



Universidad Autónoma de Baja California

Instituto de Investigación y Desarrollo Educativo

**Propuesta técnica del Estudio de Pruebas de Logro de Aprendizaje del
Cuarto Estudio Regional Comparativo y Explicativo (ERCE 2019). Primera
fase: Desarrollo de instrumentos para el piloto ERCE**

Elaborada por:

Dr. José Alfonso Jiménez Moreno, Dr. Juan Carlos Pérez Morán, Dr. Juan Carlos Rodríguez
Macías, Dr. Joaquín Caso Niebla y Mtro. Julio César Cano Gutiérrez

Ensenada, B. C. a 9 de marzo de 2017

I. Introducción

El Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) se considera ante la comunidad internacional como un referente en el ámbito de la evaluación educativa y marco regional de concertación y cooperación entre los países de Latino América y el Caribe para el logro de la calidad de los sistemas educativo. De forma especial, LLECE representa un apoyo técnico para la formación y capacitación de los equipos responsables de los sistemas de medición y evaluación educativa de la región, y como una fuente de acceso a la información para que los países alimenten su toma de decisiones y en especial la elaboración de políticas para la mejora educativa basada en evidencia empírica (UNESCO-OREALC, 2015).

Cabe recordar, que el LLECE nació en 1994, en Ciudad de México, como una red de unidades de evaluación de la calidad de los sistemas educativos de América Latina y el Caribe y que actualmente es coordinado por la Oficina Regional de Educación de la UNESCO para la América Latina y el Caribe (OREALC). También es importante recordar, que el LLECE ha desarrollado tres estudios sobre el logro educativo y sus factores asociados con la intención de promover el desarrollo de políticas educativas de los países participantes. Entre los aspectos relevantes de dichos estudios, es importante mencionar que el LLECE se basa en el concepto de calidad educativa propuesto por la UNESCO, el cual contempla desde una postura amplia más allá de la simple evaluación del logro de los aprendizajes incluyendo aspectos como la inclusión, la equidad, la relevancia y la pertinencia en educación.

Para dar cuenta de los propósitos, encomendados por la comunidad internacional y regional en cuanto a la evaluación de la calidad de los sistemas educativos de la región, el LLECE en 1997 el Laboratorio aplica el Primer Estudio Regional Comparativo y Explicativo (PERCE) a estudiantes de tercer y cuarto año de Educación Primaria en las áreas de Lectura y Matemática. Dicho esfuerzo, es reconocido por los actores educativos de distintos niveles como uno de los mayores logros políticos en educación en el continente, convocando exitosamente la participación de 13 países: Argentina, Bolivia, Brasil, Chile, Colombia, Costa Rica, Cuba, Honduras, México, Paraguay, Perú, República Dominicana y Venezuela. Los resultados del PERCE fueron publicados en noviembre de 1998.

En la actualidad, el LLECE continúa cosechando logros y capitalizando su presencia en el tema de la evaluación educativa, además de consolidar sus esfuerzos por convertirse en una fuente de información confiable, pertinente y oportuna para la toma de decisiones y generación

de política para la mejora, la equidad y la calidad educativa. Una evidencia de ello, es la continuidad y mejora que presenta el Estudio Regional de Calidad Educativa (ERCE), al cual los investigadores del Instituto de Investigación y Desarrollo Educativo (IIDE) consideran una distinción el ser invitados a participar como socios implementadores, que se espera aplicar en 2019 con la participación de diversas agencias de evaluación latinoamericanas, y que incluirá la participación de 19 sistemas educativos (OREALC-UNESCO, 2016b).

El IIDE, como parte de la Universidad Autónoma de Baja California (UABC), una de las Instituciones de Educación Superior más reconocidas de México, cuenta con amplia experiencia en procesos de investigación y evaluación educativa a nivel local, regional, nacional e internacional. En ese sentido, la participación del IIDE como socio implementador del Estudio de Factores Asociados del Cuarto Estudio Regional Comparativo y Explicativo (ERCE 2019) representa la oportunidad de abonar, bajo una perspectiva metodológica robusta, hacia el análisis de la calidad educativa bajo un enfoque de equidad, la cual va acorde con los Objetivos de Desarrollo Sustentable, preocupación latente en la región latinoamericana.

A modo de postulación, el IIDE cuenta con un grupo de investigadores con trayectoria relevante y experiencia amplia en el desarrollo y validación de evaluaciones educativas de escala pequeña y grande. En especial, en el campo de las evaluaciones con el propósito de aportar información de calidad a la toma de decisiones y la mejora del sistema educativo estatal. Dicha experiencia, facilita la coordinación de equipos especializados de trabajo, la gestión de recursos para el diseño, desarrollo y aplicación de una diversidad de estrategias evaluativas. Además, más que una obligación, es una tradición y rasgo distintivo de los proyectos del IIDE, el apego a estándares internacionales de calidad. Lo anterior, ha significado para los investigadores del Instituto el reconocimiento y respeto de la comunidad académica nacional e internacional en el campo de la evaluación educativa. Es por ello, que los investigadores del IIDE consideran una oportunidad y privilegio el poder participar en el proceso de selección para formar parte de los socios implementadores del Cuarto Estudio Regional Comparativo y Explicativo (ERCE 2019).

II. Propuesta metodológica, enfoque y plan de acción

Enfoque y características generales de la propuesta metodológica

La presente propuesta metodológica ofrece mejoras e innovaciones que incidan en la reducción de errores de diseño, desarrollo y aplicación de los instrumentos, y que aportan a la definición de una ruta crítica y visión a futuro de mejora progresiva. Además, se atienden las recomendaciones relacionadas con la optimización de la medida; la estandarización de los procesos de aplicación, captura y análisis eficiente de los resultados; el abordaje de temas relevantes para garantizar una educación inclusiva, equitativa, de calidad y que promueva oportunidades de aprendizaje durante toda la vida; la referencia explícita a las niñas, las personas con discapacidades y los pueblos indígenas dentro del idioma objetivo; la incorporación de variables pertinentes para el propósito de comparación entre países; el uso de los resultados de autoridades educativas y diseñadores de políticas educativas nacionales y regionales; la toma de decisiones de actores educativos de primer orden de relevancia para la mejora educativa; y la actualización de procesos y técnicas de análisis en consonancia con los avances y métodos de vanguardia en el campo de la evaluación educativa de gran escala (ver UNESCO-OREALC, 2016a).

En general, la metodología que se propone integra en un modelo de desarrollo y validación la participación de comités de expertos (ver Contreras, 2004; y Nitko, 1994), el aseguramiento de la calidad técnica de los instrumentos con la implementación de un programa fuerte de validez (ver OECD, 2012; Mislevy, 2006; y Messick, 1989), la pertinencia de los resultados mediante la definición de esquemas de toma de decisiones de actores educativos relevantes para la mejora educativa, la inclusión, la equidad y la calidad de la educación en la región.

Plan de acción

El plan de acción de la propuesta metodológica, se puede ver en la Tabla 1, el conjunto de actividades que se prevén realizar a lo largo del proyecto de desarrollo y la validación de las pruebas del logro de aprendizaje. Para llevar a cabo el plan de acción se conformará un equipo base que podrá extenderse sin problema alguno en caso de ser necesario:

Tabla 1. Propuesta de modelo metodológico para el desarrollo de las Pruebas de Logro de Aprendizaje del ERCE

Fases	Actividades
Fase I. Coordinación con el socio implementador del Estudio de Análisis Curricular del ERCE	Revisar los productos del socio implementador del Estudio de Análisis Curricular del ERCE
	Discutir con el socio implementador del Estudio de Análisis Curricular del ERCE las observaciones y ajustes a los productos desarrollados
	Determinar un diseño de las especificaciones de las pruebas y de los ítems en particular
	Elaborar el manual para la construcción de ítems
	Presentar el informe de observaciones a los productos desarrollados por el socio implementador del Estudio de Análisis Curricular
Fase II. Construcción de instrumentos de evaluación de logros de aprendizaje	Determinar el perfil de los integrantes de los equipos nacionales para el desarrollo de ítems
	Capacitar (en línea y presencial) a los equipos nacionales de elaboración de ítems
	Capacitar en línea y presencial a los coordinadores
	Capacitar en línea y presencial a los expertos disciplinarios
	Capacitar en línea y presencial a los expertos en medición
	Supervisar la revisión que hagan los coordinadores a la construcción individual de los ítems
	Supervisar la revisión de los ítems entre pares
	Supervisar la revisión del experto disciplinar
	Supervisar la revisión del grupo de expertos en medición
	Supervisar la edición de los ítems
	Supervisar el diseño gráfico de los ítems
Fase III. Adaptaciones nacionales y traducción al portugués	Presentar el informe sobre el proceso de construcción y desarrollo de las pruebas
	Capacitar a los equipos nacionales para la adaptación de los ítems *
	Revisar las propuestas de adaptación
Fase IV. Aplicación del estudio piloto	Revisar la traducción al portugués
	Apoyar en el diagramado de instrumentos, elaboración de manuales para piloteo ¹ , protocolos de aplicación y supervisión, reportes de incidencias; así como en el diseño de protocolos de captura, procesamiento y análisis de resultados*
	Revisar los manuales de operación de software en conjunto con el proveedor y la Coordinación Técnica del LLECE y definir los principales contenidos a adaptar
	Preparar un manual para distribuir entre los países con los principales contenidos, aprobados por las contrapartes anteriormente establecidas
	Diseñar protocolos de integración y actuación de grupos de observadores nacionales e internacionales para vigilar la correcta aplicación de los instrumentos*
	Capacitar (en línea y presencial) y certificar a los aplicadores de los países para el piloteo de instrumentos
	Asesorar y supervisar el piloteo de los instrumentos
	Revisar y analizar los reportes de incidencias, así como el proceso de aplicación con el fin de detectar mejoras en el diseño de los ítems y errores de aplicación
	Presentar el informe de la aplicación del piloto para los países con calendario <i>norte</i>
	Capacitar a los equipos nacionales en la corrección de respuestas
	Supervisar el proceso de corrección de respuestas
Fase V Análisis de datos del estudio piloto y plan de análisis	Realizar los análisis psicométricos de las pruebas
	Presentar los resultados de los análisis psicométricos de las pruebas
	Atender las recomendaciones de mejora de los instrumentos que hagan los equipos nacionales
	Elaborar un plan de análisis de las pruebas
	Presentar informe de análisis de resultados y las pruebas definitivas

A continuación, se resaltan algunos de los elementos más relevantes de cada una de las fases propuestas.

¹ Se propone un esquema de 5 niveles de jerarquía para el piloteo y la aplicación conformado por coordinadores nacionales, coordinadores locales, coordinador de centro escolar (aplica exclusivamente al director) y aplicadores de instrumentos.

Fase I. Coordinación con el Socio Implementador del Estudio de Análisis Curricular del ERCE

Uno de los elementos que hacen al ERCE un proyecto de gran validez para la implementación de políticas educativas por parte de los países participantes, reside particularmente en la comparabilidad entre currículos de los sistemas educativos involucrados. Si bien el objetivo del proyecto depende en gran medida de asumir una relativa similitud en la formación básica de la región, es menester asegurar que las coincidencias curriculares entre los sistemas educativos de cada país se mantengan en un marco de elementos mínimos e indispensables. Ello, sin duda, abonará a la validez de los ítems por desarrollar y, en consecuencia, de los resultados que de dicha medición se emitan.

En ese sentido, el diálogo con el Socio Implementador encargado del Estudio de Análisis Curricular del ERCE debe enfocarse en realimentar, bajo una perspectiva de cooperación enmarcada en la necesidad de una medición relativamente equivalente entre sistemas educativos, la tabla de especificaciones derivada de las matrices de análisis de contenidos curriculares. La discusión con el socio implementador se enfocará, principalmente, en la concordancia del producto resultante respecto a la consistencia con la agenda internacional educativa establecida en los ODS, en particular el número cuatro, mediante el Marco de Acción de Educación 2030 (UNESCO, 2015).

Partiendo de la relevancia que las especificaciones de la prueba tienen en términos metodológicos para el proyecto en cuestión, el objetivo general de esta fase se enfoca en la identificación de la pertinencia de las especificaciones desarrolladas por el Socio Implementador encargado del Estudio de Análisis Curricular del ERCE. Para su cumplimiento se establecen los siguientes objetivos específicos.

- Identificar las diferencias de dominios respecto a TERCE, con la finalidad de asegurar la comparabilidad entre estudios.
- Determinar la congruencia de las especificaciones propuestas respecto a los ODS, en particular el número cuatro, relacionado con la característica inclusiva de la educación para promover oportunidades de aprendizaje.

El primero de los objetivos específicos pretende asegurar la equivalencia entre estudios, de tal suerte que este proceso de revisión representa una oportunidad de contar con evidencias de validez de los instrumentos. Para este análisis, se llevará a cabo la identificación de correspondencia TERCE-ERCE a partir de los siguientes criterios (Webb, 2012):

1. Concurrencia categórica. La intención es identificar si las categorías de contenido entre ambos elementos convergen. Este primer análisis implica la consideración de la definición conceptual de las áreas de evaluación en cada proyecto, así como del listado de contenidos en cada una de ellas.
2. Concurrencia de la profundidad del conocimiento. Este análisis se basa en la identificación de la profundidad del conocimiento en ambos estudios; por ejemplo, la identificación de la clasificación cognitiva expuesta en cada uno de los contenidos que conforman la tabla de especificaciones.
3. Correspondencia del rango de conocimiento. Se pretende identificar si el espectro de contenido en cada área evaluada mantiene una convergencia.

Para atender el segundo de los objetivos específicos de esta primera actividad, se aplicará una metodología basada en el análisis de contenido, orientado a determinar la correspondencia de los contenidos de las tablas de especificaciones de las áreas por evaluar respecto a las metas del ODS enfocado en una educación de calidad. En ese sentido, el insumo de clasificación del contenido de las especificaciones son, en principio, las 10 metas del ODS número cuatro. Por supuesto, la búsqueda de alienación entre ambos aspectos resalta un discurso de equidad, inclusión y justicia; lo cual es particularmente relevante dada la característica internacional del ERCE y la coyuntura política en la que se instaura.

La operación de la dinámica de análisis se apoyará en el uso de herramientas digitales para el repositorio de información generada, en el cual se viertan las observaciones de los participantes –tanto en la vertiente del enfoque de derechos en la educación, así como como en lo que respecta a elementos de comparabilidad con TERCE–. Adicionalmente, el informe que se desarrolle resaltará en su contenido, y en particular en sus conclusiones, las posibles implicaciones que la tabla de especificaciones pueda tener respecto a la comparación con el TERCE y el marco del derecho a la educación que los resultados reflejarán, en aras de asegurar el desarrollo de un insumo de rendición de cuentas apropiado para los sistemas educativos participantes.

La información generada en esta primera actividad se discutirá de manera presencial con el Socio Implementador encargado del Estudio de Análisis Curricular del ERCE, con fines de realimentar y preparar los insumos necesarios para el desarrollo de ítems. La discusión con el Socio Implementador encargado del Estudio de Análisis Curricular del ERCE favorecerá no solo

el aseguramiento de la correspondencia de la tabla de especificaciones respecto a los ODS y TERCE, sino también la perspectiva comprensiva de la evaluación, a través del diálogo abierto entre instancias implementadoras.

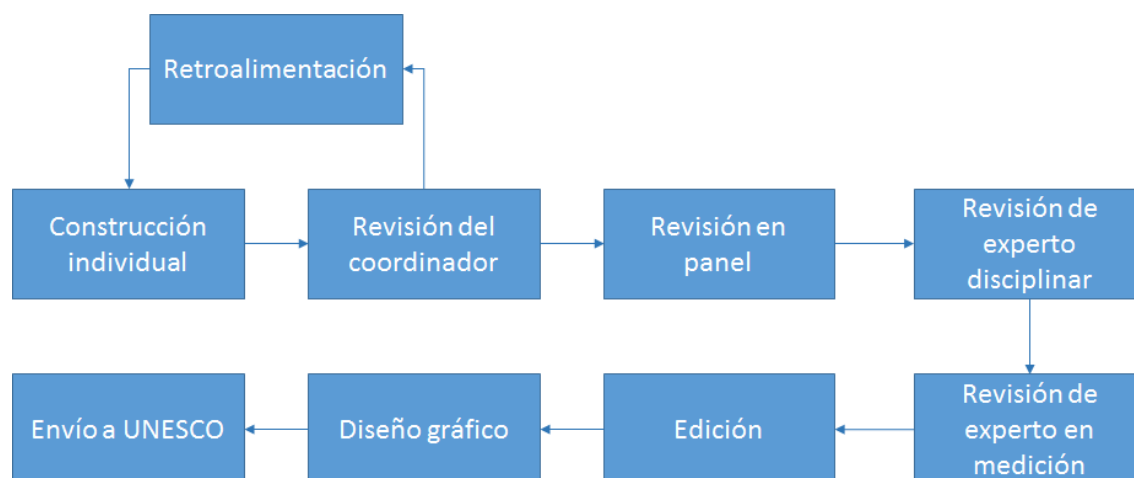
Dentro de las actividades necesarias para el desarrollo de ítems para el estudio, se plantea la determinación de un diseño de las especificaciones para las pruebas, así como especificaciones para los ítems, que serán los principales insumos de trabajo de las próximas fases. El diseño se propondrá en función de las necesidades de sistematización requerida en los procesos de evaluación a gran escala, que favorezcan la construcción de ítems, particularmente a partir de la necesidad de disminuir los errores de medida entre los sistemas educativos participantes.

Como parte de los productos generados por esta fase se elaborará un manual para la construcción de ítems, así como el informe de observaciones a los productos desarrollados por el socio implementador del Análisis Curricular del ERCE.

Fase II. Construcción de instrumentos de evaluación de logros de aprendizaje

El desarrollo de este compromiso requiere de la contratación de especialistas que fungirán como coordinadores de las áreas que comprende la evaluación: lenguaje, matemática y ciencias naturales. Cada uno de ellos deberá contar con una amplia experiencia en el desarrollo de evaluaciones relacionadas con cada una de las áreas por evaluar en el ERCE. Cada coordinador se encargará de dirigir los esfuerzos de los equipos de trabajo nacionales relacionados con el desarrollo de ítems de las áreas por evaluar (tanto de tercer y sexto grados), tomando como insumo principal el resultado final de la tabla de especificaciones, así como las especificaciones de ítems desarrolladas por el Socio Implementador encargado del Estudio de Análisis Curricular del ERCE. La organización transversal de cada uno de los equipos disciplinares de trabajo se realizará de la siguiente manera.

Diagrama 1. Flujo del proceso de construcción de ítems para cada área por evaluar.



Fuente: elaboración propia.

a) Construcción individual

Para el desarrollo de este elemento en particular se definirán los perfiles de los integrantes de los equipos nacionales que elaborarán los ítems. El perfil de participantes se dará a conocer a los diversos equipos nacionales.

A través de un taller presencial con los miembros de los equipos nacionales, cada coordinador de disciplina capacitará a los asistentes en el desarrollo de ítems, de acuerdo las especificaciones desarrolladas, así como con los lineamientos establecidos en ejercicios previos del proyecto de evaluación (UNESCO-OREALC, 2016). Como producto del taller se espera la realización de un mínimo de ítems que funjan como punto de partida para el trabajo individual de los asistentes a la capacitación.

El desarrollo del total de los ítems esperados por participante (en función de la tabla de especificaciones, la estructura de los instrumentos por desarrollar y el número de participantes en los equipos nacionales) se realizará a distancia. El envío de los ítems se realizará a través del uso de una plataforma informática, que fungirá como elemento de gestión, repositorio y retroalimentación de archivos, como más adelante se detalla.

Con la intención de identificar posibles sesgos en los ítems de respuesta construida, rescatar ejemplos de respuestas representativos, y validar la clasificación de pautas de corrección, se llevará a cabo un ejercicio de laboratorio cognitivo a través de técnicas de resolución en voz alta (Navas y Urdaneta, 2011), lo cual favorece la validez de la prueba y una capacitación de corrección adecuada para todos los países participantes. En ese sentido, se requiere el apoyo de los equipos nacionales para el desarrollo de un ejercicio de laboratorio

cognitivo previo a la capacitación en cuestión. En este ejercicio el equipo implementador recopilará y organizará los resultados, comunicando a través del manual de corrección las líneas de trabajo para la capacitación presencial. El desarrollo de este ejercicio requerirá que cada equipo nacional recopile información de las respuestas de niñas y niños de tercer y sexto grados a los ítems de respuesta construida de las áreas por evaluar (incluyendo escritura). De manera general, el desarrollo de este ejercicio implicará:

- Seleccionar de una muestra de estudiantes de tercero de primaria y otra de sexto; al menos un grupo de cada grado deberá estudiar en escuela pública y otro en escuela privada.
- Aplicar los ítems de respuesta construida a los estudiantes, solicitando –bajo un protocolo en común para los equipos nacionales– que los estudiantes hablen en voz alta lo que piensan una vez que se enfrentan al ítem (o a la tarea de escritura, en su caso).
- Contar con el consentimiento de las escuelas y los padres o tutores para grabar el audio de las respuestas de los estudiantes.
- Enviar al equipo implementador las respuestas de los estudiantes a los ítems en cuestión.

b) Revisión del coordinador

Como parte de la primera verificación de la validez elaboración de los ítems por desarrollar por parte de los equipos nacionales, el coordinador disciplinar se encargará de realizar una primera revisión de los elementos conceptuales de cada ítem elaborado, así como de su correspondiente convergencia con la tabla de especificaciones. La revisión se realizará con el apoyo de una lista de verificación dentro de la plataforma de gestión de ítems, con la posibilidad de introducir comentarios que retroalimenten la labor del desarrollo de ítems. Este proceso resulta iterativo entre quienes elaborarán los ítems y cada coordinador.

c) Revisión en panel

Una vez que el coordinador disciplinar verificó que los ítems desarrollados cubren los requisitos de calidad de contenido y estén en concordancia con la tabla de especificaciones, se realizarán diversas reuniones entre el coordinador y los miembros de los equipos nacionales. En ellas se validarán: la concordancia de cada ítem respecto a las especificaciones, precisión conceptual y elementos de redacción. Cada ítem deberá analizarse en panel.

d) Revisión mediante experto disciplinario

Para cada área por evaluar en la prueba, la versión de los ítems resultante de la revisión entre pares será enviada a un experto disciplinar para cada una de las áreas de la prueba. Se revisarán cada uno de los ítems con la ayuda de una lista de verificación que permita dar evidencia de la validez técnica y conceptual de cada uno de los ítems ajustados previamente en la revisión entre pares. Por supuesto, este proceso puede realizarse a distancia con el uso de la plataforma de gestión de ítems por desarrollar *ex professo* para este proyecto.

e) Revisión mediante un grupo de expertos en medición

El grupo de expertos en medición será el último filtro de validación de los ítems por incluir en las pruebas para la aplicación piloto. La discusión en panel sobre los ítems aceptados en fases previas permitirá asegurar un desarrollo en concordancia con las especificaciones elaboradas previamente.

f) Edición

La versión resultante de los ítems posterior a la revisión mediante expertos se enviará a una agencia contratada para que realice la revisión editorial. Por supuesto, el proceso de revisión y edición será realizado bajo el acompañamiento del personal del IIDE, en aras de asegurar que las modificaciones no alteren de forma sustancial la propuesta de ítems previamente revisada.

g) Diseño gráfico

Como parte final del proceso de desarrollo, los ítems –ya ajustados de estilo– se enviarán a una agencia de diseño gráfico que permita unificar elementos visuales respecto a versiones previas del estudio, así como en consideración de la edad de los participantes.

Las listas de verificación por desarrollar para valorar estos elementos en cada uno de los ítems en las diversas etapas del proceso deberán ser lo más detalladas posibles. Adicionalmente, se tomará en cuenta las consideraciones resultantes del proceso de desarrollo de TERCE para cada área evaluada (UNESCO-OREALC, 2016), con la finalidad de mejorar los ítems por desarrollar, a saber:

- Lectura:
 - Cuidar la diagramación de los textos.
 - Asegurar una adecuación visual de los textos discontinuos.

- Atender posibles problemas de distribución de información de los textos.
- Clarificar las instrucciones de los reactivos, relacionándolas siempre con los textos mostrados.
- Cuidar la extensión de las opciones de respuesta y el orden de las mismas.
- Las adaptaciones deben cuidar evitar errores léxicos de los textos originales.
- Matemática
 - Cuidar que las pautas de corrección de los ítems de respuesta construida eviten el modelo de penalización.
- Ciencias naturales
 - Evitar el desarrollo de ítems conceptuales.

Estas líneas permitirán generar instrumentos que orienten la verificación de *face validity* de los ítems por desarrollar, atendiendo así al aseguramiento de condiciones de medición acordes con el propósito del proyecto, la población objetivo, la diversidad cultural a la que se enfrenta, así como a las oportunidades de mejora detectadas en TERCE.

Fase III. Adaptaciones nacionales y traducción al portugués

El proceso de adaptación es uno de los ejes que aseguran la compatibilidad entre países de todos los ítems por incluir en la prueba. Para dicho proceso se hará uso de archivos digitales que se distribuirán a los equipos nacionales con fines de adaptación. Los archivos contendrán en filas los términos por adaptar, mientras que en las columnas se incluirán las argumentaciones de adaptación de los términos, así como los acuerdos logrados entre los equipos nacionales y el equipo del IIDE. El proceso de adaptación deberá seguir una lógica establecida en fases, a saber:

- a) Revisión y propuesta de adaptación de términos por parte de los equipos nacionales.
- b) Revisión de propuestas de adaptación por parte del equipo coordinador, mediante el apoyo de un experto en lingüística.
- c) Acuerdo de adaptación entre equipo nacional experto en lingüística.
- d) Revisión por parte del equipo coordinador del IIDE.
- e) Acuerdo final entre equipo nacional y equipo coordinador del IIDE

Una vez con los acuerdos establecidos con cada uno de los países participantes, cada equipo nacional es responsable del ajuste de los términos adaptados en sus versiones de ítems con fines de integración.

La adaptación se realizará mediante el uso de manuales desarrollados para tal fin, además que requerirá una reunión presencial con los equipos nacionales para su capacitación. En el caso de Brasil se requerirá la consideración de una contraparte local que funja como experto de adaptación e interlocutor con el equipo nacional de dicho país.

Fase IV. Aplicación del estudio piloto

Previo a la aplicación piloto se realizará una capacitación presencial dirigida a los equipos nacionales con el objetivo de presentar los lineamientos y protocolos necesarios para el desarrollo de la aplicación piloto, resaltando el uso y ajuste nacional de los manuales requeridos para ello, sugerencias de operación, registros de incidencias, entre otros. Una vez desarrollada la aplicación piloto, se revisarán y analizarán los reportes de incidencias, así como el proceso de aplicación para detectar mejoras en el diseño de los ítems y errores de aplicación.

Posterior a estas actividades, el equipo implementador, en conjunto con las coordinaciones disciplinares, desarrollarán un taller para la corrección de respuestas construidas para cada una de las áreas que comprende el estudio, con ejercicios prácticos para la corrección de ítems. El desarrollo de esta capacitación requerirá la elaboración de manuales para el piloteo.

Fase V. Análisis de datos del estudio piloto y plan de análisis

La labor del IIDE en el análisis de datos resultantes de la prueba piloto se enfocará en determinar los valores psicométricos de los ítems utilizados, en aras de delimitar las posibles modificaciones por realizar a cada uno de ellos. Como parte de ese proceso, se enviará a los equipos nacionales un reporte de análisis psicométrico con los ítems *dodgy*, con el fin de identificar posibles modificaciones que pretendan la mejora de los instrumentos para la versión definitiva. Adicionalmente, se entregará una propuesta de plan de análisis para la aplicación definitiva, junto con las versiones finales de los instrumentos de ERCE. El tiempo destinado para el desarrollo de las fases del proyecto, así como los productos por entregar se describen en la tabla 2:

Tabla 2. Cronograma y productos de las Pruebas de Logro de Aprendizaje del ERCE.

Fases	Actividades	Fechas de trabajo	Productos
Fase I. Coordinación con el socio implementador del Estudio de Análisis Curricular del ERCE	Revisar los productos del socio implementador del Estudio de Análisis Curricular del ERCE	Abril – junio 2017	<ul style="list-style-type: none"> ○ Informe 1. Observaciones a los productos desarrollados por el Socio Implementador del Estudio de Análisis Curricular
	Discutir con el socio implementador del Estudio de Análisis Curricular del ERCE las observaciones y ajustes a los productos desarrollados		
	Determinar un diseño de las especificaciones de las pruebas y de los ítems en particular		
	Elaborar el manual para la construcción de ítems		
	Presentar el informe de observaciones a los productos desarrollados por el socio implementador del Estudio de Análisis Curricular		
Fase II. Construcción de instrumentos de evaluación de logros de aprendizaje	Determinar el perfil de los integrantes de los equipos nacionales para el desarrollo de ítems	Junio – noviembre 2017	<ul style="list-style-type: none"> ○ Informe 2. Proceso de construcción y desarrollo de pruebas ○ Informe 3. Manual para la elaboración y revisión de ítems
	Capacitar (en línea y presencial) a los equipos nacionales de elaboración de ítems		
	Capacitar en línea y presencial a los coordinadores		
	Capacitar en línea y presencial a los expertos disciplinarios		
	Capacitar en línea y presencial a los expertos en medición		
	Supervisar la revisión que hagan los coordinadores a la construcción individual de los ítems		
	Supervisar la revisión de los ítems entre pares		
	Supervisar la revisión del experto disciplinar		
	Supervisar la revisión del grupo de expertos en medición		
Fase III. Adaptaciones nacionales y traducción al portugués	Supervisar la edición de los ítems	Diciembre 2017 – febrero 2018	<ul style="list-style-type: none"> ○ Informe 4. Metas de construcción para la fase piloto del estudio (en español y portugués)
	Supervisar el diseño gráfico de los ítems		
	Presentar el informe sobre el proceso de construcción y desarrollo de las pruebas		
Fase IV. Aplicación del estudio piloto	Capacitar a los equipos nacionales para la adaptación de los ítems	Marzo – diciembre 2018	<ul style="list-style-type: none"> ○ Informe 5. Manual de aplicación para la fase piloto, describiendo los procedimientos y las normas para la preparación, distribución, aplicación y captura de datos para la aplicación ○ Informe 6. Informe de aplicación del pilotaje
	Revisar las propuestas de adaptación		
	Revisar la traducción al portugués		
	Apoyar en el diagramado de instrumentos, elaboración de manuales para piloteo ² , protocolos de aplicación y supervisión, reportes de incidencias; así como en el diseño de protocolos de captura, procesamiento y análisis de resultados		
	Revisar los manuales de operación de software en conjunto con el proveedor y la Coordinación Técnica del LLECE y definir los principales contenidos a adaptar		
	Preparar un manual para distribuir entre los países con los principales contenidos, aprobados por las contrapartes anteriormente establecidas		
	Diseñar protocolos de integración y actuación de grupos de observadores nacionales e internacionales para vigilar la correcta aplicación de los instrumentos		
	Capacitar (en línea y presencial) y certificar a los aplicadores de los países para el piloteo de instrumentos		
	Asesorar y supervisar el piloteo de los instrumentos		
	Revisar y analizar los reportes de incidencias, así como el proceso de aplicación con el fin de detectar mejoras en el diseño de los ítems y errores de aplicación		
	Presentar el informe de la aplicación del piloto para los países con calendario norte		
	Capacitar a los equipos nacionales en la corrección de respuestas		
	Supervisar el proceso de corrección de respuestas		

² Se propone un esquema de 5 niveles de jerarquía para el piloteo y la aplicación conformado por coordinadores nacionales, coordinadores locales, coordinador de centro escolar (aplica exclusivamente al director) y aplicadores de instrumentos.

Fases	Actividades	Fechas de trabajo	Productos
Fase V	Realizar los análisis psicométricos de las pruebas		
Análisis de datos del estudio piloto y plan de análisis	Presentar los resultados de los análisis psicométricos de las pruebas		○ Informe 7- Documento que contiene las cinco pruebas (Matemática 3, Lenguaje 3, Matemáticas 6, Lenguaje 6 y Ciencias naturales 6)
	Atender las recomendaciones de mejora de los instrumentos que hagan los equipos nacionales		○ Informe 8. Plan de análisis
	Elaborar un plan de análisis de las pruebas		
	Presentar informe de análisis de resultados y las pruebas definitivas		

III. Planteamiento de valor agregado

Las mejoras e innovaciones metodológicas para el desarrollo, validación y aplicación de las pruebas de logro que se proponen en el presente documento, se encuentran basadas en experiencias de los investigadores del IIDE en el campo de la evaluación educativa a gran escala. Dichas mejoras e innovaciones metodológicas tratan de atender distintas recomendaciones presentes en los informes y reportes técnicos de TERCE, así como de observaciones de investigadores con conocimiento del proyecto que fueron recolectadas a lo largo de la elaboración de la presente propuesta. En especial, las mejoras e innovaciones metodológicas se pueden organizar en distintas líneas de acción:

- Implementación de un modelo de desarrollo y validación de ítems que cuida la coherencia respecto a estudios previos, así con el interés latinoamericano de realizar una evaluación que responda a los ODS. En este proceso se atiende como prioritario el diseño de especificaciones que faciliten la sistematización de la información por desarrollar por cada uno de los países participantes en el estudio.
- Definición de modelos de uso de resultados con base en esquemas de toma de decisiones de actores educativos de primer orden de relevancia para la mejora educativa.
- Definición de ruta crítica y visión a futuro con la definición de informes de resultados basados en modelos de usos y esquemas de tomas de decisiones de actores educativos de primer orden de relevancia para la mejora educativa, así como de la definición *a priori* de objetivos para estudios secundarios.
- Implementación de un programa de validez basado especialmente en evidencias del diseño de los ítems y la implementación de técnicas de pensamiento en voz alta en diferentes subpoblaciones para verificar el diseño y adaptaciones de los ítems, la equidad de la prueba, procesos y tiempos de respuesta.
- Elaboración de protocolos de actuación y manuales de aplicación.
- Programas de desarrollo de capacidades especializadas en evaluación educativa con mayor amplitud y alcance a los necesarios para el logro del proyecto³.

³ El IIDE puede proporcionar programas educativos desde cursos y talleres en línea como presenciales, así como diplomados y posgrados de maestría y doctorado en ciencias educativas con especial énfasis en evaluación educativa.

- Certificación y conformación de grupos de observadores nacionales e internacionales para vigilar la correcta aplicación de los instrumentos.

Referencias

- Contreras (2004) Metodología para elaborar exámenes criterios alienados al currículum. En S. Castañeda (Ed.). Educación, aprendizaje y cognición: teoría en la práctica. México, Manual Moderno.
- UNESCO-OREALC (2016a). Reporte Técnico. Tercer Estudio Regional Comparativo y Explicativo, TERCE. Santiago, Chile. UNESCO – OREALC.
- UNESCO-OREALC (2016b). Términos de referencia: Factores asociados. Cuarto Estudio Regional Comparativo y Explicativo (ERCE 2019). Santiago, Chile. UNESCO – OREALC.
- OECD (2012), PISA 2009 Technical Report, PISA, OECD Publishing.<http://dx.doi.org/10.1787/9789264167872-en>
- Koretz, D. M. (2010). El ABC de la evaluación educativa (Measuring Up), México: Ceneval
- Messick, S. (1989). Validity. En R. L. Linn (Ed.), Educational measurement (3rd ed.). (pp. 13-103). New York, NY, England: Macmillan Publishing Co, Inc.
- Mislevy, R. J. (2006). Cognitive psychology and educational assessment. In R. L. Brennan (Ed.), Educational Measurement (4th ed.) (pp.257-305). Wetsport, CT: American Council on Education/Paeger Publishers.
- Navas, M. y Urdaneta, E. (2011). PISA y el triángulo de la evaluación. Psicothema. 23(4). 701-706.
- Nitko, A. (1994, Julio). A model for developing curriculum-driven criterion referenced and norm-referenced national examinations for certification and selection of students. Ponencia presentada en la conferencia internacional sobre evaluación y medición educativas, de la Asociación para el Estudio de la Evaluación Educativa en Sudáfrica (ASSESA).
- Treviño et al. (2010). Factores asociados al logro cognitivo de los estudiantes de América Latina y el Caribe. Segundo Estudio Regional Comparativo y Explicativo. Oficina Regional de Educación para América Latina y el Caribe, Santiago, Chile.
- UNESCO-OREALC (2015). Flotts, M. Manzi, J. Jiménez, D. Abarzúa, A. Cayuman, C. y García, M. Informe de Resultados del Tercer Estudio Regional Comparativo y Explicativo TERCE. Logros de Aprendizaje. Santiago, Chile.
- Webb, N. (2012). Identificación del contenido para las pruebas de logro de estudiante. En: Downing, S. y Haladyna, T. (Comp.). Manual para el desarrollo de pruebas a gran escala. México: Ceneval.