Qué decisiones se deben tomar previo al lanzamiento de la evaluación?
- ¿Cómo se mide el aprendizaje?
- ¿Cómo se seleccionan las escuelas y los estudiantes?
- ¿Cómo se administra una evaluación?
- ¿Cómo se crea una base de datos?
- ¿Cómo se calculan los resultados?
- ¿Cómo se comunican los resultados?
- ¿Qué tipo de personal e instalaciones se requieren?
- ¿Cómo se elabora un presupuesto?
- ¿Cuáles son los principales problemas y desafíos?

Desarrollar una evaluación nacional del aprendizaje implica una larga trayectoria que requiere de compromisos políticos, financiamiento estable y capacidad técnica local. Los países que por primera vez implementan una evaluación nacional tienen mucho que aprender de las experiencias de otros países. El caso de Mozambique, un país de bajos ingresos que ha ido gradualmente implementando su evaluación nacional del aprendizaje, ha sido todo un éxito (véase el Recuadro 1).

Los países que logren implementar efectivas evaluaciones del aprendizaje se encontrarán en una posición más ventajosa en términos de apoyar el aprendizaje para todos

RECUADRO 1. EVALUACIÓN NACIONAL DEL APRENDIZAJE DE MOZAMBIQUE

Mozambique es un país de bajos ingresos ubicado en la costa sureste de África. Luego de una explosiva expansión de su cobertura escolar, el país volvió su atención al aprendizaje. Este nuevo enfoque promovió el desarrollo de una evaluación del aprendizaje orientada a monitorear las tendencias de aprendizaje identificadas en áreas clave del currículo nacional.

La primera evaluación nacional del aprendizaje fue administrada el año 2013. Una muestra representativa de estudiantes nacionales de tercer grado (6.764 estudiantes de 400 escuelas) fue sometida a una prueba de lectura en Portugués (el idioma de instrucción). La prueba mide la habilidad de los estudiantes para leer letras, sílabas, palabras, oraciones simples y párrafos breves, y su administración ha sido estandarizada. Los administradores externos leen en voz alta las instrucciones y la mayoría de las preguntas a todas las clases. Los maestros y directores de escuelas responden los cuestionarios de antecedentes.

El año 2016, la evaluación fue nuevamente administrada, esta vez en lectura y matemáticas. Se ha programado una nueva evaluación para el año 2019.

Los resultados de la evaluación muestran el porcentaje de estudiantes por nivel de competencia, donde el nivel más alto (Nivel III) corresponde a las expectativas curriculares. También muestran las diferencias de puntuaciones entre los años 2013 y 2016. Los resultados se publican en un informe nacional.

El INDE, un instituto semiautónomo de investigación dentro del Ministerio de Educación tiene a su cargo la evaluación nacional. El instituto es responsable del diseño y administración de las pruebas, procesamiento de datos, análisis y elaboración de informes.

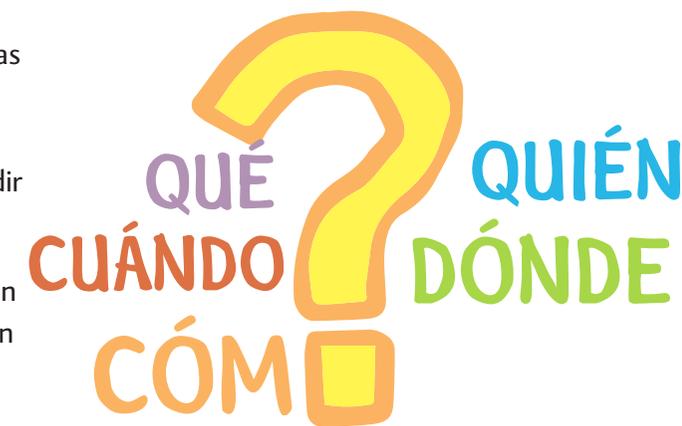
El INDE cuenta con un pequeño equipo de trabajo que gradualmente ha ido fortaleciendo su capacidad tecnológica en materia de evaluaciones. Previamente, el equipo había participado en la implementación de la evaluación internacional SACMEQ y un proyecto formativo de evaluación del aula (Quizz I Can Read) que sentó las bases para el desarrollo de la evaluación nacional. La experiencia del equipo también se enriqueció a través de su participación en un programa Magíster haciendo visitas de estudio y recibiendo capacitación práctica (con el apoyo del Fondo fiduciario READ).

Los costos de la evaluación nacional han sido cubiertos tanto por el gobierno como por fondos mancomunados.



¿Qué decisiones se deben tomar previo al lanzamiento de la evaluación?

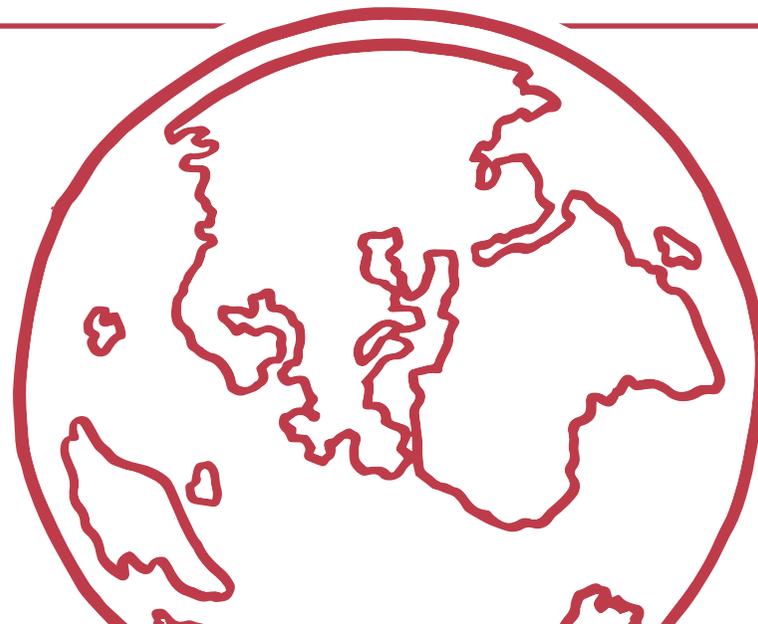
- **¿CUÁL es el propósito de la evaluación?** Los países desarrollan evaluaciones del aprendizaje con el fin de monitorear la medida en que los estudiantes están alcanzando los objetivos fundamentales como se describen en el currículo nacional y de esta forma promover el aprendizaje para todos. Algunos países también utilizan los resultados de las evaluaciones como reflejo del grado de responsabilidad que corresponde a las escuelas y estrategia para proporcionar a padres y estudiantes información sobre el avance del aprendizaje. Es esencial identificar las preguntas de política que la evaluación debe responder (por ej., ¿Qué porcentaje de los estudiantes ha alcanzado los objetivos curriculares de matemáticas al finalizar la educación primaria? ¿Se evidencia alguna mejora con el paso del tiempo?). La evaluación debe estar diseñada para responder estas preguntas.
- **¿QUÉ conocimientos cubrirá la prueba?** Las evaluaciones generalmente miden las áreas curriculares (dominios del aprendizaje) consideradas críticas para tener éxito en la vida, tales como comprensión lectora, matemáticas o ciencias. Otras miden áreas innovativas como escritura, salud o finanzas para así obtener un panorama integral del aprendizaje de los estudiantes. También es importante decidir en qué idioma(s) se administrará la prueba.
- **¿QUÉ otra información será recabada?** Los países deberán decidir si se utilizarán cuestionarios para recabar información sobre variables contextuales, elementos importantes para entender el aprendizaje (por ej., prácticas pedagógicas y



la disponibilidad de libros). Muchos países administran cuestionarios a los estudiantes, a sus maestros, a los directores escolares y a los padres.

- **¿QUIÉNES serán evaluados?** Los países generalmente seleccionan uno o más grados (por ej., el último grado de educación primaria) como objetivo de la evaluación. Dependiendo del uso que el país quiera darles a los resultados de la evaluación, podría administrar las pruebas a una muestra representativa de estudiantes nacionales o bien a la totalidad de las escuelas y alumnos (censo). En ambos casos, los países deben generar una lista completa de todas las escuelas participantes de manera que puedan tomar las muestras adecuadas o bien evaluar a la totalidad de la población estudiantil. En países con altos porcentajes de niños y jóvenes no escolarizados, sería aconsejable incluir a estos grupos en la prueba a través de encuestas de hogares.
- **¿CÓMO se evaluará a los estudiantes?** En la actualidad, la mayoría de las evaluaciones del aprendizaje se administran en el formato “lápiz y papel” y utilizan una combinación de preguntas de opción múltiple y respuestas abiertas. Algunos países administran pruebas similares en línea. La mayoría de los países implementan administraciones grupales (es decir, de la totalidad de la clase), mientras que otros administran las pruebas a cada estudiante en forma individual. Los administradores de las pruebas deben seguir pautas estrictas que garanticen que todos los estudiantes serán evaluados bajo las mismas condiciones, procedimiento conocido como administración “estandarizada”.

Cada país deberá responder estas preguntas con acuerdo a sus metas más amplias de educación, al contexto local y a los recursos disponibles.



- **¿QUIÉN administrará la evaluación?** Una opción sería capacitar a administradores de pruebas externos que aseguren que todos los estudiantes rendirán la prueba bajo las mismas condiciones. Los maestros también pueden colaborar en este sentido siempre que observen las pautas de administración.
- **¿CUÁNDO será administrada la evaluación?** Algunos países administran la evaluación todos los años, otros cada dos o más años. En algunos países se alterna el grado evaluado (por ej., el tercer grado se evalúa un año y el sexto grado el año siguiente). La frecuencia de la evaluación variará dependiendo del propósito perseguido y del financiamiento. Algunos países administran la evaluación al comienzo del año escolar y otros al final. La cronología de la evaluación debe tener en cuenta eventos de trascendencia a nivel escolar y nacional. La evaluación debe ser administrada cada vez en la misma oportunidad de manera que los resultados puedan ser comparables en el tiempo.
- **¿CÓMO se publicarán los resultados?** Idealmente, los resultados deberían publicarse tan pronto como sea posible y responder preguntas clave de política (por ej., ¿Obtienen los niños y las niñas resultados similares en lectura?). Estos resultados pueden ser publicados como el porcentaje de estudiantes que alcanzan distintos niveles de competencia (avanzado, básico, inferior a básico), puntuaciones medias, porcentaje de respuestas correctas u otros. La frecuente publicación y análisis de los resultados ayudan a informar a las distintas audiencias y partes interesadas sobre cómo está el país abordando los temas de interés educativo. Esto puede promover la confianza en la evaluación nacional.

Una evaluación nacional del aprendizaje puede costar entre 200.000 y 1.000.000 de dólares, dependiendo de una serie de factores.

- **¿CUÁNTO costará?** Una evaluación nacional del aprendizaje puede costar entre 200.000 y 1.000.000 de dólares, dependiendo de una serie de factores, por ej., la población objetivo (niños y jóvenes escolarizados comparados con niños y jóvenes no escolarizados), costo local de los servicios (impresión) y personal (administradores de pruebas). Es importante estimar los costos totales y asegurar un financiamiento apropiado y estable (proveniente del gobierno y donantes).
- **¿QUIÉN implementará la evaluación?** La entidad podría ser el Ministerio o Departamento de Educación, una agencia gubernamental semiautónoma o una universidad que cuente con un equipo especialista en evaluaciones. Es importante especificar el personal y las instalaciones de la entidad participante que estarán disponibles para la evaluación. Esta entidad debería responsabilizarse ante un órgano claramente definido (por ej., un departamento del gobierno).
- **¿QUIÉN estará encargado de tomar estas decisiones clave?** Se recomienda que un comité directivo tome las decisiones críticas respecto a qué debe ser evaluado, por quién y con qué frecuencia. El comité podría ser patrocinado por el servicio nacional de educación y estar integrado por representantes de partes interesadas clave (formuladores de política, educadores, padres y especialistas en evaluaciones). La participación de partes interesadas externas al servicio nacional de educación podría ayudar a concretar un compromiso nacional gracias a su mayor comprensión de los temas educativos y del proceso de evaluación. El Comité de Dirección debería ser asesorado por expertos técnicos.

Cada país deberá responder estas preguntas con acuerdo a sus metas más amplias de educación, al contexto local y a los recursos disponibles. Las decisiones tomadas como resultado de estas preguntas deberían ser documentadas en un marco de evaluación que debe estar disponible al público general. Dicho marco especifica las características principales de la evaluación nacional del aprendizaje, orienta el diseño e implementación de la evaluación, cumple la función de informar a las partes interesadas (formuladores de políticas, educadores, padres) y describe el enfoque que el país debe adoptar en su evaluación.

¿Cómo se mide el aprendizaje?

- **Identificar componentes, competencias, objetivos, contenidos o habilidades clave.** Por ejemplo, la alfabetización lectora podría desglosarse en componentes de decodificación y comprensión lectora y las matemáticas en áreas de contenido numérico, geometría, medición y álgebra. Este paso tiene implicaciones para el formato de la evaluación. Por ejemplo, una evaluación de lápiz y papel permite medir la comprensión lectora, no así la comprensión oral. En la medida de lo posible, el formato de la evaluación debe reflejar el nivel real de conocimientos y habilidades que los alumnos demuestran en el aula.
- **Identificar niveles clave de competencias.** Cada nivel (avanzado, básico, inferior a básico) describe lo que los estudiantes saben y son capaces de hacer en las distintas etapas de su trayectoria escolar. Por ejemplo, el nivel más alto de competencia puede describir los objetivos curriculares, en tanto que el nivel más bajo solo puede describir lo que los estudiantes de pobre rendimiento pueden efectivamente hacer. Los resultados de las evaluaciones se pueden reportar en base al porcentaje de estudiantes que alcanza cada nivel.



- **Redactar especificaciones para cada dominio del aprendizaje incluido en la prueba.** Identificar categorías analíticas clave para orientar el diseño de la prueba, es decir, componentes, competencias, objetivos, contenidos o habilidades; formato de los ítems (preguntas de opción múltiple o respuesta abierta); tipo de texto (narrativo, informativo); y grado de complejidad o dificultad de los ítems (fácil, intermedia, difícil). Especificar el número de ítems o preguntas que se necesitarán en total y en cada categoría. La distribución de ítems debe ser consistente con el currículo.
- **Contratar y capacitar a redactores de ítems.** Los redactores de ítems deben tener una clara comprensión del marco de evaluación y de las especificaciones de la prueba antes de comenzar su tarea. Estas personas pueden ser integrantes de una organización nacional de evaluación o externas a ella (por ej., maestros o especialistas curriculares de otra unidad de Ministerio o Departamento de Educación).
- **Redactar los ítems de la prueba.** Los ítems (preguntas o tareas) deben cubrir un amplio rango de niveles de dificultad de manera que puedan medir los diferentes niveles del aprendizaje. Un ítem de buena calidad: (a) Presenta tareas que pueden ser trazadas a las especificaciones de la prueba; (b) Es conceptualmente correcto; (c) Se centra en temas significativos, no en detalles triviales y; (d) Es claro respecto a lo que los estudiantes deben hacer. Los ítems de opción múltiple de buena calidad también (a) Incluyen una clave que es indudablemente correcta; (b) Incluyen distractores que son indisputablemente incorrectos

aunque parezcan razonables y plausibles; y (c) Presentan opciones de respuesta equivalentes en términos de estilo y extensión. Los ítems de respuesta abierta deben ser redactados junto con sus pautas de puntuación y recibir puntuación inmediatamente después de la administración de la prueba.



- **Redactar las instrucciones de la prueba.** Las instrucciones de la prueba deben dar la bienvenida a los estudiantes, explicar cómo realizar la prueba y en cuanto tiempo y proporcionar algunos ítems de práctica. Los estudiantes deben entender cómo responder las preguntas de la prueba, incluyendo las de opciones múltiples. Si se proporciona una hoja de respuesta separada los estudiantes deben practicar sus respuestas en ella. El administrador de la prueba debe leer las instrucciones antes que los estudiantes comiencen a responder.
- **Actualizar y realizar una prueba con los ítems antes de la prueba piloto.** Los ítems deben ser actualizados por los redactores y paneles externos integrados por personas no involucradas en la redacción. Los ítems actualizados deben ser primero administrados a un grupo reducido de estudiantes (por ej., 50 estudiantes). Antes de ensamblar los cuadernillos que se utilizarán en la prueba piloto se deberían realizar rondas de actualizaciones y revisiones.
- **Realizar una administración piloto.** Esto es importante ya que permite conocer las características psicométricas (dificultad o discriminación) de la prueba en general y de cada ítem en particular. Se considera buena práctica incluir en la prueba piloto dos o tres veces el número de ítems requeridos para la prueba final y administrar cada ítem a un mínimo de 200 estudiantes con características similares a los que rendirán la prueba final. La prueba piloto también es importante como forma de comprobar la eficiencia del manual de administración, la logística (por ej., el transporte) y otros procedimientos relacionados con esta actividad.
- **Analizar los resultados de la prueba piloto.** Calcular las estadísticas de los ítems tales como el nivel de dificultad (porcentaje de respuestas correctas, puntuación escalonada) y de discriminación (punto de correlación biserial); porcentaje de estudiantes por opción de respuesta (para ítems de opción múltiple) o porcentaje que ha logrado crédito total o parcial (preguntas de respuestas abiertas); y porcentaje que omitió la respuesta. Las principales estadísticas de la prueba incluyen la puntuación media, fiabilidad y precisión (por ej., el Alfa de Cronbach). Se considera buena práctica que el ítem cubra un amplio rango de dificultades y que el Alfa de Cronbach sea superior a 0.70. El programa computacional IATA es gratis y se puede utilizar para realizar estos análisis.

- **Ensamblaje de cuadernillos de prueba para la administración principal.** Los cuadernillos que se utilicen en la prueba deben ser atractivos sin ser distractivos. Deben incluir números identificadores (ID) que permitan localizar a estudiantes y escuelas. El número de ítems en el cuadernillo debe permitir que el estudiante finalice la prueba dentro de un plazo razonable. Es ventajoso presentar los ítems más fáciles al inicio de los cuadernillos. Los países suelen utilizar distintos cuadernillos con diferentes ítems para incrementar el número total de ítems incluidos en la prueba. En este caso, los distintos cuadernillos deben ser equivalentes en términos de contenido y características psicométricas (dificultad), si bien también pueden compartir ítems comunes.
- **Preparación para la etapa de impresión.** Imprimir es bastante costoso y puede demorar varias semanas. Las cotizaciones se deberán solicitar tan pronto se conozcan los parámetros clave (número de cuadernillos y número de páginas en cada uno). Todos los cuadernillos de prueba y otros instrumentos (cuestionarios y manuales de administración) deben ser rigurosamente revisados antes de imprimir.



¿Cómo se seleccionan las escuelas y los estudiantes?

Las escuelas y los estudiantes se deben seleccionar de manera que los resultados de la evaluación sean aplicables a todo el país. En ciertos casos, la selección de escuelas y estudiantes podría verse limitada debido a restricciones logísticas o financieras. En estos casos, los resultados de la evaluación deben ser interpretados como pertenecientes a un nivel subnacional del país (por ej., el porcentaje de estudiantes que alcanzan el nivel mínimo de competencia en escuelas urbanas de la región X). Toda decisión relativa a la selección de escuelas y estudiantes debe ser documentada.

Los siguientes pasos son esenciales al momento de seleccionar escuelas y estudiantes para participar en una evaluación del aprendizaje.

- **Especificar la población objetivo de estudiantes.** Indicar a qué estudiantes del país aplicarían los resultados (por ej., a todos los estudiantes de sexto grado de escuelas públicas y privadas).
- **Crear una lista completa de escuelas.** Idealmente, esta lista (marco del muestreo) debería incluir información de contacto e identificación de todas las escuelas, información relativa a la clasificación de la escuela (urbana/rural, pública/privada, regional/distrital) y el número de clases y de estudiantes que integran el grado objetivo.
- **Definir la mejor estrategia para seleccionar**



escuelas y estudiantes:

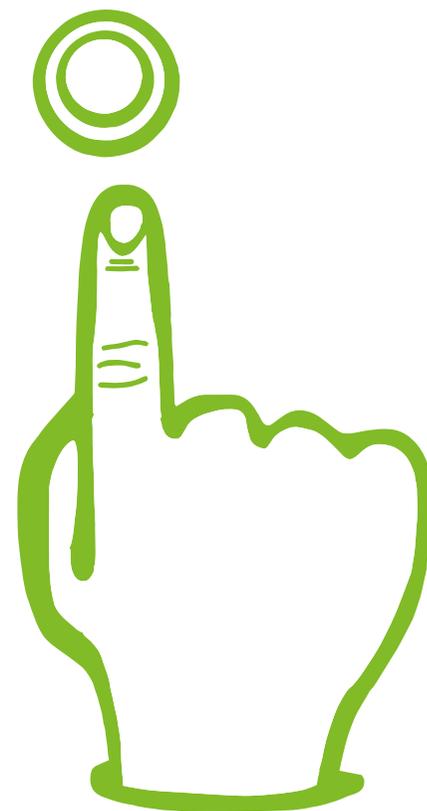
- a. Seleccionar una muestra representativa de estudiantes a nivel nacional, o subnacional. Esta representa la mejor estrategia si el propósito de la evaluación es estimar el nivel general de aprendizaje de un país. Los métodos de muestreo normalmente involucran: i) estratificación de la muestra de acuerdo con la información de clasificación de la escuela (ej., por región); (ii) muestreo aleatorio sistemático de las escuelas incluidas en el marco de muestreo, donde las escuelas más grandes tiene una mayor probabilidad de ser seleccionadas; y (iii) muestreo de estudiantes o de una clase completa en las escuelas. Es aconsejable seleccionar escuelas de reemplazo ante la eventualidad que algunas de las escuelas originalmente seleccionadas no puedan participar.
- b. Seleccionar la totalidad de las escuelas y estudiantes (censo) del grado objetivo. Esta generalmente representa la mejor estrategia si el propósito de la evaluación es informar resultados a nivel de escuelas o si la evaluación se implementa en un país relativamente pequeño.



¿Cómo administrar la evaluación?

La administración de la evaluación del aprendizaje es una iniciativa de trascendental importancia. Habitualmente, requiere contactarse con cientos de escuelas, reclutar y capacitar a cientos de administradores y supervisores de pruebas, distribuir miles de cuadernillos en forma segura a todo el país y subcontratar diversos servicios (por ej., transporte). Las operaciones en terreno son todas las actividades consideradas necesarias para una exitosa administración de la prueba. Estas incluyen:

- **Diseño de un plan operativo.** Este debería identificar actividades, responsabilidades y plazos y establecer cómo se asignarán las escuelas a los coordinadores en terreno y administradores de la prueba (por ej., cinco escuelas por coordinador). Un manual de operaciones en terreno y su debida documentación (formularios y listas de verificación) deberían proporcionar las pautas requeridas y funcionar como mecanismos de control de calidad.
- **Capacitación del personal clave.** La capacitación debería enfatizar los objetivos que persigue la evaluación, la importancia de contar con condiciones uniformes durante la prueba, y definir los roles y responsabilidades de los nuevos reclutas durante las operaciones en terreno. La capacitación debería realizarse antes que la recolección de datos y, dentro de lo posible, replicar fielmente las condiciones reales de la administración

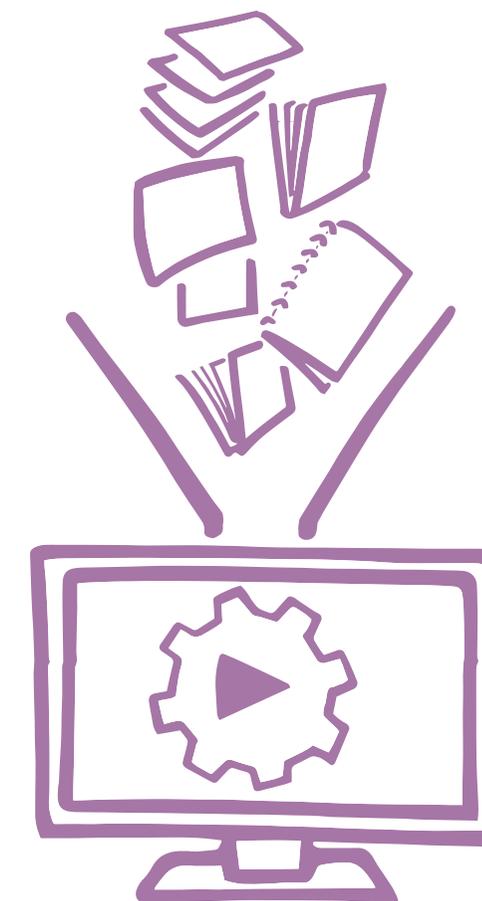


de la prueba. Esto se puede lograr, por ejemplo, leyendo en voz alta el manual de administración durante la sesión de capacitación y completando los formularios pertinentes.

- **Establecer contacto con las escuelas.** Invitar a un grupo seleccionado de escuelas a participar en la evaluación. Garantizar que es posible establecer contacto con un integrante de la escuela que se encuentre habilitado para organizar la administración de la prueba (por ej., informando a los maestros acerca de la evaluación, almacenando los cuadernillos de prueba, etc.).
- **Imprimir, verificar y distribuir los instrumentos de evaluación.** Revisar las muestras de los instrumentos impresos (pruebas y cuestionarios) para comprobar que los estándares de calidad se han cumplido. El empaque y distribución pueden tomar varias semanas. A la seguridad y confidencialidad de los instrumentos siempre se les debe otorgar la más alta prioridad.
- **Administrar la evaluación con acuerdo a procedimientos estandarizados.** Todos los estudiantes deberán rendir la prueba bajo las mismas condiciones. El manual de administración de pruebas debe establecer dichas condiciones y servir como fundamento del control de calidad.
- **Enviar observadores del control de calidad a las escuelas.** Es importante verificar que la evaluación ha sido administrada bajo las mismas condiciones en todas las escuelas. Asimismo, es importante inspirar confianza en la evaluación. Los observadores pueden utilizar el manual de administración o una lista de verificación como referencia. Entre los observadores se pueden incluir funcionarios del Ministerio o Departamento de Educación, donantes, padres y otros.
- **Comprobar el retorno de los instrumentos de evaluación.** Generalmente, los países deben recoger todos los cuadernillos y demás instrumentos utilizados para la prueba (cuestionarios y formularios) en las distintas escuelas. La totalidad de los cuadernillos e instrumentos deben ser registrados (con números de identificación) y contabilizados antes y después de la prueba.

¿Cómo se crea una base de datos?

- **Preparar una estructura idónea para la captura de datos.** La base de datos debe incluir todas las variables que aparecen en los cuadernillos de prueba y en otros instrumentos (por ejemplo, números de identificación, respuestas al ítem 1, ítem 2, etc.). También debe especificar el tipo de información que debe registrarse en cada variable (de texto, numérica y rangos válidos). Esta información debe documentarse en un manual de base de datos (libro de códigos).
- **Preparar protocolos para el procesamiento de datos.** Los protocolos tienen la función de establecer reglas para la captura, limpieza y verificación de los datos. Por ejemplo, instruye a los dactilógrafos sobre cómo ingresar los datos, establece reglas para realizar verificaciones de consistencia interna (por ej., si un estudiante se encuentra ausente todos los ítems de respuesta deberían aparecer como faltantes), y correcciones automatizadas.
- **Captura de datos en bruto en la base de datos.** Entre los métodos utilizados se incluye la recolección de datos en línea, el escaneo de cuadernillos de prueba u hojas de calificaciones y el teclado de información. Según el método seleccionado será

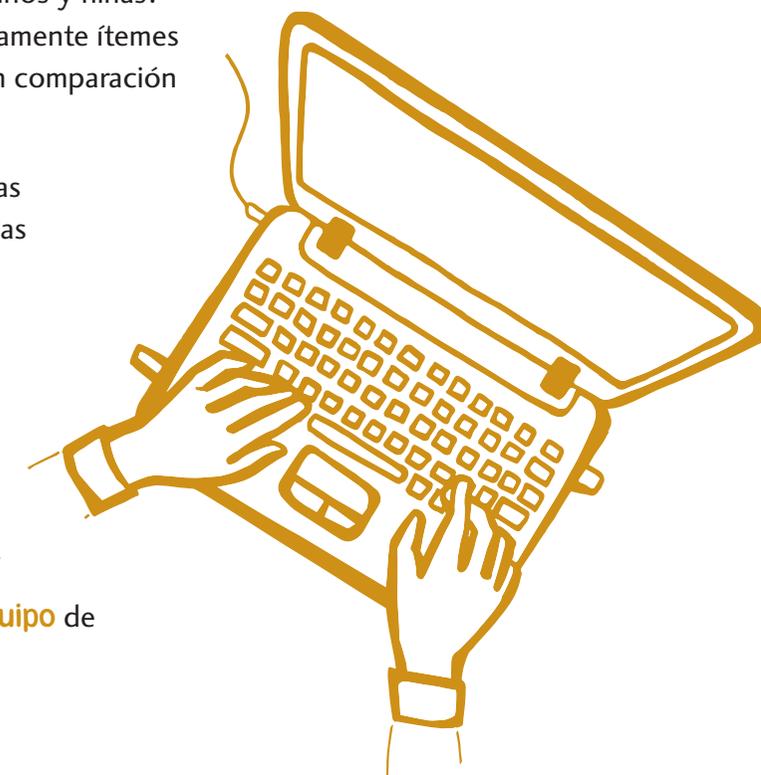


necesario utilizar hardware (por ej., escaneadores de alta velocidad) y software (Access o Excel) especializados. Se debería destinar los fondos necesarios para adquirir los equipos y programas requeridos para esta actividad.

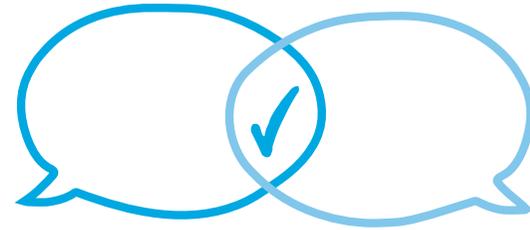
- **Limpieza de datos.** Incluye procedimientos de verificación y corrección de códigos únicos de identificación, registros duplicados o eliminados y comprobación que los valores ingresados sean válidos y estén dentro del rango. La limpieza de datos habitualmente requiere varias rondas de iteraciones antes de producir la base de datos final u oficial.
- **Calcular o agregar nuevas variables.** Estas pueden incluir las siguientes:
 - a. Puntuaciones IRT (Ítem de Respuesta al Ítem) escalonadas. Estas permiten situar las puntuaciones de diferentes cuadernillos en la misma escala para efectos de comparación.
 - b. Niveles de competencia. Los estudiantes se clasifican en un nivel (avanzado, básico, inferior a básico) de acuerdo con su puntuación final en la prueba.
 - c. Pesos para posibles ajustes al diseño de la muestra. Estos elementos son importantes para producir resultados válidos a nivel de países.

¿Cómo se calculan los resultados?

- **Identificar preguntas clave de política que orienten los análisis.** Por ejemplo: ¿Cuál es la puntuación media del país en matemáticas? ¿Qué porcentaje de los estudiantes alcanzó el nivel mínimo de competencia en lectura? ¿Existen diferencias de puntuación entre niños y niñas? ¿Qué porcentaje de los estudiantes respondió correctamente ítemes específicos de la prueba? ¿Mejoraron los resultados en comparación con la evaluación anterior?
- **Analizar los datos.** Calcular las estadísticas descriptivas (puntuaciones medias y porcentaje de estudiantes) y las estadísticas inferenciales básicas (comparaciones de medias) para responder las preguntas de investigación más elementales. Aceptar que las puntuaciones pueden no ser absolutamente precisas (fidedignas) reportando errores estándar o señalando si las diferencias son estadísticamente significativas.
- **Aplicar pesos** de acuerdo con el diseño de la muestra.
- **Documentar todos los análisis realizados por su equipo** de manera que otros puedan entenderlos y replicarlos.



¿Cómo se comunican los resultados?



Los países implementan evaluaciones del aprendizaje con el propósito de monitorear y apoyar el aprendizaje para todos. Para lograr estos objetivos, los resultados de las evaluaciones deben ser comunicados rápida y eficientemente a las partes interesadas. Entre las acciones esenciales que se deben tomar para garantizar una buena comunicación se cuentan las siguientes:

- **Elaborar un plan de Comunicaciones.** Este plan requiere la identificación de: mensajes y preguntas clave de política que se desean comunicar; los productos que deben ser comunicados (informes o folletos nacionales); las partes interesadas hacia quienes los productos van dirigidos (padres, maestros); y la periodicidad de estas publicaciones (por ej., 3 meses después de la administración de la evaluación). Es importante obtener el financiamiento que permita la implementación de este plan.
- **Identificar mensaje y preguntas clave de política.** Por ejemplo, ¿cuántos estudiantes cumplen los objetivos curriculares actualmente? ¿Cómo se comparan los resultados regionales con los resultados nacionales? ¿Se aprecia algún avance en el tiempo?
- **Identificar audiencias o partes interesadas esenciales.** Por ejemplo, padres, maestros, el Ministerio o Departamento de Educación, agencias de financiamiento, los medios de comunicación y el público general. Priorice la atención a sus necesidades de información. Desarrolle una cultura de la evaluación que promueva entre las partes interesadas la comprensión, valorización y utilización de la información derivada de las evaluaciones para mejorar el aprendizaje.
- **Seleccionar los mejores productos de comunicación.** Los distintos productos (reportes, folletos o talleres) persiguen diversos objetivos y son útiles para llegar a diversas audiencias o partes interesadas. Asegúrese que los productos de comunicación seleccionados respondan las preguntas clave de política

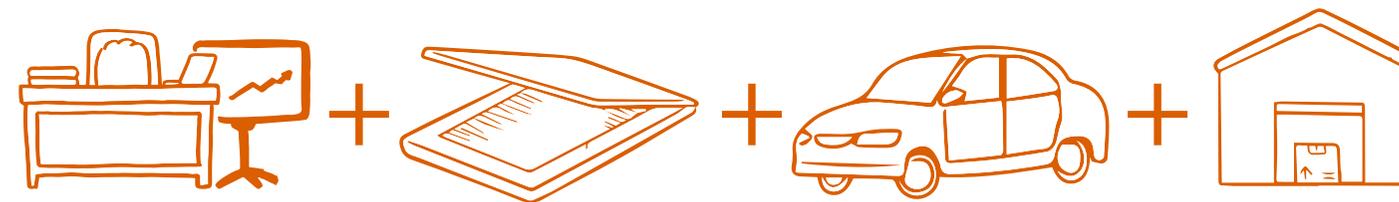
(por ej., actualmente, ¿qué porcentaje de los estudiantes alcanzan el nivel mínimo de competencia en lectura?) y se hagan llegar a las distintas partes interesadas (por ej., a través de la distribución de informes escolares a todas las escuelas). Asegúrese que la información haya sido reportada en forma clara y correcta. Ciertos informes escolares pueden requerir bases de datos adicionales (por ej., una base de datos que permita imprimir puntuaciones específicas para cada escuela). Algunos productos de comunicación que debería tener presente son los siguientes:

- Informes nacionales para el público general. Estos proporcionan una visión panorámica de la evaluación y de sus principales resultados.
 - Talleres para docentes y líderes escolares. Estas iniciativas ofrecen una oportunidad para explicar la importancia de la evaluación y cómo usar los resultados para mejorar el aprendizaje de los estudiantes.
 - Informes para los medios de comunicación. Estos contienen toda la información necesaria para comunicar los resultados a una audiencia más extensa, e incluye mensajes clave, gráficos y tablas.
 - Sitios web, videos, blogs y los medios sociales para motivar el continuo interés del público general.
 - Los informes para padres y escolares representan herramientas efectivas para propósitos de asignación de responsabilidades.
 - Las conferencias y presentaciones (por ej., en Power Point) ofrecen una oportunidad para involucrar al mundo académico y a otras instituciones (por ej., agencias de financiamiento).
 - Los folletos, panfletos y carteles se prestan para comunicar mensajes breves a una audiencia más amplia y a quienes no tienen tiempo de leer resultados más elaborados (por ej., los formuladores de políticas).
- **Presentar acciones para mejorar los resultados de la evaluación.** Esto es particularmente importante cuando los resultados informados son de baja calidad. Por ejemplo, el Ministerio de Educación puede explicar que los resultados representan un diagnóstico o punto de partida e iniciar un nuevo programa en apoyo del aprendizaje. Por su parte, las escuelas pueden precisar los objetivos de aprendizaje que se deberán haber logrado de aquí a la próxima publicación de resultados.
 - **Monitorear cómo están siendo utilizados los resultados de las evaluaciones.** Verifique si los informes están efectivamente llegando a las escuelas, padres o partes interesadas como está previsto. Reúname con las partes interesadas y pregúnteles acerca de los distintos productos de evaluación: ¿Han sido recibidos como planificado? ¿Fueron los resultados informativos para su escuela en particular? ¿Se pudieron entender claramente? ¿Cómo usaron los resultados en apoyo del aprendizaje?

¿Qué personal e instalaciones se requieren?

El personal esencial para implementar una evaluación nacional del aprendizaje está integrado por:

- **Un coordinador nacional**, responsable de entregar orientaciones generales y asumir funciones de liderazgo, garantizando que la evaluación aborde preguntas clave de política, sea técnicamente adecuada, y se realice dentro de los plazos acordados y de acuerdo con el presupuesto disponible. En algunos países se podrían requerir coordinadores subnacionales. El personal administrativo y los especialistas en informática también pueden trabajar directamente con el coordinador nacional.
- **Coordinadores de desarrollo de pruebas**. Estas personas estarán a cargo de la producción de pruebas diseñadas para medir el aprendizaje en cada uno de sus dominios (por ej., matemáticas, lectura) y deben estar bien informadas sobre los objetivos curriculares y los niveles reales de aprendizaje de los estudiantes en el aula. Entre sus principales responsabilidades se incluyen la contratación, capacitación y apoyo de los equipos de redacción de ítems y la puntuación de preguntas de respuesta abierta.
- **Coordinador de operaciones en terreno**. La persona responsable de la administración estandarizada de todos los instrumentos de evaluación estará a cargo de la contratación y capacitación de los administradores de las pruebas, supervisores, coordinadores a nivel regional o distrital y observadores del control de calidad. Esta persona podría ser un empleado a tiempo completo de la organización nacional para evaluación de la educación, integrante de otras ramas del Ministerio o Departamento de Educación, o bien un profesional temporalmente contratado. El coordinador de operaciones en terreno también será responsable de la



producción del manual de administración de la evaluación y de sus aspectos logísticos (por ej., distribución y recolección de instrumentos).

- **Coordinador del administrador de datos**. Persona responsable de la preparación de bases de datos, es decir, captura y limpieza de datos, cálculo de pesos, puntuaciones y puntos de corte asociados con niveles de competencia y resultados. Esta persona debe garantizar que las bases de datos cumplan los requisitos de los productos de comunicación y demostrar competencia en la realización de análisis estadísticos y psicométricos utilizando diferentes programas informáticos (por ej., SAS, SPSS, STATA y IATA).
- **Coordinador de comunicaciones**. Responsable de elaborar e implementar un plan de comunicaciones efectivo. Es altamente recomendable trabajar con un diseñador y un especialista en comunicaciones.

Para operar en forma eficiente, la organización nacional para evaluación de la educación requerirá tener acceso a instalaciones básicas, tales como:

- **Oficinas y salas de reuniones** para personal clave dotadas de conexión a internet, impresoras y otros artículos de uso normal en oficinas.
- **Soporte físico** (por ej., escaneadores) y **programas informáticos** (por ej., SPSS, IATA) con sus respectivas licencias de operación.
- **Vehículos para el transporte** de cajas con instrumentos de evaluación.
- **Instalaciones de almacenamiento seguras** con capacidad para un gran número de cajas con instrumentos de evaluación.

¿Cómo se elabora un presupuesto?

Elabore un presupuesto identificando los costos de los principales ítems de la evaluación y estimando los costos de cada uno. La Tabla 1 muestra los principales ítems de costo y el desglose del costo proporcional por ítem. Los costos totales y el desglose pueden variar significativamente por país, dependiendo de las características de la evaluación (por ej., censo frente a muestreo) la capacidad institucional existente y los costos locales, entre otras cosas. Nótese que la tabla no incluye los costos asociados con el personal, las oficinas o el equipamiento. Usted podrá ver cómo otros países elaboraron presupuestos para sus evaluaciones del aprendizaje [aquí](#).

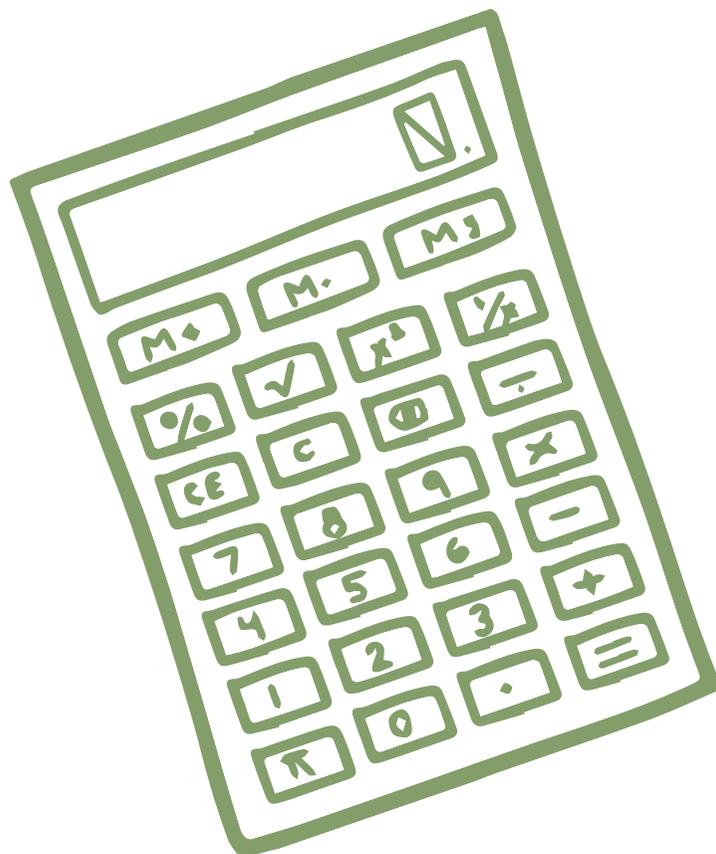
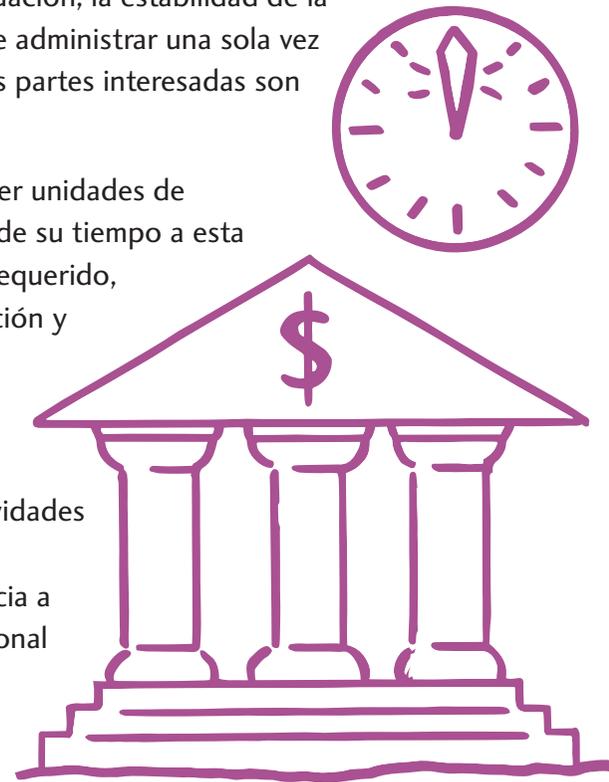


TABLA 1. Costos de los principales ítems y estimación desglosada de los costos asociados con una evaluación del aprendizaje

| ÍTEM | DESGLOSE DE COSTOS |
|---|--------------------|
| DISEÑO Y PRUEBA PILOTO DE LA EVALUACIÓN | |
| Marco de la evaluación, especificaciones de la prueba | 20% |
| Redacción de ítems | |
| Desarrollo de cuestionarios | |
| Muestreo | |
| Impresión de materiales para la prueba piloto | |
| Administración de la prueba piloto | |
| ADMINISTRACIÓN PRUEBA PRINCIPAL | |
| Muestreo | 50% |
| Diseño de cuadernillos y cuestionarios de la prueba | |
| Impresión de instrumentos (cuadernillos, cuestionarios, manuales) | |
| Distribución a las escuelas | |
| Administración de la prueba principal | |
| PROCESAMIENTO Y ANÁLISIS DE DATOS | |
| Captura y limpieza de datos | 15% |
| Puntuación de preguntas de respuesta abierta | |
| COMUNICACIÓN | |
| Diseño de sitios web, informes, folletos, carteles | 15% |
| Talleres, conferencias | |
| Distribución a escuelas, padres y otras partes interesadas | |
| TOTAL | 100% |

¿Cuáles son los principales problemas y desafíos?

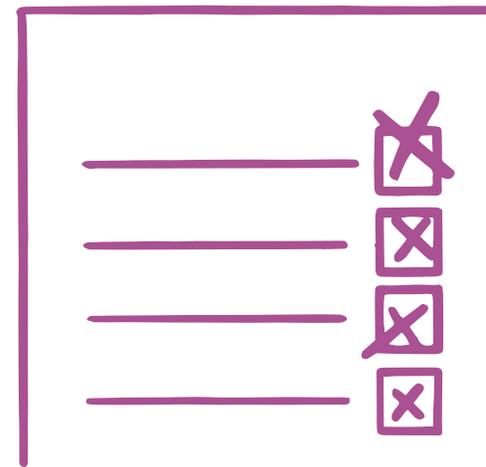
- **Incapacidad de obtener apoyo político y financiamiento estable.** Es posible que los costos de la evaluación no estén totalmente cubiertos por el presupuesto anual del Ministerio o Departamento de Educación. Como consecuencia de esta situación, la estabilidad de la evaluación podría encontrarse en riesgo. La evaluación se puede administrar una sola vez o en forma irregular. La transparencia y el involucramiento de las partes interesadas son esenciales para minimizar este riesgo.
- **Necesidad de contar con el personal necesario.** Es común ver unidades de evaluación cuyos integrantes pueden dedicar solo una fracción de su tiempo a esta actividad. Es importante llegar a un acuerdo sobre el personal requerido, la cantidad de tiempo que cada uno puede destinar a la evaluación y planear su implementación en base a este compromiso.
- **Necesidad de desarrollar la capacidad local.** La mejor estrategia para lograr esto es proporcionar capacitación práctica durante la implementación de la evaluación. Otras actividades de reforzamiento de la capacidad pueden involucrar cursos o talleres (presenciales o a distancia), visitas de estudio y asistencia a conferencias. Se debe hacer todo lo posible por retener al personal capacitado.
- **Las pruebas son demasiado difíciles.** Las pruebas de evaluación suelen ser precisas cuando se trata de medir



objetivos curriculares o expectativas, pero no así cuando miden lo que los estudiantes en la parte inferior de la distribución del aprendizaje pueden hacer. En consecuencia, los resultados pueden mostrar que la mayoría de los estudiantes se encuentran agrupados en los niveles más bajos de competencia. Por lo tanto, asegúrese que el diseño de la prueba incluya información tanto del currículo como de los niveles reales de aprendizaje de los estudiantes.

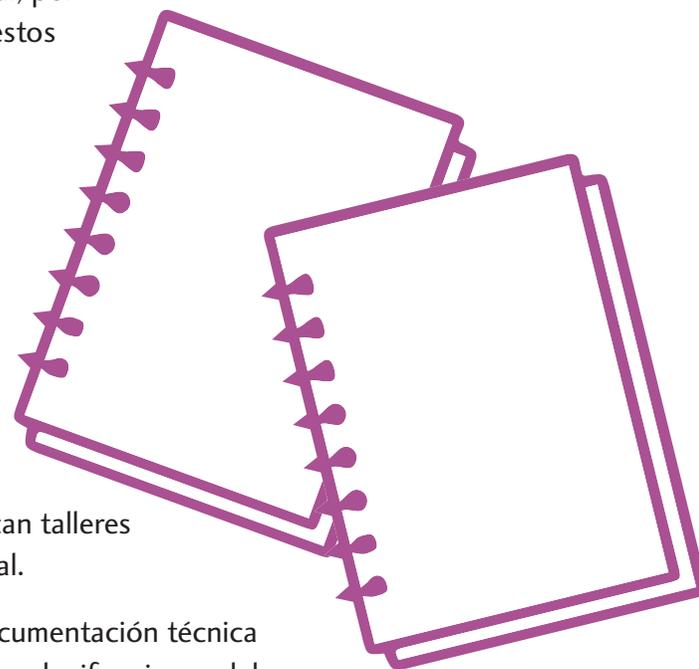


- **Muestreo deficiente.** Evaluaciones que supuestamente son representativas a nivel del país pueden terminar utilizando muestras representativas de solo algunas regiones o de ciertos tipos de establecimientos educativos (escuelas urbanas). O, durante el trabajo en terreno, la prueba podría administrarse solo a escuelas de “fácil acceso”. Esta situación se puede dar como resultado de una mala planificación, dificultades logísticas, o un presupuesto insuficiente. El muestreo y el trabajo en terreno deben ser planificados en detalle y con bastante antelación para evitar este tipo de problemas.
- **Falta de procedimientos estandarizados.** Cuando no se dispone de pautas y manuales claros para administrar la evaluación, cuando los administradores de las pruebas no han recibido la capacitación adecuada, cuando las reglas de puntuación no están claras, cuando no existen reglas o protocolos para crear bases de datos, la totalidad del esfuerzo realizado podría terminar siendo solo una pérdida de tiempo. Se requiere un nivel mínimo de estandarización para garantizar que las puntuaciones de la prueba se puedan interpretar como estaba previsto. Los manuales, la capacitación y los procedimientos de control de calidad constituyen importantes herramientas para garantizar dicha estandarización.
 - **Los resultados de las evaluaciones no son comparables.** Esto se convierte en un serio problema cuando el propósito es informar sobre los cambios que ha tenido el aprendizaje a través de



los años. La comparabilidad se ve en riesgo cuando las escuelas y los estudiantes muestreados no son equivalentes entre una y otra administración o las pruebas no son paralelas o no se rigen por la misma escala de puntuación. Otro problema relacionado con lo anterior es que la incertidumbre o error asociado con las puntuaciones de las pruebas es desconocido, hecho que hace imposible saber si las diferencias son significativas o estadísticamente relevantes. Al momento de diseñar la evaluación, debe existir la certeza que las características técnicas requeridas para comparar los resultados han sido incorporadas.

- **Los resultados de la evaluación no se han publicado.** Es muy probable que esta situación ocurra en un contexto de malos resultados. Es importante anticipar este escenario ante el Ministerio de Educación o ente equivalente y planificar, por ejemplo, una estrategia de comunicación que utilice estos resultados como punto de referencia desde donde se implementarán medidas de mejoramiento.
- **Falta de una cultura de la evaluación.** Es posible que las partes interesadas (maestros, padres, figuras políticas), no entiendan, valoren o utilicen los resultados de la evaluación como se había previsto. Esto podría suceder si los informes de los resultados se caracterizan por ser demasiado extensos y complejos. La solución está en producir informes simples, folletos y sitios web que aborden un número reducido de preguntas clave de investigación y ofrezcan talleres para explicar los resultados de la evaluación en general.
- **Falta de documentación.** Se requiere contar con documentación técnica (marco de la evaluación, especificaciones de la prueba, planificaciones del muestreo, protocolos de procesamiento de datos) como evidencia que la evaluación puede ser interpretada y utilizada como se anticipaba. Esto tiene una importancia clave para garantizar la credibilidad de la evaluación.



Dónde encontrar más información.

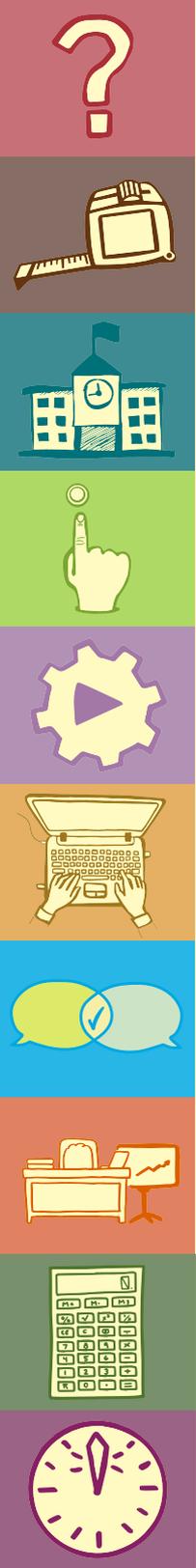
Esta Guía Rápida está basada en las siguientes referencias que cubren en mucho más detalle los temas presentados en este documento:

Banco Mundial (2008). *Serie: Evaluaciones Nacionales del Logro Educativo* (disponible en Inglés, Francés, Portugués, Español y Ruso). Washington, D.C.: Banco Mundial.

- Evaluación de los niveles nacionales del logro educativo (Vol. 1)
- Desarrollo de pruebas y cuestionarios para una evaluación nacional del logro educativo (Vol. 2)
- Implementación de una evaluación nacional del logro educativo (Vol. 3)
- Análisis de datos de una evaluación nacional del logro educativo (Vol. 4).
- Uso de los resultados derivados de una evaluación nacional del logro educativo (Vol 5)

UNESCO-UIS y ACER (2017). *Principles of Good Practice in Learning Assessment*. Montreal: UNESCO Institute for Statistics (UIS).

Wolff, Laurence (2007). *The Costs of Student Assessments in Latin America*. Washington, D.C.: PREAL.



Organización
de las Naciones Unidas
para la Educación,
la Ciencia y la Cultura



INSTITUTO
DE ESTADÍSTICA
DE LA UNESCO



ALIANZA
MUNDIAL *para*
LA EDUCACIÓN

@UNESCOstat
<http://uis.unesco.org>