

COMPETENCIA
LOGROS
LECTURA
MATEMÁTICA
MEDICIÓN
MUESTREO
DESEMPEÑO
PSICOMETRÍA
APRENDIZAJE
EDUCACIÓN
CALIDAD

Reporte técnico de las Evaluaciones Censales y Muestrales de Estudiantes 2018



COMPETENCIA
LOGROS
LECTURA
MATEMÁTICA
MEDICIÓN
MUESTREO
DESEMPEÑO
PSICOMETRÍA
APRENDIZAJE
EDUCACIÓN
CALIDAD

Reporte técnico de las Evaluaciones Censales y Muestrales de Estudiantes 2018





PERÚ

Ministerio
de Educación

Flor Aideé Pablo Medina

Ministra de Educación del Perú

Guido Alfredo Rospigliosi Galindo

Viceministro de Gestión Institucional

Ana Patricia Andrade Pacora

Viceministra de Gestión Pedagógica

José Carlos Chávez Cuentas

Secretario de Planificación Estratégica

Humberto Pérez León Ibáñez

Jefe de la Oficina de Medición de la Calidad de los Aprendizajes

Reporte técnico de las Evaluaciones Censales y Muestrales de Estudiantes 2018

Responsables del documento

Andrés Burga León (coordinador)

Luis Mejía Campos

Tania Pacheco Valenzuela

Yoni Arámbulo Mogollón

Gabriela Santibáñez Rojas

Yuriko Sosa Paredes

Wilmer Hernández Cabrera

Ángel Zegarra López

Esta publicación es el producto final del esfuerzo institucional de la UMC por medio de sus diferentes equipos de especialistas.

©Ministerio de Educación del Perú, 2019

Calle Del Comercio 193, San Borja

Lima, Perú

Teléfono: (511) 615-5800

www.minedu.gob.pe

Todos los derechos reservados. Prohibida la reproducción de este libro por cualquier medio, total o parcialmente, sin permiso expreso.

En el presente documento, se utilizan de manera inclusiva términos como “el docente”, “el estudiante” y sus respectivos plurales (así como otras palabras equivalentes en el contexto educativo) para referirse a hombres y mujeres. Esta opción se basa en una convención idiomática y tiene por objetivo evitar las formas para aludir a ambos géneros en el idioma castellano (“o/a”, “los/las” y otras similares), debido a que implican una saturación gráfica que puede dificultar la comprensión lectora.

Índice

Introducción	6
Capítulo 1: Diseño y construcción de las pruebas aplicadas	7
1.1 Construcción de los ítems y las pruebas	10
1.2 Procesos de validación de los instrumentos de evaluación	12
1.2.1 Validación de expertos de la UMC	13
1.2.2 Validación de campo: aplicación piloto	13
1.2.3 Análisis de las propiedades psicométricas y de la distribución de los ítems en la escala	14
1.2.4 Validación de expertos de otras oficinas del Minedu	14
1.2.5 Adaptación de instrumentos para estudiantes con discapacidad	15
1.2.6 Aplicación definitiva	15
1.3 Ensamblaje de las pruebas mediante el diseño bloques	16
1.3.1 Diseño de bloques de las evaluaciones censales	17
1.3.2 Diseño de bloques de las evaluaciones muestrales	25
Capítulo 2: Población y muestra	27
2.1 Evaluaciones censales	28
2.1.1 Población objetivo	28
2.1.2 Marco poblacional	29
2.1.3 Cobertura	30
2.1.4 Ajuste por no respuesta y cálculo de resultados	33
2.2 Evaluaciones muestrales	34
2.2.1 Diseño muestral	34
2.2.2 Cobertura	35
2.2.3 Pesos muestrales	36
Capítulo 3: Operativo de campo	39
3.1 Organización territorial	41
3.1.1 Divisiones jurisdiccionales para la ECE	41
3.1.2 Divisiones jurisdiccionales para la EM	43
3.2 Conformación de la red administrativa de aplicación	44
3.3 Proceso de selección y capacitación del personal de la RA	47
3.4 Características de los instrumentos y documentos utilizados	50
3.5 Proceso de aplicación	53
Capítulo 4: Procesamiento de datos y análisis psicométrico	61
4.1 Gestión y depuración de datos	62
4.1.1 Gestión física	62
4.1.2 Captura de datos	63
4.1.3 Depuración de datos	63

4.2	Procesamiento psicométrico	64
4.2.1	Modelo Rasch para ítems dicotómicos	64
4.3	Modelo Rasch de crédito parcial	67
4.4	Análisis de ítems	71
4.4.1	Calibración de ítems	73
4.4.2	Evidencias de validez vinculadas a la estructura interna de las pruebas	107
4.4.3	Confiabilidad y consistencia de la clasificación	109
4.4.4	Equiparación de medidas	113
4.5	Niveles de logro y presentación de resultados	124
	Referencias	129

Introducción

La Oficina de Medición de la Calidad de los Aprendizajes (UMC) de la Secretaría de Planificación Estratégica del Ministerio de Educación (Minedu) tiene entre sus funciones el diseño e implementación de las evaluaciones de logros de aprendizaje de los estudiantes de Educación Básica a nivel nacional. En ese marco, y de acuerdo a la RM 116-2018-Minedu, en el 2018 la UMC llevó a cabo la Evaluación Censal de Estudiantes (ECE) y la Evaluación Muestral (EM). La ECE evaluó, en 2.º y 4.º grado de primaria las áreas de Comunicación (Lee diversos tipos de textos escritos en su lengua materna) y Matemática (todas las competencias); en 4.º grado de primaria EIB, las áreas de Comunicación (Lee diversos tipos de textos escritos en su lengua materna) y Castellano como segunda lengua (Lee diversos tipos de texto escritos en castellano como segunda lengua); y en 2.º grado de secundaria, las área de Comunicación (Lee diversos tipos de textos escritos en su lengua materna), Matemática (todas las competencias), Ciencias Sociales (todas las competencias), y Ciencia y Tecnología (todas las competencias). Asimismo, la EM evaluó, en 2.º grado de secundaria, el área de Comunicación (Escribe diversos tipos de texto en su lengua materna) y la competencia de Ciudadanía (Convive y participa democráticamente) en 6.º grado de primaria.

En el primer capítulo de este documento, se describe el proceso de construcción de los instrumentos aplicados. Se señalan los criterios que se siguen en la elaboración de los ítems, el recojo de evidencias de validez vinculadas al contenido de los ítems, y el proceso de aplicación piloto y definitiva.

En el segundo capítulo, se describen la población objetivo y el marco muestral utilizado. Además, se señala cómo fueron definidos los tamaños muestrales y los pesos utilizados en la estimación de los diversos parámetros. Este capítulo incluye además el análisis de la cobertura.

El tercer capítulo aborda los aspectos relacionados con el operativo de campo. Describe la forma cómo se organizó la red administrativa, el modo cómo fueron seleccionados y capacitados los aplicadores, los instrumentos utilizados y la manera cómo se recolectaron los datos en el contexto de las evaluaciones 2018.

Finalmente, el cuarto capítulo describe el proceso de gestión y depuración de datos, y el modelo Rasch utilizado para los análisis psicométricos. Se pone énfasis en la forma cómo se analizaron los ítems, y cómo se obtuvieron evidencias de confiabilidad y validez. Además, se describe cómo se realizó el proceso de equiparación de medidas, la forma cómo se establecieron los niveles de logro a partir de los puntos de corte en las evaluaciones 2018.

Diseño y construcción de las pruebas aplicadas

Capítulo 1

La UMC cuenta con un equipo de evaluación conformado por distintos subequipos organizados por especialidades: Comunicación, Lenguas originarias, Ciudadanía y Ciencias Sociales, y Ciencia y Tecnología. Estos son los encargados de diseñar y construir las pruebas y demás instrumentos para las diversas evaluaciones y estudios que están bajo la responsabilidad de la UMC como instancia dependiente del Minedu. Además, la oficina cuenta con un subequipo de Atención a la Discapacidad que orienta y contribuye en la adaptación de los diversos instrumentos a las características particulares de los estudiantes con discapacidad incluidos en las escuelas de Educación Básica Regular (EBR) con la finalidad de que ellos puedan participar de las evaluaciones. Cada año, las pruebas que se aplican en el marco de la ECE y/o de la EM son especificadas mediante una resolución ministerial que es publicada en los primeros meses del año en el diario oficial El Peruano. Así, en 2018, mediante la R.M. 116-2018-Minedu, se aprobó la implementación de la ECE y de la EM tal como se detalla en las tablas 1.1 y 1.2.

Tabla 1.1 Evaluaciones censales implementadas por la UMC en el 2018

Nivel	Grado	Área	Competencias a evaluar	Referente curricular
Primaria	4.º	Matemática	Resuelve problemas de cantidad.	Currículo Nacional de la Educación Básica RM 281-2016-MINEDU
			Resuelve problemas de regularidad, equivalencia y cambio.	
	Resuelve problemas de forma, movimiento y localización.			
	Resuelve problemas de gestión de datos e incertidumbre.			
		Comunicación	Lee diversos tipos de textos escritos en lengua materna.	
	4.º EIB	Comunicación	Lee diversos tipos de textos escritos en lengua materna: Aimara Ashaninka Awajún Quechua chanca Quechua cusco collao Shipibo-konibo Lee diversos tipos de textos escritos en castellano como segunda lengua.	

Tabla 1.1 Evaluaciones censales implementadas por la UMC en el 2018

Nivel	Grado	Área	Competencias a evaluar	Referente curricular
Secundaria	2.º	Matemática	Actúa y piensa matemáticamente en situaciones de cantidad.	Diseño Curricular de la Educación Básica Regular RM 0440-2008-ED, modificado parcialmente por la RM 199-2015-MINEDU
			Actúa y piensa matemáticamente en situaciones de regularidad, equivalencia y cambio.	
			Actúa y piensa matemáticamente en situaciones de forma, movimiento y localización.	
			Actúa y piensa matemáticamente en situaciones de gestión de datos e incertidumbre.	
		Comunicación	Comprende textos escritos.	
Ciencias Sociales (Historia, Geografía y Economía)	<p>Construye interpretaciones históricas.</p> <p>Actúa responsablemente en el ambiente.</p> <p>Actúa responsablemente respecto a los recursos económicos.</p>			
Ciencia y Tecnología	<p>Indaga mediante métodos científicos, situaciones que pueden ser investigadas por la ciencia.</p> <p>Explica el mundo físico basado en conocimientos científicos.</p> <p>Diseña y produce prototipos tecnológicos para resolver problemas de su entorno.</p> <p>Construye una posición crítica sobre la ciencia y la tecnología en sociedad.</p>			

Tabla 1.2 Evaluaciones muestrales implementadas por la UMC en el 2018

Nivel	Grado	Área	Competencias a evaluar	Referente curricular
Primaria	6.º	Personal Social	Ciudadanía (Convive y participa democráticamente).	Currículo Nacional de la Educación Básica RM 281-2016-MINEDU
Secundaria	2.º	Comunicación	Produce textos escritos.	Diseño Curricular de la Educación Básica Regular RM 0440-2008-ED, modificado parcialmente por la RM 199-2015-MINEDU

1.1 Construcción de los ítems y las pruebas

Las pruebas de la ECE y de la EM están alineadas a los documentos curriculares vigentes; es decir, se construyen considerando las competencias y las capacidades establecidas en el currículo. Además, el diseño y la construcción tanto de los ítems como de las pruebas siguen los principios de validez, confiabilidad y diseño universal de evaluación. Estos establecen que los instrumentos de evaluación deben recoger información de los estudiantes de tal manera que se pueda estimar de forma fiable su nivel de logro de las competencias evaluadas y que dicha información pueda ser usada para los fines propios del proceso educativo (American Educational Research Association, American Psychological Association y National Council of Measurement in Education, 2014). Asimismo, debe reflejar una concepción inclusiva de la educación, conforme a los lineamientos de la política educativa nacional (Ministerio de Educación, 2015).

En la ECE y en la EM, se emplean ítems de dos formatos: de opción múltiple y de respuesta construida. En seguida, se describen estos formatos de ítems.

Ítem de opción múltiple. Este formato de ítem está constituido por un estímulo, un enunciado o una pregunta que debe ser respondida eligiendo solamente una opción (o alternativa) de entre varias que se presentan. Esto implica que una de las opciones funciona como clave (la que responde o completa correctamente el enunciado o pregunta), mientras que el resto de las alternativas funcionan como distractores (respuestas incorrectas). En segundo grado de secundaria, alguno de los distractores puede constituir una respuesta intermedia o incompleta y, por lo tanto, modelada como un ítem de crédito parcial (puntaje parcial).

Para cada ítem de opción múltiple, la construcción de su enunciado y sus alternativas sigue las pautas ofrecidas por Moreno, Martínez y Muñiz (2004). De acuerdo con lo propuesto por estos autores, se consideran las siguientes pautas:

Respecto de la construcción del ítem en general

- Debe haber correspondencia entre el ítem y el indicador que evalúa.
- La complejidad de la tarea solicitada debe ser adecuada para el estudiante según el grado.
- El ítem debe estar libre (en la medida de lo posible) de todo sesgo sociocultural, socioeconómico, de género o de religión.
- El ítem no debe estar relacionado con ningún otro ítem de la prueba; es decir, la respuesta correcta no puede inferirse a partir de otro ítem.

Respecto de la construcción del enunciado

- Lo central debe expresarse en el enunciado. Cada opción es un complemento de este.
- El enunciado debe concordar gramaticalmente con todas las alternativas.
- La sintaxis o estructura gramatical debe ser clara y precisa; es decir, el enunciado debe estar claramente redactado. Se deben evitar enunciados demasiado escuetos, profusos, ambiguos o confusos.
- En lo posible, debe evitarse redactar el enunciado de forma negativa. Si hubiera necesidad de usar una expresión negativa, esta debe resaltarse.
- El enunciado debe estar libre de elementos irrelevantes e innecesarios.

Respecto de la construcción de las alternativas

- La opción correcta debe ser solo una y debe estar acompañada por distractores plausibles.
- A lo largo de la prueba, la opción correcta debe estar distribuida entre las distintas ubicaciones.
- Las opciones deben ser preferiblemente tres o cuatro.
- Las opciones deben disponerse, de preferencia, de forma vertical.
- El conjunto de opciones de cada ítem debe aparecer estructurado.
- Las opciones deben ser autónomas entre sí, sin solaparse ni referirse unas a otras.
- Se deben evitar en todo momento las opciones “Todas las anteriores” y “Ninguna de las anteriores”.
- Ninguna opción debe destacar entre las demás, ni en contenido ni en apariencia.
- Las alternativas deben estar libres de palabras, frases o imágenes irrelevantes.
- Se deben emplear como distractores los errores posibles y más frecuentes en la resolución del ítem.

Ítem de respuesta construida. Este formato de ítem se compone de un estímulo y de un enunciado que demanda una tarea al estudiante evaluado. Se espera que el estudiante redacte su respuesta. La inclusión de este tipo de ítem responde a la necesidad de explorar tanto los procedimientos involucrados en la resolución de un ítem como las habilidades complejas de argumentación, evaluación, juicio crítico y toma de decisiones razonada. Las respuestas de los estudiantes a este tipo de ítem son codificadas por codificadores expertos en el área y especialmente capacitados para este fin. Estos, de preferencia, poseen experiencia en procesos similares. Los

codificadores pasan por un riguroso proceso de capacitación que les permite apropiarse de la lógica de los ítems, los criterios de codificación y las posibles respuestas. Tanto para el proceso de capacitación previa como para la codificación en sí, se utiliza un manual de codificación que especifica el indicador y el descriptor que el ítem evalúa, la descripción de la respuesta correcta y, al menos, cuatro ejemplos de respuestas reales de los estudiantes. También, especifica la descripción de las respuestas incorrectas y sus respectivos ejemplos. Asimismo, algunos de los ítems de este tipo admiten respuestas incompletas o que reflejan un logro parcial del estudiante en la tarea. En este caso, se describe también en el manual la respuesta parcial y se incluyen ejemplos de respuestas parciales.

En las ECE, los ítems de respuesta construida se encuentran en menor proporción que los ítems de opción múltiple; esto debido a la complejidad del procesamiento de las respuestas a este tipo de ítems en procesos masivos. En cambio, en la EM, debido al menor volumen de respuestas a codificar, se incluye una mayor cantidad de ítems de respuesta construida.

1.2 Procesos de validación de los instrumentos de evaluación

Estos procesos se realizan también para obtener evidencias de validez vinculadas al contenido de las pruebas. Este tipo de evidencias de validez se vincula a la necesidad de garantizar que el test sea una muestra adecuada y representativa de los contenidos que pretende evaluar (McGartland, Berg-Weger, Tebb, Lee y Rauch, 2003; Muñiz, 1998). En esta misma línea, Suen (1990) indica que los ítems en un test, supuestamente, constituyen una muestra representativa de todos los ítems que se pueden emplear para medir el constructo de interés. Si este supuesto se cumple, entonces se podrá decir que la prueba tiene adecuadas evidencias de validez referidas a su contenido, quedando esta definida como la representatividad de los ítems empleados en un test para medir un constructo particular. La representatividad de un test puede definirse como la precisión con que se pueden hacer inferencias acerca de la puntuación de la persona en el universo de indicadores a partir de la puntuación en la muestra de ítems que constituyen la prueba (Paz, 1996).

Muñiz (1998) señala que la práctica más usual para lograr este tipo de validez consiste en enumerar todas las áreas de contenido que se consideren importantes o imprescindibles y, luego, asegurarse de que la prueba contenga ítems que hagan referencia a cada una de ellas en una proporción adecuada. En ese sentido, el proceso de validación implica el mismo proceso de construcción, pues se debe seguir una tabla de especificaciones. En los siguientes acápite se describen los procesos de validación que se han realizado.

1.2.1 Validación de expertos de la UMC

Después de que los equipos de especialistas de las áreas (de Matemática, de Comunicación, de Ciudadanía y Ciencias Sociales, de Ciencia y Tecnología, y de Lenguas originarias) han construido los ítems, estos son revisados por expertos de la UMC. Para esto, se cuenta con guías que orientan el proceso y fichas que permiten evaluar cada ítem en aspectos como la calidad, vigencia y veracidad de la información según cada disciplina científica, la correspondencia con la tabla de especificaciones, la adecuación de la complejidad del ítem a la población evaluada, y la construcción del enunciado y las alternativas, tanto en lo formal como en su eficacia para la medición del constructo a evaluar. Asimismo, se revisa que no se presenten sesgos socioeconómicos, culturales ni de género en la construcción de los ítems. En el caso de las pruebas de Lectura, correspondientes al área de Comunicación (cuarto grado de primaria, segundo grado de secundaria y cuarto grado de primaria EIB), se revisa y evalúa también la pertenencia de los textos a los tipos y géneros textuales señalados, la densidad de la información, el vocabulario, la sintaxis, la familiaridad del tema, la extensión y el formato.

1.2.2 Validación de campo: aplicación piloto

Después de la revisión de expertos de la UMC, los ítems que no presentan problemas pasan a ser empleados para la construcción de las pruebas destinadas a la aplicación piloto. Esta consiste en la aplicación de las pruebas a una muestra representativa de estudiantes a nivel nacional. Un diseño estadístico permite determinar las IE que forman parte de la muestra de la aplicación piloto. De acuerdo con el procedimiento estadístico, las IE son elegidas aleatoriamente y corresponden a los distintos estratos (según las características determinadas en el diseño): rural y urbano, estatal y no estatal, polidocente y multigrado/unidocente. Asimismo, cabe señalar que los estudiantes de estas IE presentan similares condiciones a las de la población objetivo (edad, avance curricular, desarrollo de capacidades, etc.). Además, las pruebas piloto presentan estructura y características similares (orden de presentación de los ítems según su dificultad; extensión, dificultad, diagramación, tiempo de aplicación de la prueba) a las de las pruebas definitivas.

La aplicación piloto busca recoger información que permita asegurar que los ítems, en particular, y los instrumentos en su conjunto presentan adecuadas evidencias de validez referidas a cada uno de los constructos que se evalúan en las diversas pruebas. Asimismo, se recogen evidencias acerca del funcionamiento de los ítems y su ajuste al modelo de medición empleado en la UMC.

1.2.3 Análisis de las propiedades psicométricas y de la distribución de los ítems en la escala

Sobre la base de los resultados de la aplicación piloto, se analizan las propiedades y otros aspectos psicométricos de las pruebas y los ítems; además, se realiza una revisión pedagógica de la distribución de los ítems en la escala. Entre los aspectos psicométricos, se analizan el ajuste de los ítems y de los distractores, la confiabilidad de las medidas, la unidimensionalidad, la correspondencia entre la media de habilidad de los estudiantes y la media de dificultad de cada prueba, y el comportamiento diferencial de los ítems según los distintos estratos (urbano y rural, estatal y no estatal, hombre y mujer). Igualmente, se ensayan hipótesis sobre las posibles causas por las que un ítem no ajustó al modelo psicométrico, tuvo un funcionamiento diferencial según los estratos o su dificultad no concordó con la complejidad propuesta. Para una mayor profundización de este aspecto, en el capítulo 4 de este documento, se detallan dichas propiedades psicométricas.

En la revisión pedagógica, se analiza la correspondencia entre la complejidad pedagógica estimada de un ítem (antes de la aplicación piloto) y la dificultad dada por el modelo psicométrico (después de la aplicación piloto), la pertinencia de incluir un determinado ítem en la prueba definitiva y las posibles causas por las que los estudiantes no contestaron correctamente un ítem. En el caso de Comunicación, tanto en primaria como en secundaria, también se analizan los textos para encontrar posibles explicaciones a las respuestas no esperadas o a un error recurrente de los estudiantes. Este proceso de análisis concluye con la selección de los ítems de mejor desempeño y con la modificación de los ítems que presentaron algún tipo de problema con vistas a la aplicación definitiva de las pruebas.

1.2.4 Validación de expertos de otras oficinas del Minedu

Según lo establecido por la UMC, el proceso de validación incluye que un conjunto de expertos en cada una de las áreas evaluadas emita su juicio respecto de la construcción, la correspondencia con el indicador y la dificultad de la tarea solicitada en cada ítem empleado para medir el constructo implicado (criterio de jueces). Para ello, los ítems seleccionados a partir de la aplicación piloto y el análisis posterior son organizados en formas o cuadernillos para que sean validados por expertos de las distintas direcciones del Minedu involucradas, tales como la Dirección de Educación Primaria (DEP), la Dirección de Educación Secundaria (DES), la Dirección General de Educación Básica Alternativa, Intercultural Bilingüe y de Servicios Educativos (Digeibira)¹. En el proceso de validación, también participan especialistas de la Dirección de Evaluación Docente (DIED) por su experiencia en la

¹La prueba de cuarto grado de primaria EIB, además, es revisada por parte de especialistas UGEL de las regiones correspondientes a las lenguas originarias evaluadas, así como docentes de aula. Asimismo, los ítems de esta prueba son sometidos a un proceso de doble traducción (con expertos traductores propios de las culturas originarias) con la finalidad de que sean pertinentes a las características socioculturales y lingüísticas de los estudiantes evaluados.

construcción de ítems y su nivel de aporte relacionado con esa experiencia. En esta validación, se solicita a los jueces que evalúen los aspectos mencionados en una ficha técnica.

Esta ficha cuenta con cuatro secciones: tres en las que se presentan los indicadores que evalúan tanto las características generales del ítem como las características específicas del enunciado y de las alternativas, y una sección en que el juez escribe sugerencias puntuales sobre el ítem. Las fichas presentan ciertas particularidades de acuerdo a las características de la competencia que evalúan. Por ejemplo, en el caso de Lectura (área de Comunicación), se incluyen indicadores referidos a cada uno de los textos y no solamente a los ítems. Cada experto completa la ficha y formula las observaciones que, de acuerdo con su juicio, estime convenientes. Luego, las fichas son sistematizadas, los cuestionamientos absueltos y las sugerencias incorporadas. Ello implica, en algunos casos, realizar algún tipo de ajuste al ítem o al texto.

1.2.5 Adaptación de instrumentos para estudiantes con discapacidad

Desde el 2015, la UMC adapta (es decir, realiza cambios) los instrumentos de evaluación de las distintas competencias evaluadas, así como los procedimientos de aplicación de las pruebas. Estas adaptaciones se realizan con el fin de asegurar la participación plena de los estudiantes con discapacidad sensorial (baja visión, ceguera, hipoacusia y sordera), discapacidad motora y trastorno del espectro autista de las pruebas nacionales a gran escala.

1.2.6 Aplicación definitiva

La aplicación definitiva de la ECE y la EM se realizó en el último trimestre del año (octubre-noviembre-diciembre). En el caso de la ECE, la aplicación se llevó a cabo durante dos días en 4.º grado de primaria y tres días en 2.º grado secundaria. En el primer día, tanto en 4.º grado de primaria como en 2.º grado de secundaria, se aplicaron, primero, las pruebas de Lectura y, luego, las de Matemática. En el segundo día, el orden de aplicación de las pruebas es inverso. Este cambio obedece a la necesidad de neutralizar los posibles efectos del cansancio de los estudiantes, quienes rinden dos pruebas en un día. En el caso de secundaria, el tercer día se destinó a la aplicación de las pruebas de Ciencias Sociales, y Ciencia y Tecnología. Asimismo, las pruebas de Ciudadanía (Convive y participa democráticamente) en 6.º grado de primaria y de Escritura en 2.º grado de secundaria se aplicaron en un solo día.

A partir del 2015, los instrumentos de la ECE, incluidos los cuadernillos de las pruebas, retornan en su totalidad al Minedu y se mantienen bajo confidencialidad. Este procedimiento, aunque complicado, se realiza porque permite alimentar un banco de ítems creciente y asegurar la comparabilidad de resultados en el tiempo. Asimismo, cabe señalar que las aplicaciones piloto permiten renovar un porcentaje

de los ítems de las pruebas cada año, sobre todo aquellos ítems que se liberan, es decir, que son hechos públicos en los reportes de resultados que produce y distribuye la UMC.

En resumen, mediante los procesos de construcción y validación, se logró que los instrumentos de evaluación de la ECE y la EM 2018 cumplan con los requerimientos técnicos y consideraciones de contenido que permiten recoger datos válidos y confiables sobre los niveles de logro de los estudiantes evaluados. Asimismo, se aseguró que los instrumentos guarden correspondencia con los diversos documentos curriculares del sistema educativo y que sus resultados sean comparables a lo largo del tiempo.

1.3 Ensamblaje de las pruebas mediante el diseño bloques

Conforme avanzan en su escolaridad, los estudiantes amplían y profundizan los aprendizajes logrados en los grados anteriores. Así, por ejemplo, en Lectura, el lenguaje de los textos se complejiza y se especializa progresivamente, los temas tratados son más diversos, y las capacidades de inferencia y reflexión se consolidan. Por su parte, en Matemática, las nociones de número, variación y espacio, en un inicio ligadas a situaciones más particulares y objetos concretos, demandan de mayor formalización, abstracción y aprendizaje de conceptos nuevos, necesarios para seguir construyendo el pensamiento matemático.

Por esta razón, es fundamental proponer una evaluación que cubra un amplio espectro de capacidades y conocimientos, de manera que los resultados recojan de la manera más precisa lo que realmente logran hacer los estudiantes. Sin embargo, las evaluaciones de carácter censal suelen enfrentar la restricción de tener que medir con una prueba única (y las consiguientes limitaciones del tiempo de evaluación y del número de ítems) las habilidades de los estudiantes. Para lidiar con esta restricción, en la evaluación —de 4.º grado de primaria en adelante—, se estudió la posibilidad de utilizar diseños que contengan bloques de ítems comunes, que sean aplicados por todos los estudiantes, y bloques no comunes, que sean aplicados solo por un conjunto de estudiantes. La combinación de bloques comunes y no comunes debía permitir el ensamblaje de formas (o “cuadernillos”) que se distribuyan aleatoriamente entre la población. De este modo, y gracias al uso de técnicas de equiparación del modelo Rasch, una prueba podría contar con mayor cantidad de ítems en la misma métrica, sin que ello implicara incrementar la cantidad de ítems a los que cada estudiante se enfrenta.

Con la asistencia técnica del Berkeley Evaluation and Assessment Research Center (BEAR), de la Universidad de California (Berkeley), se llevó a cabo un conjunto de estudios psicométricos que plantearon diversos escenarios en los que se combinaban bloques de ítems de maneras distintas para ensamblar las formas de las pruebas, a fin de ampliar el número total de ítems aplicados (aunque no todos los estudiantes

respondieran el total de estos ítems). En estos escenarios se consideraban bloques comunes, bloques rotados y un número determinado de ítems por bloque. A partir de estos escenarios, se realizaron simulaciones estadísticas con la base de datos de la aplicación piloto, para verificar que la confiabilidad de las personas en estos diseños se mantuviera estable en múltiples iteraciones, de manera que sea factible, dentro de márgenes razonables, reportar resultados a nivel individual y/o de agregado por escuela. En todos los escenarios planteados, se obtuvieron niveles altos en la media de la confiabilidad (entre 0,90 y 0,96), así como en la media de la correlación entre el valor verdadero del rasgo latente y su valor estimado (entre 0,95 y 0,96).

En relación con la conformación de los bloques, es preciso señalar que el bloque común tiene por finalidad optimizar la equiparación de las medidas y es construido de tal forma que los ítems reflejen la tabla de especificaciones, así como la dispersión de las medidas a lo largo de la escala de dificultad.

Por otro lado, en una evaluación muestral es posible utilizar diseños de bloques balanceados incompletos (BIBD, por sus siglas en inglés) sin las restricciones que supone una devolución de resultados a nivel de estudiante. Esta ventaja ha sido aprovechada para el diseño de las pruebas de Lectura y Matemática de segundo de primaria y de Ciudadanía de sexto de primaria.

En la EM de Escritura en 2.º grado de secundaria, se empleó un diseño de prueba única. Esta prueba planteó a los estudiantes escribir cuatro textos de diferente tipo: narrativo (anécdota y noticia), descriptivo (artículo descriptivo) y argumentativo (carta argumentativa formal) a partir de cuatro consignas. Cada consigna describe una situación comunicativa y especifica la tarea que se debe realizar (qué tipo de texto se debe escribir, para qué destinatario, en qué género debe encuadrarse y qué extensión debe tener).

Por último, en el caso de las pruebas con adaptaciones para evaluar estudiantes con discapacidad, también se ha considerado el diseño de prueba única.

1.3.1 Diseño de bloques de las evaluaciones censales

En las siguientes tablas, se presenta el diseño de bloques de las pruebas aplicadas a nivel censal. Las tablas de especificaciones de estos instrumentos pueden ser consultadas en los anexos.

Diseño de bloques de las pruebas de 4.º grado de primaria

La prueba de Lectura de 4.º grado de primaria elaborada para la ECE 2018 cuenta con un total de 12 textos de diferente tipo, género, formato y complejidad. A partir de este conjunto de textos, se elaboraron 72 ítems. Dichos textos e ítems se distribuyeron en 6 bloques comunes a todas las formas y 6 bloques no comunes. Las diferentes combinaciones de los bloques permitieron ensamblar 6 formas que

reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

Tabla 1.3 Matriz de bloques de la prueba de Lectura de 4.º grado de primaria. Día 1

Forma	Bloque	Ítems por forma
1	B01	24
	B07	
	B02	
	B03	
2	B01	24
	B02	
	B08	
	B03	
3	B01	24
	B03	
	B02	
	B09	

Tabla 1.4 Matriz de bloques de la prueba de Lectura de 4.º grado de primaria. Día 2

Forma	Bloque	Ítems por forma
4	B04	24
	B10	
	B05	
	B06	
5	B04	24
	B11	
	B05	
	B06	
6	B04	24
	B06	
	B05	
	B12	

De las 6 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 72 ítems que tiene la prueba, cada estudiante respondió 48.

La prueba de Matemática de 4.º grado de primaria elaborada para la ECE 2018

cuenta con un total de 90 ítems. Dichos ítems se distribuyen en 2 bloques comunes a todas las formas y 6 bloques no comunes. Las diferentes combinaciones de los bloques permitieron ensamblar 6 formas que reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

Tabla 1.5 Matriz de bloques de la prueba de Matemática de 4.º grado de primaria. Día 1

Forma	Bloque	Ítems por forma
1	B01	25
	BC1	
2	B02	25
	BC1	
3	B03	25
	BC1	

Tabla 1.6 Matriz de bloques de la prueba de Matemática de 4.º grado de primaria. Día 2

Forma	Bloque	Ítems por forma
4	B04	25
	BC2	
5	B05	25
	BC2	
6	B06	25
	BC2	

De las 6 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 90 ítems que tiene la prueba, cada estudiante respondió 50.

Diseño de bloques de las pruebas de 4.º grado de primaria EIB

La evaluación de los estudiantes de Educación Intercultural Bilingüe supone una aplicación diferenciada, según el grupo lingüístico al que pertenezcan. Así, los estudiantes hablantes maternos de aimara, ashaninka, quechua cusco collao, quechua chanca y shipibo-konibo son evaluados en Lectura en su lengua originaria y en castellano como segunda lengua durante dos días de aplicación. Cada día, aplican una forma de Lectura en lengua originaria de 22 ítems y otra de Lectura en castellano como segunda lengua, también de 22 ítems. En cambio, los estudiantes de los demás grupos lingüísticos son evaluados solo en Lectura en castellano como segunda lengua en un solo día de aplicación en dos sesiones. En la primera sesión,

resuelven un conjunto de 22 ítems, y en la segunda, otro conjunto, también de 22 ítems.

Lectura en lengua originaria y en castellano como segunda lengua

Tanto la prueba de Lectura en lengua originaria (aimara, ashaninka, quechua cusco collao, quechua chanca y shipibo-konibo) como la de castellano como segunda lengua elaboradas para la ECE 2018 cuentan con un total de 12 textos de diferente tipo, género, formato y complejidad cada uno. A partir de este conjunto de textos, se han elaborado 68 ítems para cada prueba. Dichos textos e ítems se distribuyeron en 2 bloques comunes a todas las formas y 8 bloques no comunes. Las diferentes combinaciones de los bloques permitieron ensamblar 4 formas que reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

Tabla 1.7 Matriz de bloques de la prueba de Lectura (en lengua originaria y castellano como segunda lengua). Día 1

Forma	Bloque	Ítems por forma
1	BC1	22
	B01	
	BC2	
	B02	
2	BC1	22
	B03	
	BC2	
	B04	

Tabla 1.8 Matriz de bloques de la prueba de Lectura (en lengua originaria y castellano como segunda lengua). Día 2

Forma	Bloque	Ítems por forma
3	BC3	22
	B05	
	BC4	
	B06	
4	BC3	22
	B07	
	BC4	
	B08	

De las 4 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). De este modo, de los 68 ítems que tiene la prueba,

un estudiante respondió 44.

Por otro lado, para los estudiantes cuya competencia lectora es evaluada solo en castellano como segunda lengua, se ensamblaron 2 formas que tuvieron el siguiente diseño:

Tabla 1.9 Matriz de bloques de la prueba de Lectura (en castellano como segunda lengua). Día 2

Forma	Sesión	Bloque	Ítems por forma
1	Primera	BC1	22
		B01	
		B02	
	Segunda	BC2	22
		B05	
		B06	
2	Primera	BC1	22
		B03	
		B04	
	Segunda	BC2	22
		B07	
		B08	

De las 2 formas ensambladas, cada estudiante evaluado se enfrentó a una de ellas. De este modo, de los 68 ítems en total que tiene la prueba, un estudiante respondió 44 ítems.

Diseño de bloques de las pruebas de 2.º grado de secundaria

La prueba de Lectura de segundo grado de secundaria elaborada para la ECE 2018 cuenta con un total de 14 textos de diferente tipo, género, formato y complejidad. A partir de este conjunto de textos, se elaboraron 86 ítems. Dichos textos e ítems se distribuyeron en 2 bloques comunes a todas las formas y 10 bloques no comunes. Las diferentes combinaciones de los bloques permitieron ensamblar 10 formas que reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

**Tabla 1.10 Matriz de bloques de la prueba de Lectura de 2.º grado de secundaria.
Día 1**

Forma	Bloque	Ítems por forma
1	BC1	25
	B01	
	B02	
2	BC1	25
	B02	
	B03	
3	BC1	25
	B03	
	B04	
4	BC1	25
	B04	
	B05	
5	BC1	25
	B05	
	B01	

**Tabla 1.11 Matriz de bloques de la prueba de Lectura de 2.º grado de secundaria.
Día 2**

Forma	Bloque	Ítems por forma
6	BC2	25
	B06	
	B07	
7	BC2	25
	B07	
	B08	
8	BC2	25
	B08	
	B09	
9	BC2	25
	B09	
	B10	
10	BC2	25
	B10	
	B06	

De las 10 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 86 ítems que tiene la prueba, cada estudiante respondió 50.

La prueba de Matemática elaborada para la ECE 2018 cuenta con un total de 90 ítems, los cuales están distribuidos en 2 bloques comunes a todos los cuadernillos y 10 bloques no comunes. Las diferentes combinaciones de los bloques permitieron ensamblar 10 formas que reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

Tabla 1.12 Matriz de bloques de la prueba de Matemática de 2.º grado de secundaria. Día 1

Forma	Bloque	Ítems por forma
1	B01	25
	BC1	
2	B02	25
	BC1	
3	B03	25
	BC1	
4	B04	25
	BC1	
5	B05	25
	BC1	

Tabla 1.13 Matriz de bloques de la prueba de Matemática de 2.º grado de secundaria. Día 2

Forma	Bloque	Ítems por forma
6	B06	25
	BC2	
7	B07	25
	BC2	
8	B08	25
	BC2	
9	B09	25
	BC2	
10	B10	25
	BC2	

De las 10 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 90 ítems que tiene la prueba, cada estudiante respondió 50.

La prueba de Ciencias Sociales elaborada para la ECE 2018 cuenta con un total de 87 ítems, 84 ítems de opción múltiple distribuidos en 6 bloques de 14 ítems cada uno, y 3 ítems de respuesta construida (B07, B08, B09) que forman un bloque independiente cada uno. Las diferentes combinaciones de los bloques permitieron ensamblar 6 formas que reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

Tabla 1.14 Matriz de bloques de la prueba de Ciencias Sociales de 2.º grado de secundaria. Día 3

Forma	Bloque	Ítems por forma
1	B01	29
	B07	
	B02	
2	B02	29
	B08	
	B03	
3	B03	29
	B07	
	B04	
4	B04	29
	B08	
	B05	
5	B05	29
	B09	
	B06	
6	B06	29
	B09	
	B01	

De las 6 formas ensambladas, cada estudiante evaluado se enfrentó a una de ellas. En consecuencia, de los 87 ítems que tiene la prueba, cada estudiante respondió 29.

La prueba de Ciencia y Tecnología elaborada para la ECE 2018 cuenta con un total de 87 ítems, 84 ítems de opción múltiple distribuidos en 6 bloques de 14 ítems cada uno, y 3 ítems de respuesta construida (B07, B08, B09) que forman un bloque independiente cada uno. Las diferentes combinaciones de los bloques permitieron

ensamblar 6 formas que reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

Tabla 1.15 Matriz de bloques de la prueba de Ciencia y Tecnología de 2.º grado de secundaria. Día 3

Forma	Bloque	Ítems por forma
1	B01	29
	B07	
	B02	
2	B03	29
	B08	
	B04	
3	B05	29
	B09	
	B06	
4	B02	29
	B08	
	B03	
5	B04	29
	B09	
	B05	
6	B06	29
	B07	
	B01	

De las 6 formas ensambladas, cada estudiante evaluado se enfrentó a una de ellas. En consecuencia, de los 87 ítems que tiene la prueba, cada estudiante respondió 29.

1.3.2 Diseño de bloques de las evaluaciones muestrales

En las siguientes tablas, se presenta el diseño de bloques de las pruebas aplicadas a nivel muestral. Las tablas de especificaciones de estos instrumentos pueden ser consultadas en los anexos.

Diseño de bloques de las pruebas de 6.º grado de primaria

La prueba de Ciudadanía elaborada para la EM 2018 cuenta en total con 120 ítems distribuidos en 8 bloques de 15 ítems, 13 ítems de opción múltiple y 2 de respuesta construida. Las diferentes combinaciones de los bloques permitieron ensamblar 8 formas que reflejaban los indicadores de la tabla de especificaciones y presentaban una dificultad similar entre sí.

**Tabla 1.16 Matriz de bloques de la prueba de Ciudadanía de 6.º grado de primaria.
Día 1**

Forma	Bloque	Ítems por forma
1	B01	30
	B02	
2	B03	30
	B04	
3	B05	30
	B06	
4	B07	30
	B08	
5	B02	30
	B03	
6	B04	30
	B05	
7	B06	30
	B07	
8	B08	30
	B01	

De las 8 formas ensambladas, cada estudiante evaluado se enfrentó a una de ellas. En consecuencia, de los 120 ítems que tiene la prueba, cada estudiante respondió 30.

Población y muestra

Capítulo 2

2.1 Evaluaciones censales

2.1.1 Población objetivo

En las evaluaciones censales 2018, existieron tres poblaciones objetivos: los estudiantes que asistieron a clases en el cuarto grado de primaria en una IE de Educación Básica Regular (EBR); los que asistieron a clases en el cuarto grado de primaria en una IE que desarrolla el Programa de Educación Intercultural Bilingüe (EIB); y los que asistieron a clases en el segundo grado de secundaria de una IE de Educación Básica Regular (EBR).

Se evaluó a los estudiantes de cuarto grado de primaria, porque en este grado se concluye el ciclo IV de la EBR, mientras que, en segundo grado de secundaria, se concluye el ciclo VI de la EBR. Evaluar estos grados brinda la oportunidad de planificar intervenciones en el sistema para mejorar los grados y ciclos posteriores, en los cuales el conocimiento se especializa y se hace más complejo.

Atendiendo el hecho de que no todos los estudiantes manejan el castellano como lengua materna, se evaluó a los estudiantes de cuarto grado de primaria de las IE EIB, por ser este el grado en que finaliza el ciclo IV de la EBR.

Se excluyó a los estudiantes que asisten a una IE que atiende a menos de cinco estudiantes matriculados² en alguno de los grados evaluados. En cuarto grado de primaria, esta exclusión fue de alrededor del 5 % de los estudiantes a nivel nacional y, en relación con la cantidad de IE, fue de alrededor del 40 % a nivel nacional. En cambio, en segundo grado de secundaria, la exclusión fue del 0,5 % de los estudiantes a nivel nacional. En cuarto grado de primaria EIB, esta exclusión fue de alrededor del 16 % a nivel de estudiantes, mientras que a nivel de IE fue de 49 %. Las razones de esta exclusión se encuentran detalladas en el *Marco de trabajo de la Evaluación Censal de Estudiantes (ECE)* (Ministerio de Educación, 2009).

En el 2018, la ECE tenía programado evaluar 20 584 IE en cuarto grado de primaria, 1 857 IE en cuarto grado de primaria EIB y 13 494 IE en segundo grado de secundaria. A nivel de estudiantes, se tenía programado evaluar, aproximadamente, 550 609 en cuarto grado de primaria, 20 484 en cuarto grado de primaria EIB y 558 674 en segundo grado de secundaria.

²La cantidad de estudiantes matriculados tiene como fuente el dato registrado en el Sistema de Información de Apoyo a la Gestión de la Institución Educativa (Siagie). En caso de no contar con esta información, el dato proviene del Censo Educativo o de la aplicación de la ECE del año anterior.

2.1.2 Marco poblacional

El marco poblacional ha sido elaborado con información proveniente de tres fuentes diferentes: el Sistema de Información de Apoyo a la Gestión de la Institución Educativa (Siagie)³, el Padrón de Instituciones Educativas de la Unidad de Estadística Educativa (UEE) y el Censo Escolar 2017.

Anualmente, la Unidad de Estadística Educativa (UEE) del Minedu elabora y difunde el padrón de instituciones educativas con fines estadísticos, el cual es actualizado de manera constante. La información sobre las IE tiene como fuente las operaciones estadísticas que aplican a nivel nacional y los procedimientos asociados a estas operaciones. Es importante precisar que la información de la cantidad de estudiantes y secciones que proviene del censo escolar es declarada por el director de la IE. En el Censo Escolar 2017, existen algunas IE que no reportan datos. Por ello, el 1,5 % de la información proveniente de esta fuente fue imputada.

En cuanto a la elaboración del listado de IE a evaluar en la ECE 2018, el procedimiento aplicado fue el siguiente:

1. Se utilizó el Padrón de Instituciones Educativas para determinar el listado de instituciones educativas activas⁴.
2. Se identificaron las IE del programa de Educación Intercultural Bilingüe (EIB) en primaria según lo declarado por los directores. Para ello, se utilizó la definición operacional construida para efectos de la ECE (Ministerio de Educación, 2009), la cual sostiene que una IE es EIB siempre que enseñe a leer y a escribir a sus estudiantes en su lengua originaria y en castellano como segunda lengua.
3. Se asignó el grado a evaluar. Las IE EIB son evaluadas en cuarto grado de primaria en Lectura en castellano como segunda lengua y en Lectura en lengua originaria. Las demás IE son evaluadas en cuarto grado de primaria (Lectura y Matemática) o en segundo grado de secundaria (Lectura, Matemática, Ciencias Sociales, y Ciencia y Tecnología).
4. Se calculó la cantidad de secciones y estudiantes programados. Esta información corresponde al valor registrado en Siagie en el caso de que las nóminas de matrículas se encuentren aprobadas al momento de la construcción del marco poblacional. Si una IE no contaba con las nóminas de matrícula aprobadas, se tomó el valor mayor entre la información del Censo Educativo 2017, Siagie 2018 y la cantidad de estudiantes evaluados en la ECE 2016.
5. Se seleccionaron las IE que tenían cinco o más estudiantes asignados en el punto 4.

³Específicamente, se consideró la base de datos correspondiente al 2018.

⁴El Padrón de Instituciones Educativas y el Censo Educativo se encuentran disponibles en la web Escale de la Unidad de Estadística Educativa: <http://escale.minedu.gob.pe/uee/>

Culminado este procedimiento, la distribución de IE y estudiantes para la ECE 2018 quedó definida tal y como se muestra en la Tabla 2.1

Tabla 2.1 Cantidad de IE, secciones y estudiantes programados en la ECE 2018

Tipo de evaluación	IE	Secciones	Estudiantes
4.º grado de primaria	20 687	29 906	550 856
4.º grado de primaria EIB	1 879	1 964	20 527
2.º grado de secundaria	13 500	24 476	558 692

2.1.3 Cobertura

Cuarto grado de primaria

La cobertura a nivel de estudiantes se muestra en la Tabla 2.2. El porcentaje de estudiantes evaluados respecto de los programados a nivel nacional es de 97,1 %⁵. Dicha cobertura se calculó considerando la cantidad de estudiantes que llegan a la evaluación final del año escolar en cada IE según la información del Siagie para el 2018.

⁵Esta cobertura no considera aquellas instituciones educativas que se encontraban cerradas o sin estudiantes en el grado a evaluar durante la aplicación.

Tabla 2.2 *Porcentaje de estudiantes evaluados por estratos a nivel nacional y en cada DRE*

	Estatad	No estatal	Urbana	Rural	Polidocente completo	Unidocente / Multigrado
Nacional	97,4	96,2	97,3	95,8	97,5	94,6
Amazonas	97,0	96,2	97,3	96,4	98,1	94,8
Áncash	97,7	96,9	97,7	97,0	97,8	95,7
Apurímac	98,9	98,2	98,8	98,9	98,9	98,7
Arequipa	98,3	98,1	98,2	96,4	98,3	96,5
Ayacucho	99,0	97,9	98,9	98,6	99,1	97,6
Cajamarca	97,4	97,1	97,5	97,3	97,5	97,1
Callao	97,3	94,9	96,4	-	96,6	90,9
Cusco	98,9	97,5	98,9	97,4	99,0	96,1
Huancavelica	98,7	98,8	99,0	98,1	98,9	97,9
Huánuco	97,4	97,8	98,1	96,0	98,0	95,9
Ica	98,5	96,4	97,9	98,7	98,1	94,1
Junín	97,4	97,3	97,6	96,1	97,8	95,1
La Libertad	96,9	96,1	96,7	96,7	96,9	95,6
Lambayeque	97,6	95,9	97,3	96,3	97,4	95,5
Lima Metropolitana	97,7	95,5	96,6	98,0	96,9	89,5
Lima Provincias	97,6	97,5	97,6	97,5	97,6	97,1
Loreto	92,4	93,3	94,0	88,1	94,2	87,6
Madre de Dios	97,7	89,5	96,9	93,7	97,0	92,0
Moquegua	99,4	98,5	99,3	95,5	99,4	94,8
Pasco	96,5	98,3	98,4	90,7	98,4	90,7
Piura	98,2	97,0	98,0	97,9	98,2	97,4
Puno	99,0	99,1	99,0	99,0	99,0	98,8
San Martín	97,3	98,4	97,8	95,7	97,7	95,9
Tacna	99,4	99,1	99,3	100	99,3	99,3
Tumbes	97,5	94,7	97,0	99,4	97,3	93,8
Ucayali	94,1	94,3	95,7	84,4	95,7	85,0

Cuarto grado de primaria EIB

Para el caso de cuarto grado de primaria EIB, el porcentaje de IE evaluadas fue de 98,9 %, mientras que el de estudiantes fue de 92,1 %.

Segundo grado de secundaria

La cobertura de estudiantes se muestra en la Tabla 2.3. El porcentaje de estudiantes evaluados respecto de los programados a nivel nacional es de 97,4 %. Al igual que en cuarto grado de primaria, dicha cobertura se calculó considerando la cantidad

de estudiantes que llegan a la evaluación final del año escolar en cada IE, según la información de Siagie para el 2018.

Tabla 2.3 *Porcentaje de estudiantes evaluados por estratos a nivel nacional y en cada DRE*

	Estatal	No estatal	Urbana	Rural	Polidocente completo
Nacional	97,7	96,5	97,8	93,7	97,4
Amazonas	94,7	88,4	97,5	88,9	94,6
Áncash	98,3	96,6	98,2	97,2	98,0
Apurímac	98,2	97,7	98,5	97,2	98,2
Arequipa	99,3	98,4	99,0	98,8	99,0
Ayacucho	98,7	97,0	98,6	98,4	98,6
Cajamarca	97,8	93,7	98,5	95,1	97,4
Callao	98,9	94,8	97,4	-	97,4
Cusco	98,9	97,2	98,9	97,9	98,7
Huancavelica	97,0	98,3	97,7	95,8	97,0
Huánuco	96,7	97,7	97,9	93,0	96,8
Ica	98,6	97,8	98,4	100	98,4
Junín	97,6	94,8	97,5	93,2	97,1
La Libertad	97,5	96,2	97,8	93,5	97,2
Lambayeque	97,7	96,2	97,4	94,7	97,3
Lima Metropolitana	98,6	96,2	97,5	86,9	97,5
Lima Provincias	98,1	97,8	98,1	96,8	98,1
Loreto	91,2	95,5	94,4	80,6	91,5
Madre de Dios	96,6	94,0	96,3	97,3	96,4
Moquegua	99,4	99,8	99,5	97,6	99,4
Pasco	96,3	94,0	98,3	90,0	96,2
Piura	98,8	97,7	98,7	97,9	98,6
Puno	99,0	98,1	98,9	98,7	98,8
San Martín	96,4	98,6	97,7	89,7	96,4
Tacna	99,6	98,4	99,3	99,4	99,3
Tumbes	98,8	97,1	98,5	100	98,5
Ucayali	93,0	95,0	95,7	80,1	93,1

2.1.4 Ajuste por no respuesta y cálculo de resultados

Este ajuste es realizado a nivel de estrato, el cual es construido a partir de la combinación entre DRE, gestión, área y característica. Para el cálculo de resultados en cuarto grado de primaria y en segundo grado de secundaria, es necesario multiplicar los datos de cada estudiante por el ajuste por no respuesta. Los resultados se reportan según los siguientes niveles:

- Nacional
- Por sexo de los estudiantes (femenino, masculino)
- Por gestión de la IE (estatal, no estatal)
- Por área geográfica de la IE (urbana, rural)
- Por característica de la IE (polidocente completo, unidocente-multigrado)
- Por Dirección Regional de Educación (DRE)
- Por Unidad de Gestión Educativa Local (UGEL)
- Por distrito

El cálculo del ajuste por no respuesta se realizó de la siguiente manera. Primero, se realizó el ajuste por estudiantes no evaluados a nivel de IE.

$$aj_IE_i = \frac{p_IE_i}{e_IE_i},$$

donde:

e_IE_i = Cantidad de alumnos evaluados en la i-ésima IE

p_IE_i = Cantidad de alumnos programados en la i-ésima IE

aj_IE_i = Ajuste por no respuesta en la i-ésima IE

Luego, se calculó el ajuste por estudiantes no evaluados a nivel de estrato.

$$aj_E_j = \frac{p_E_j}{e_E_j},$$

donde:

e_E_j = Cantidad de alumnos programados en el j-ésimo estrato

p_E_j = Cantidad de alumnos evaluados ajustada en el j-ésimo estrato, la cual es calculada ponderando la cantidad de alumnos evaluados por el ajuste de no respuesta en la IE

aj_E_j = Ajuste por no respuesta en el j-ésimo estrato

Finalmente, se calculó el ajuste por no respuesta para cada estudiante.

$$aj_{NR} = aj_{IE_i} \times aj_{E_j}$$

2.2 Evaluaciones muestrales

Entre las evaluaciones de rendimiento realizadas el 2018, se llevaron a cabo tres operativos muestrales. Específicamente, se evaluaron Lectura y Matemática en segundo grado de primaria (detallado en Ministerio de Educación, [2019]), Ciudadanía en sexto grado de primaria y Escritura en segundo grado de secundaria.

En estos casos, la construcción del marco muestral se llevó a cabo siguiendo el mismo procedimiento que en las evaluaciones censales. Por lo tanto, se excluyeron aquellas IE en las que había menos de cinco estudiantes matriculados en el grado a ser evaluado.

2.2.1 Diseño muestral

El muestreo empleado fue del tipo probabilístico, estratificado, por conglomerados y bietápico, donde la unidad primaria de muestreo corresponde a las instituciones educativas y la unidad secundaria corresponde a las secciones completas dentro de cada IE. De esta manera, cada IE es considerada un conglomerado, pues agrupa a un conjunto de secciones (que también son conglomerados), que a su vez agrupan a un conjunto de estudiantes. El muestreo de conglomerados surge como una necesidad práctica, económica y de eficiencia administrativa. Su mayor ventaja es que no es necesario contar con un marco muestral de cada elemento por analizar, solo es necesario tener el marco muestral a nivel de conglomerados (IE y secciones) para realizar la selección de los mismos (Lehtonen y Pahkinen, 2004). Luego, en cada conglomerado seleccionado, se puede obtener el listado de todas las secciones y seleccionar una muestra de estas.

Volviendo a la característica bietápica del muestreo realizado en la primera etapa, la selección de escuelas fue proporcional al tamaño, por lo cual las escuelas con mayor cantidad de estudiantes tuvieron una mayor probabilidad de ser seleccionadas. En la segunda etapa, se realizó una selección aleatoria simple de hasta dos secciones en segundo y sexto grado de primaria, y una sección en segundo grado de secundaria.

Además, la muestra compleja trabajada implicó el uso de estratos. Específicamente, en las evaluaciones muestrales de 2018 se realizó una estratificación porque, en un contexto de tanta diversidad como el peruano, la variable de interés (rendimiento) toma diferentes valores en diferentes subpoblaciones, como aquellas diferenciadas por el área geográfica, el tipo de gestión de la IE o su característica⁶. La combinación de estas tres subpoblaciones permite generar diferentes estratos

⁶Solo aplicable para primaria.

mutuamente excluyentes y colectivamente exhaustivos donde cada elemento pertenece únicamente a un estrato.

En el caso de primaria, se conformaron los siguientes estratos:

- Estrato 1: Urbano - Estatal - Polidocente completo
- Estrato 2: Urbano - Estatal - Unidocente / Multigrado
- Estrato 3: Rural - Estatal - Polidocente completo
- Estrato 4: Rural - Estatal - Unidocente / Multigrado
- Estrato 5: No estatal (las IE no estatales, casi en su totalidad, son urbanas y polidocentes)

En el caso de secundaria, se conformaron los siguientes estratos:

- Estrato 1: Urbano - Estatal
- Estrato 2: Rural - Estatal
- Estrato 3: No estatal

Una vez estratificado el marco muestral y determinado el tamaño de muestra, se empleó la asignación de Neyman para determinar la cantidad de muestra en cada uno de los estratos, la cual es directamente proporcional a la variabilidad y el tamaño del mismo.

En cuanto a los niveles de inferencia de la EM 2018, se pueden reportar resultados representativos en los siguientes niveles: a) nivel nacional, b) por sexo de los estudiantes, c) por gestión de las IE (estatales / no estatales), d) por ubicación geográfica de las IE (urbanas / rurales), y e) por característica (polidocente completo / unidocente-multigrado). Esto último solo es aplicable en los grados de primaria, dado que todas las IE en secundaria son polidocentes completas.

2.2.2 Cobertura

En las siguientes tablas, se presenta el tamaño final de la muestra desagregada según gestión, área y característica para cada uno de los grados evaluados, así como la cantidad efectivamente evaluada y la cobertura alcanzada.

Tabla 2.4 Estudiantes programados y evaluados en 6.º grado de primaria

6.º grado de primaria	Evaluados		Programados		Cobertura (%)	
	IE	alumnos	IE	alumnos	IE	alumnos
Nacional	359	10 524	360	10 801	99,7	97,4
Estatad	289	8 227	290	8 431	99,7	97,6
No estatal	70	2 297	70	2 370	100	96,9
Urbana	226	9 271	226	9 479	100	97,8
Rural	133	1 253	134	1 322	99,3	94,8
Polidocente completo	222	9 406	222	9 629	100	97,7
Unidocente / Multigrado	137	1 118	138	1 172	99,3	95,4

Tabla 2.5 Estudiantes programados y evaluados en 2.º grado de secundaria

2.º grado de secundaria	Evaluados		Programados		Cobertura (%)	
	IE	alumnos	IE	alumnos	IE	alumnos
Nacional	251	5 968	258	6 399	97,3	93,3
Estatad	191	4 538	194	4 864	98,5	93,3
No estatal	60	1 430	64	1 535	93,8	93,2
Urbana	211	5 322	217	5 688	97,2	93,6
Rural	40	646	41	711	97,6	90,9

2.2.3 Pesos muestrales

Los pesos ayudan a corregir la distribución de la muestra en los estratos y a expandir la información muestral a la población. Para su cálculo (peso final), se obtuvo una serie de pesos intermedios, cuya combinación lleva a definir el peso final por estudiante.

Peso por IE. En la primera etapa, el peso es el inverso de la probabilidad de selección de un conglomerado (IE) en el interior de cada estrato. En el caso del muestreo proporcional al tamaño, se utilizan probabilidades conjuntas de selección de la IE. La inversa de esta probabilidad vendría a ser el peso de la IE.

$$pIE_i = \frac{1}{p_{ij}}$$

donde:

p_{ij} = Probabilidad de selección conjunta de la i -ésima IE en el j -ésimo estrato

pIE_i = Peso de la i -ésima IE en el j -ésimo estrato

Peso por sección. Es el inverso de la probabilidad de selección de las secciones. Como ya se mencionó antes, la selección de secciones fue aleatoria simple.

$$psec_i = \frac{secIE_i}{seceva_i},$$

donde:

$psec_i$ = Peso por sección en la i-ésima IE

$secIE_i$ = Total de secciones de la i-ésima IE

$seceva_i$ = Secciones evaluadas en la i-ésima IE

Ajuste por estudiantes no evaluados en la sección. Este ajuste se realiza de manera separada para cada área evaluada.

$$a_{ki} = \frac{t_{ki}}{s_{ki}},$$

donde:

a_{ki} = Ajuste por estudiantes no evaluados en la k-ésima sección y en la i-ésima IE

t_{ki} = Total de estudiantes que asisten a la k-ésima sección en la i-ésima IE

s_{ki} = Total de estudiantes evaluados en la k-ésima sección y en la i-ésima IE

Peso final. Se obtiene un peso para Lectura y otro para Matemática. Para cada caso, este se obtiene a partir de la siguiente fórmula:

$$pf_{kij} = pIE_{ij} \times psec_i \times a_{ki},$$

donde:

pf_{kij} = Peso final en la k-ésima sección, en la i-ésima IE y el j-ésimo estrato

pIE_{ij} = Peso de la i-ésima IE en el j-ésimo estrato

$psec_i$ = Peso por sección en la i-ésima IE

a_{ki} = Ajuste por estudiantes no evaluados en la k-ésima sección en la i-ésima IE

Operativo de campo

Capítulo 3

La organización, la implementación y los procedimientos del operativo de campo de las evaluaciones se rigen bajo tres principios:

- Mantener la confidencialidad de las pruebas
- Asegurar la estandarización de los procedimientos de evaluación
- Promover la probidad en las acciones de todos los actores durante todo el proceso

Las estrategias propuestas para alcanzar estos principios se plasmaron en las especificaciones técnicas de los términos de referencia del servicio de aplicación de las evaluaciones. En el caso de la ECE, desde el 2007, el Instituto Nacional de Estadística e Informática (INEI) es el operador logístico responsable de la ejecución de la aplicación de esta evaluación en todo el país. En cambio, la Evaluación Muestral (EM) es ejecutada en función de la estrategia específica del operativo cada año; para el 2018, se optó por un operador logístico privado.

Los operadores logísticos tienen bajo su responsabilidad garantizar lo siguiente:

- La correcta impresión, modulado y embalaje de los instrumentos de evaluación, así como de los materiales de aplicación y capacitación.
- La distribución, y la entrega oportuna y segura de los materiales de capacitación e instrumentos de aplicación a cada una de las sedes a nivel nacional, así como el retorno de todos los instrumentos, aplicados y no aplicados, a la empresa responsable de la captura de datos
- La seguridad y confidencialidad de los instrumentos requeridos en todas las fases del operativo
- La adecuada selección de la red administrativa (RA) según el perfil establecido
- La capacitación de la RA en los procedimientos estandarizados de aplicación
- El contacto previo con las instituciones educativas (IE), y la actualización de la información de las IE y los estudiantes
- La correcta aplicación de los instrumentos de evaluación en las IE

Por su parte, el Minedu puso en práctica una estrategia de monitoreo y supervisión de todas las fases del operativo de campo. Para ello, contó con un equipo de monitores Minedu: 102 para la ECE y 20 para la EM. Ellos supervisaron in situ una o más sedes del operativo. El equipo de monitoreo tuvo como principales funciones asegurar que los procesos de capacitación cumplieran con el objetivo de transmitir

de manera estandarizada los procedimientos de aplicación y velar por que el operador logístico llevara a cabo las acciones necesarias para asegurar la confidencialidad de los instrumentos de evaluación, así como apoyar en la solución de los incidentes que se presentaran en las sedes. Asimismo, este equipo fue el principal interlocutor con las instancias de gestión educativa descentralizada.

En el caso de la ECE, desde el 2018 se implementó un nuevo modelo operativo con el propósito de optimizar los procesos y garantizar el cumplimiento de los principios de la evaluación (estandarización, confidencialidad y probidad). Este proceso consistió en que ninguna caja⁷ con instrumentos de aplicación se abriera antes del primer día de evaluación correspondiente a las fechas principales.

El modelo consideró dos grandes momentos de aplicación:

- La aplicación en fechas principales, en la que el aplicador retorna cada día de aplicación al local de la sede
- La aplicación en fechas posteriores, en la que el aplicador pernocta en una zona cercana a la IE, debido a la accesibilidad, tiempo de traslado, frecuencia de movilidad y seguridad de la ruta

3.1 Organización territorial

Para gestionar territorialmente las evaluaciones, el país se divide en jurisdicciones regionales. La conformación de estas jurisdicciones regionales no se ciñe estrictamente a los límites del departamento donde se ubican, sino a la accesibilidad de los lugares de aplicación.

3.1.1 Divisiones jurisdiccionales para la ECE

El país se dividió en 41 jurisdicciones regionales de aplicación, en las cuales se consideraron 280 sedes provinciales o distritales.

En la tabla 3.1, se puede observar el número de IE programadas en cada jurisdicción regional de aplicación.

Tabla 3.1 Jurisdicciones regionales de aplicación y distribución de IE en la ECE 2018

Jurisdicción regional	2S	4EIB	4P
Amazonas - Bagua	427	243	667
Amazonas - Chachapoyas	109		191
Áncash - Chimbote	230	19	347
Áncash - Huaraz	397	143	476

⁷Cada caja contiene los materiales necesarios para realizar la aplicación en una sección de una IE

Jurisdicción regional	2S	4EIB	4P
Apurímac	314	151	327
Arequipa - Camaná	135	8	182
Arequipa - La Joya	63		84
Arequipa - Norte	245		326
Arequipa - Sur	192		309
Ayacucho	397	131	454
Cajamarca	810		1 474
Callao	287		457
Cusco	616	255	719
Huancavelica	337	101	338
Huánuco	461	49	749
Ica	368	10	589
Junín	667	133	1 012
La Libertad - Noroeste	340		519
La Libertad - Sureste	466		886
Lambayeque	435	59	728
Lima Metropolitana 1	446		709
Lima Metropolitana 2	450		693
Lima Metropolitana 3	317		406
Lima Metropolitana 4	414		634
Lima Metropolitana 5	360		539
Lima Metropolitana 6	444		671
Lima Metropolitana 7	276		368
Lima - Cañete	135	1	198
Lima - Huacho	247		334
Loreto	381	68	780
Madre de Dios	70	6	91
Moquegua	66		78
Pasco	187	16	251
Piura	770		1 555
Puno - Juliaca	318	46	442
Puno - Puno	245	39	288
San Martín - Moyobamba	187	17	395
San Martín - Tarapoto	346	209	650
Tacna	117		146
Tumbes	83		130
Ucayali	345	175	495
Total	13 500	1 879	20 687

Nota. "2S" hace referencia a segundo grado de secundaria; "4P EIB", a cuarto grado de primaria de Educación Intercultural Bilingüe (EIB); y "4P", a cuarto grado de primaria.

3.1.2 Divisiones jurisdiccionales para la EM

Para este operativo, el país se organizó territorialmente en 20 sedes macro regionales. Como se observa en la tabla 3.2, las sedes tienen bajo su administración la aplicación de IE pertenecientes a más de un departamento y provincia.

Tabla 3.2 Jurisdicciones regionales de aplicación y distribución de IE en la EM 2018

Sede macroregional	Departamentos que conforman la sede	EM 6P
Arequipa	Arequipa, Tacna, Moquegua	18
Cajamarca	Amazonas, Cajamarca, La Libertad	31
Callao	Callao, Lima Provincias	10
Chachapoyas	Amazonas, Cajamarca	17
Chanchamayo	Junín	19
Chiclayo	Lambayeque	10
Coronel Portillo	Loreto, Ucayali	14
Cusco	Apurímac, Cusco, Madre de Dios, Arequipa	15
Huamanga	Apurímac, Ayacucho, Cusco	15
Huancavelica	Huancavelica, Junín	7
Huánuco	Huánuco, Pasco	20
Huaraz	Huaraz, Huánuco	17
Ica	Ayacucho, Huancavelica , Ica, Lima Provincias	9
Juliaca	Puno	10
Lima 1	Junín, Lima	35
Lima 2	Lima	31
Maynas	Loreto	12
Piura	Piura, Tumbes	29
Tarapoto	Loreto, San Martín, Ucayali	22
Trujillo	La Libertad, Lambayeque	19
Total		360

Nota. "6P" hace referencia a sexto grado de primaria.

3.2 Conformación de la red administrativa de aplicación

El personal necesario para ejecutar el operativo de campo estuvo organizado en una red administrativa (RA).

Tabla 3.3 Personal del operativo en la ECE 2018

	Cargo	N.º de personas
Nivel I	Coordinador regional	41
	Coordinador provincial	280
Nivel II	Supervisor de procesos de aplicación	280
	Supervisor de almacén y sistemas	280
Nivel III	Asistentes de procesos de aplicación ECE (2S ECE Y 4P ECE)	2 410
	Asistentes de procesos de aplicación EIB L1 y L2	83
	Asistentes de procesos de aplicación EIB solo L2	141
Nivel IV	Aplicador de 4P	38 989
	Aplicadores de 4P EIB L1 y L2	2 511
	Aplicadores de 4P EIB solo L2	1 541
	Aplicador de 2S	44 279
	Aplicadores de Necesidades Educativas Especiales (NEE)	4 174

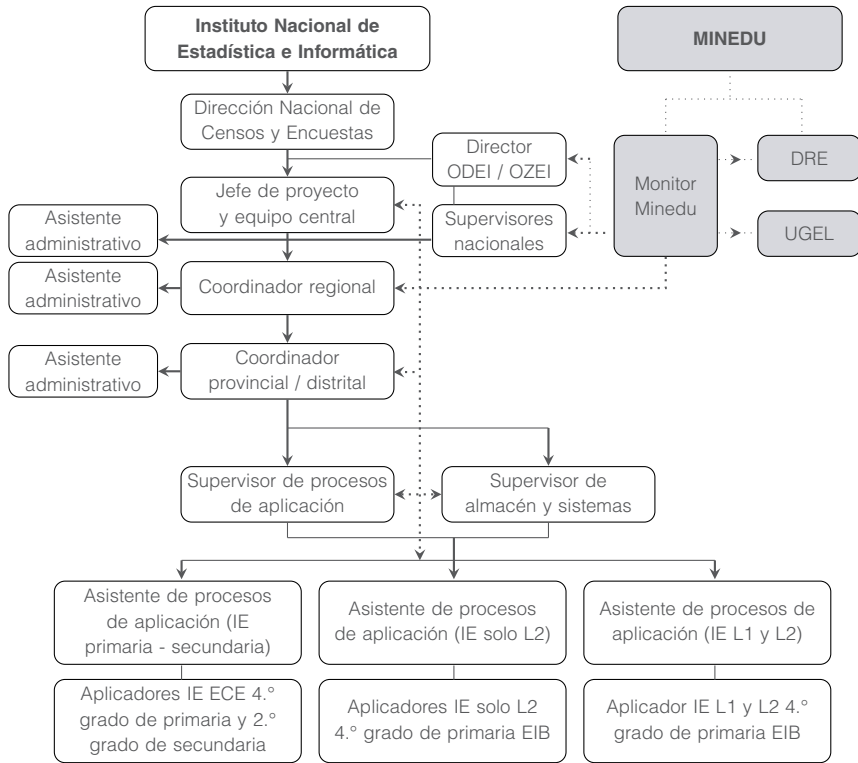
Nota. "2S" hace referencia a segundo grado de secundaria; "4P EIB", a cuarto grado de primaria de Educación Intercultural Bilingüe (EIB); y "4P", a cuarto grado de primaria. Asimismo, "L1 y L2" refiere a las escuelas EIB que enseñan en lengua originaria (L1) y en castellano como segunda lengua (L2), mientras que "solo L2" indica a las escuelas que solo enseñan en castellano como segunda lengua.

Tabla 3.4 Personal del operativo en la EM 2018

	Cargo	N.º de personas
Nivel I	Coordinadores regionales	20
	Supervisores de almacén y sistemas	20
Nivel II	Supervisores EM 6P	39
	Aplicadores	664

Nota. "6P" hace referencia a sexto grado de primaria.

Figura 3.1 Organigrama de la red administrativa ECE 2018



A continuación, se resumen las principales funciones de cada uno de los cargos.

- Los coordinadores regionales, provinciales y distritales son responsables de las siguientes actividades: organizar, gestionar y dirigir el operativo; velar por el cumplimiento de los procedimientos; y monitorear el desempeño de los miembros de la RA de su jurisdicción (región, provincia o distrito). Asimismo, son responsables de garantizar las condiciones materiales necesarias para el operativo y asegurar la implementación de medidas de seguridad en el almacén de instrumentos.
- Los supervisores de procesos de aplicación son responsables de organizar y dirigir los procesos de convocatoria, selección y capacitación de aplicadores; contactar previamente a las IE; despachar a los aplicadores; y del proceso de aplicación en general. También son los encargados de velar por el cumplimiento de los procedimientos y del monitoreo del desempeño de los asistentes de los procesos de aplicación.
- Los supervisores de sistemas y almacén son responsables de gestionar la recepción, el almacenamiento, el inventario, el registro de incidencias, el

despacho del material de aplicación y el uso de los sistemas de información implementados para estas actividades.

- Los asistentes de procesos de aplicación o supervisores de la EM son responsables de ejecutar la selección y capacitación de los aplicadores, velar por el cumplimiento de los procedimientos, supervisar la aplicación y realizar el contacto previo con las IE asignadas.
- Los aplicadores son responsables de la adecuada aplicación de los instrumentos en las IE evaluadas.

Para cada uno de los niveles de la RA, el Minedu entregó manuales de aplicación y manuales de funciones. En ese sentido, se elaboraron manuales para el personal de la RA que atiende las particularidades de cada operativo y de los grados evaluados.

ECE 2018

- Manual del coordinador regional y provincial
- Manual del supervisor de procesos de aplicación
- Manual del supervisor de almacén y de sistemas
- Manual del usuario del sistema
- Manual del asistente de procesos de aplicación ECE Educación Intercultural Bilingüe (EIB) solo en castellano como segunda lengua (L2)
- Manual del asistente de procesos de aplicación ECE EIB en lengua originaria (L1) y en castellano como segunda lengua (L2)
- Manual del asistente de procesos de aplicación ECE primaria y secundaria
- Manual del aplicador 4.º grado de primaria EIB solo L2
- Manual del aplicador 4.º grado de primaria EIB L1 y L2 awajún
- Manual del aplicador 4.º grado de primaria EIB L1 Y L2 aimara
- Manual del aplicador 4.º grado de primaria EIB L1 Y L2 ashaninka
- Manual del aplicador 4.º grado de primaria EIB L1 Y L2 quechua cusco collao
- Manual del aplicador 4.º grado de primaria EIB L1 Y L2 quechua chanca
- Manual del aplicador 4.º grado de primaria EIB L1 Y L2 shipibo-konibo
- Manual del aplicador 4.º grado de primaria
- Manual del aplicador 2.º grado de secundaria
- Manual del aplicador NEE Discapacidad visual (DV) - baja visión
- Manual del aplicador NEE Discapacidad visual (DV) - ceguera
- Manual del aplicador NEE Discapacidad auditiva (DA)

- Manual del aplicador NEE Discapacidad motora (DM)
- Manual del aplicador NEE Discapacidad de trastorno de espectro autista (D-TEA)
- Manual del aplicador NEE Discapacidad intelectual (DI) - sabe leer
- Manual del aplicador NEE Discapacidad intelectual (DI) - no sabe leer

EM 2018

- Manual del coordinador regional
- Manual del supervisor de almacén y de sistemas
- Manual del supervisor
- Manual del aplicador de 6.º grado de primaria

3.3 Proceso de selección y capacitación del personal de la RA

El proceso de selección del personal estuvo a cargo de los operadores logísticos y constó de tres etapas cancelatorias:

1. Convocatoria y revisión de hojas de vida
2. Preselección
3. Capacitación

En los términos de referencia, el Minedu estableció el perfil profesional para cada uno de los cargos. En ambos operativos, el perfil de los aplicadores fue muy similar; en la ECE fue el siguiente:

- Formación (no están colocadas en orden de prioridad):
 - Docente con experiencia en aula, sin carga docente o administrativa en el sector estatal, o
 - Egresado universitario o egresado técnico, excepto carreras vinculadas al cuidado personal u oficios, o
 - Estudiantes universitarios de carreras de humanidades de los tres últimos años, o
 - Estudiantes universitarios de otras carreras que no sean humanidades, de los tres últimos años y con experiencia demostrada en la aplicación de instrumentos de recojo de información, o
 - Estudiantes de institutos pedagógicos de los tres últimos años, o

- Estudiantes de carreras técnicas del último año con experiencia demostrada en la aplicación de instrumentos de recojo de información, excepto carreras vinculadas al cuidado personal u oficios.
- Experiencia:
 - Aplicador con buen desempeño en el operativo de la ECE 2015 o la ECE 2016, o
 - Personal sin experiencia en alguno de los operativos ECE 2015 o 2016 que, además, cumpla con los siguientes requisitos:
 - * Entre 19 y 60 años de edad (puede haber excepciones siempre que sean debidamente aprobadas por la UMC)
 - * Experiencia deseable en trabajo con niños o adolescentes en aula o en animaciones socioculturales, o aplicando pruebas estandarizadas de rendimiento escolar
 - * Habilidades para el buen manejo de grupos de niños y adolescentes
 - * Experiencia deseable en la aplicación de instrumentos estandarizados de evaluación educativa
 - * Buen nivel de lectura comprensiva
 - * Adecuada dicción
 - * Asistencia obligatoria a las sesiones de capacitación
 - * Disposición a tiempo completo durante las fechas de capacitación y aplicación (puede incluir viajes a lugares donde se los requiera)
- Para el caso de aplicadores asignados a IE EIB L1 y L2, buen dominio oral y lector de una de las siguientes lenguas originarias: aimara, awajún, ashaninka variante Perené de las zonas de Chanchamayo y Puerto Bermúdez, quechua chanca, quechua cusco collao o shipibo-konibo. El nivel de manejo de la lengua originaria fue definido a través de una entrevista.
- Para el caso de aplicadores de secciones con necesidades educativas especiales asociadas a discapacidad, se solicitó como mínimo (en orden de prioridad) uno de los siguientes:
 - Egresado de carrera de educación especial
 - Estudiante universitario de carreras de educación especial de los dos últimos años
 - Egresado de educación superior universitaria, o titulado técnico con especialidad o diplomado en educación inclusiva
 - Egresado universitario de las carreras de psicología, trabajo social o tecnología médica con mención en terapia de lenguaje

- Egresado de educación superior universitaria o titulado técnico con experiencia en la atención de estudiantes con necesidades educativas especiales asociadas a discapacidad

Además, a los postulantes se les requirió lo siguiente:

- Experiencia mínima de 3 meses en trabajo con niños o adolescentes
- Experiencia deseable en trabajo con niños o adolescentes con discapacidad
- Para atención de estudiantes que presenten discapacidad auditiva (según registro R-NEE), manejar lengua de señas peruana (LSP)
- Para atención de estudiantes con sordoceguera, manejar dactilología (deseable) o tener experiencia como mediador (3 meses)
- Para atención de estudiantes con autismo, experiencia en la atención de estudiantes con esta discapacidad y manejo deseable del sistema de comunicación por intercambio de tarjetas
- Asistir obligatoriamente a las sesiones de capacitación
- Disponibilidad para desplazarse a la IE que se les asigne

Con el objetivo de desarrollar capacidades en los equipos de trabajo y asegurar que reciban las mismas instrucciones de aplicación en todo el país, se preparó un diseño de capacitación que fue replicado en todas las aulas de capacitación a nivel nacional. En este, la metodología incluyó actividades de lectura individual y grupal, ejercicios prácticos para el uso de las fichas de recojo de información, actividades sobre manejo de procedimientos y documentos de aplicación, y actividades de simulación de los procesos de aplicación. También se detallaron los indicadores de evaluación en cada una de las actividades propuestas. Al terminar la capacitación, los candidatos rindieron una prueba de salida orientada a conocer el nivel de comprensión de los procedimientos centrales de la aplicación, así como la habilidad de los candidatos para el manejo de los documentos e instrumentos de aplicación.

La capacitación de aplicadores se realizó en cada una de las sedes establecidas y estuvo a cargo de los supervisores y asistentes de procesos de aplicación de la sede.

Para los procesos de la ECE, la capacitación tuvo una duración de dos días para los aplicadores de cuarto grado de primaria, segundo grado de secundaria, cuarto grado de primaria EIB en lengua originaria (L1) y castellano como segunda lengua (L2); y una duración de día y medio para los aplicadores de cuarto grado de primaria EIB en solo L2. En el caso de los procesos de la EM, la capacitación tuvo una duración de día y medio para los aplicadores de sexto grado de primaria. Además, los aplicadores que atendieron a estudiantes con discapacidad recibieron medio día adicional de

capacitación en sus propios manuales, según el tipo de discapacidad que iban a atender. Las aulas estuvieron conformadas por un máximo de 30 candidatos.

Durante la capacitación, se instruyó a los candidatos en la aplicación estandarizada de instrumentos. Para ello, se utilizó el manual del aplicador correspondiente al operativo y grado, así como los materiales elaborados para esta finalidad: fichas ópticas de práctica, papelógrafos, guiones de aplicación, entre otros.

3.4 Características de los instrumentos y documentos utilizados

Se utilizaron los siguientes instrumentos y documentos de aplicación:

ECE de cuarto grado de primaria

- Cuadernillos integrados de pruebas (un cuadernillo por día de evaluación)
- Ficha óptica de asistencia y respuestas (FOAR)
- Registro de préstamo
- Ficha de verificación
- Carta al director
- Cuestionario al director
- Carta al docente
- Compromiso de confidencialidad para el docente
- Cuestionario al docente de Comunicación
- Cuestionario para la familia
- Comunicado para la familia
- Ficha del aplicador líder
- Acta de conformidad del director y del aplicador

ECE de cuarto grado de primaria EIB (solo L2)

- Cuadernillos integrados de prueba
- Ficha óptica de asistencia y respuestas (FOAR)
- Registro de préstamo
- Ficha de verificación
- Carta al director
- Cuestionario para el docente
- Compromiso de confidencialidad para el docente
- Comunicado para la familia

- Ficha del aplicador líder
- Acta de conformidad del director y del aplicador

ECE de cuarto grado de primaria EIB (L1 y L2)

- Cuadernillos integrados de prueba (un cuadernillo por día de evaluación)
- Ficha óptica de asistencia y respuestas (FOAR)
- Registro de préstamo
- Ficha de verificación
- Carta al director
- Cuestionario para el docente
- Compromiso de confidencialidad para el docente
- Comunicado para la familia
- Ficha del aplicador líder
- Acta de conformidad del director y del aplicador

ECE de segundo grado de secundaria

- Cuadernillos integrados de pruebas (un cuadernillo por día de prueba)
- Cuestionario para el estudiante
- Fichas de respuestas para estudiantes
- Registro de estudiantes
- Registro de préstamo
- Ficha de verificación
- Carta al director
- Cuestionario a la institución educativa
- Carta al docente
- Cuestionario para el docente de Ciencia y Tecnología
- Compromiso de confidencialidad para el docente (por día de aplicación)
- Ficha del aplicador de sección
- Ficha del aplicador líder
- Acta de conformidad del director y del aplicador

EM de sexto grado de primaria

- Cuadernillos de pruebas
- Cuestionario al estudiante
- Registro de estudiantes
- Registro de préstamo
- Ficha de verificación
- Carta al director
- Compromiso de confidencialidad para el docente (por día de aplicación)
- Ficha del aplicador de sección
- Ficha del aplicador líder
- Acta de conformidad del director y del aplicador

Para facilitar y agilizar los procesos de aplicación, los cuadernillos de pruebas fueron personalizados con los nombres y apellidos de los estudiantes, y un código de barras único. Para ello, se utilizó la información provista por el Sistema de Información de Apoyo a la Gestión de las Instituciones Educativas (Siagie). El 95 % de los instrumentos enviados a las IE contó con información personalizada por estudiante.

La inclusión de códigos y otros elementos de identificación en todos los instrumentos de estas evaluaciones permitió realizar la trazabilidad de cada uno de ellos.

El operador logístico supervisó el avance y realizó el control de calidad de la impresión y modulado de los instrumentos (una caja para cada sección por evaluar). El Minedu monitoreó estos procesos. En la ECE se embalaron 64 164 cajas con instrumentos, y en la EM, 518 cajas. En todo momento, se garantizó la confidencialidad de los instrumentos.

Una vez moduladas, las cajas con los instrumentos fueron trasladadas en vehículos que aseguraron su arribo en adecuadas condiciones de seguridad a las sedes, donde se realizó un inventario del número de cajas. Este traslado se llevó a cabo cumpliendo un cronograma y fue monitoreado por el personal del operador logístico y del Minedu. Cuando las cajas llegaron a cada sede, fueron inventariadas y dispuestas en el almacén del local de jurisdicción de la sede, el cual debía permanecer bajo llave y estar custodiado por el supervisor de almacén y con resguardo policial o privado hasta el fin del operativo.

3.5 Proceso de aplicación

La duración y las fechas de aplicación de cada operativo de las evaluaciones 2018 se muestran en la siguiente tabla:

Tabla 3.5 Duración y fechas de la evaluación

Grados a evaluar	Duración	Fechas a evaluar ^a	Criterio
4.º grado de primaria	2 días	5 y 6 de noviembre	Un aplicador líder por IE de una sola sección que requiera pernocte de los aplicadores Un aplicador líder para IE urbanas que tengan cinco o más secciones programadas Un aplicador por sección Un aplicador NEE para secciones que cuenten con estudiantes con discapacidad según registro de información UMC_ R-NEE
4.º grado de primaria EIB solo L2	1 día	5 de noviembre	Un aplicador líder por IE Un aplicador por sección Un aplicador NEE para secciones que cuenten con estudiantes con discapacidad según registro de información UMC_ R-NEE
4.º grado de primaria EIB L1 y L2	2 días	5 y 6 de noviembre	Un aplicador líder por IE Un aplicador por sección Un aplicador NEE para secciones que cuenten con estudiantes con discapacidad según registro de información UMC_ R-NEE
2.º grado de secundaria	3 días	22, 23 y 24 de octubre	Mínimo un aplicador líder para IE que no requiera pernocte (según indicador del padrón) que tenga dos o más secciones programadas Un aplicador líder por IE de una sola sección que requiera pernocte de los aplicadores Un aplicador por sección Un aplicador adicional para secciones de 30 a más estudiantes Un aplicador NEE para secciones que cuenten con estudiantes con discapacidad según registro de información UMC_ R-NEE
6.º grado de primaria	1 día	19 de noviembre	Un aplicador líder por IE con una sección programada que requiera pernocte de los aplicadores. Un aplicador por sección Un aplicador NEE para secciones que cuenten con estudiantes en condición de discapacidad según registro de información UMC*

Nota. "EIB" refiere a Educación Intercultural Bilingüe. Asimismo, "L1 y L2" refiere a las escuelas EIB que enseñan en lengua originaria (L1) y en castellano como segunda lengua (L2), mientras que "solo L2" indica a las escuelas que solo enseñan en castellano como segunda lengua.

^a En el caso de la ECE, estas fechas aplican para las IE en las que no se requiere pernocte en el traslado de los aplicadores, en las otras IE se aplicó una semana después.

A continuación, se presenta el resumen del proceso de aplicación:

Cuarto grado de primaria y cuarto grado de primaria EIB (L1 y L2)

- I *Recepción de materiales.* En las IE que no requieren pernocte, el aplicador acude cada día de aplicación a la sede para recoger los instrumentos de evaluación. Los aplicadores de las IE que requieren pernocte se desplazan con el tiempo necesario para iniciar la aplicación en las fechas posteriores.
- II *Presentación ante el director de la IE y organización de la aplicación.* El aplicador llega temprano a la IE, se presenta ante el director, le entrega los documentos enviados por el Minedu, y le solicita la nómina de matrícula del grado a evaluar y la lista auxiliar de asistencia del docente. El docente firma el compromiso de confidencialidad.
- III *Preparación de los materiales antes de la evaluación.* El aplicador solicita al director un espacio privado para preparar los cuadernillos y demás instrumentos. Solo y sin ayuda, el aplicador verifica la información de la nómina de matrícula con los nombres de los estudiantes de la FOAR y los cuadernillos.
- IV *Preparación del aula e indicaciones.* El aplicador organiza el mobiliario del aula, ubica a los alumnos según el orden de la lista de estudiantes de la FOAR, entrega los cuadernillos de las pruebas y da lectura a las indicaciones de acuerdo con el guion de aplicación.
- V *Orden y duración de las sesiones de aplicación.* El aplicador sigue el guion de aplicación y administra los cuadernillos siguiendo la secuencia y los tiempos indicados en la tabla 3.6.

Tabla 3.6 Orden y duración de las sesiones de aplicación

4.º grado de primaria		4.º grado de primaria EIB L1 y L2	
Primer día	Segundo día	Primer día	Segundo día
Lectura (60 minutos)	Matemática (60 minutos)	Lectura en lengua originaria (45 minutos)	Lectura en castellano como segunda lengua (45 minutos)
Descanso (30 minutos)	Descanso (30 minutos)	Descanso (30 minutos)	Descanso (30 minutos)
Matemática (60 minutos)	Lectura (60 minutos)	Lectura en castellano como segunda lengua (45 minutos)	Lectura en lengua originaria (45 minutos)

- VI *Procedimiento al culminar un día de aplicación.* Cada día, el aplicador se dirige a un ambiente privado dentro de la IE para realizar el traslado de las respuestas de los cuadernillos a la FOAR. Luego, el director verifica que el traslado haya sido correcto y haya respetado los procedimientos. Al culminar la revisión, el aplicador cuenta la totalidad de cuadernillos y procede a lacrarlos utilizando una bolsa de seguridad. El director presencia el conteo y firma un acta de conformidad que da fe de la cantidad de cuadernillos que retornan al Minedu. Al terminar las actividades previstas en la IE, el aplicador retorna inmediatamente a la sede para entregar todos los instrumentos de aplicación de la ECE.
- VII *Procedimiento al culminar el último día de aplicación.* El director o personal responsable de la IE da conformidad a las condiciones de la aplicación en una declaración jurada de aplicación, la cual es firmada para cada sección evaluada. Para los casos en los que no se concretó la aplicación, se cuenta con una constancia de no aplicación, la cual detalla el motivo de la no aplicación, que es corroborado con una constancia de la UGEL.
- VIII *Procedimiento al culminar el trabajo en la IE.* El aplicador se dirige a la sede donde el asistente de procesos verifica el material devuelto teniendo en cuenta la cantidad y la calidad de la información recogida. Luego de este proceso, los asistentes de procesos, y los supervisores de almacén y sistemas hacen un inventario de los instrumentos. Las fichas ópticas son retiradas de sus cajas, y son clasificadas y ordenadas según su correlativo y tipo de instrumento. Los paquetes de cuadernillos retornan en sus mismas cajas.

Cuarto grado de primaria EIB (solo L2)

- I *Recepción de materiales.* En las IE que no requieren pernocte, el aplicador acude el mismo día de aplicación a la sede para recoger los instrumentos de evaluación. Los aplicadores de las IE que requieren pernocte se desplazan con el tiempo necesario para iniciar la aplicación en las fechas posteriores.
- II *Presentación ante el director de la IE y organización de la aplicación.* El aplicador llega temprano a la IE, se presenta ante el director, le entrega los documentos enviados por el Minedu, y le solicita la nómina de matrícula del grado a evaluar y la lista auxiliar de asistencia del docente. El docente firma el compromiso de confidencialidad.
- III *Preparación de los materiales antes de la evaluación.* El aplicador solicita al director un espacio privado para preparar los cuadernillos y demás instrumentos. Solo y sin ayuda, el aplicador verifica la información de la nómina de matrícula con los nombres de los estudiantes de la FOAR y los cuadernillos.
- IV *Preparación del aula e indicaciones.* El aplicador organiza el mobiliario del aula, ubica a los alumnos según el orden de la lista de estudiantes de la FOAR, entrega

los cuadernillos de las pruebas y da lectura a las indicaciones de acuerdo con el guion de aplicación.

V *Orden y duración de las sesiones de aplicación.* El aplicador administra los cuadernillos siguiendo la secuencia y los tiempos indicados en la tabla 3.7

Tabla 3.7 Orden y duración de las sesiones de aplicación de la ECE de 4.º grado de primaria EIB (solo L2)

Grado y operativo	Único día
4.º grado de primaria EIB (solo L2)	Lectura en castellano como segunda lengua (45 minutos)
	Descanso (30 minutos)
	Lectura en castellano como segunda lengua (45 minutos)

VI *Procedimiento al culminar el día de aplicación.* El aplicador se dirige a un ambiente privado dentro de la IE para realizar el traslado de las respuestas de los cuadernillos a la FOAR. Luego, en la dirección, el director verifica que el traslado haya sido correcto y se hayan respetado los procedimientos. Al culminar la revisión, el aplicador cuenta la totalidad de cuadernillos y procede a lacrarlos utilizando una bolsa de seguridad. El director presencia el conteo y firma un acta de conformidad que da fe de la cantidad de cuadernillos que retornan al Minedu.

Asimismo, el director o personal responsable de la IE da conformidad a las condiciones de la aplicación en una declaración jurada de aplicación, la cual es firmada para cada sección evaluada. Antes de retirarse de la IE, el aplicador se asegura de no dejar ningún instrumento en la IE. Para los casos en los que no se concretó la aplicación, se cuenta con una constancia de no aplicación, la cual detalla el motivo de la no aplicación, que es corroborado con una constancia de la UGEL.

VII *Procedimiento al culminar el trabajo en la IE.* El aplicador se dirige a la sede donde el asistente de procesos verifica el material devuelto teniendo en cuenta la cantidad y la calidad de la información recogida. Luego de este proceso, los asistentes de procesos, y los supervisores de almacén y sistemas hacen un inventario de los instrumentos. Las fichas ópticas son retiradas de sus cajas, clasificadas y ordenadas según su correlativo y tipo de instrumento. Los paquetes de cuadernillos retornan en sus mismas cajas.

Segundo grado de secundaria

- I *Recepción de materiales.* En las IE que no requieren pernocte, el aplicador acude cada día de aplicación a la sede para recoger los instrumentos de evaluación. Los aplicadores de las IE que requieren pernocte se desplazan con el tiempo necesario para iniciar la aplicación en las fechas posteriores.
- II *Presentación ante el director de la IE y organización de la aplicación.* El aplicador llega temprano a la IE, se presenta ante el director, le entrega los documentos enviados por el Minedu, y le solicita la nómina de matrícula del grado a evaluar y la lista auxiliar de asistencia del docente. Se pide que se designe un docente para que apoye en la organización de los estudiantes y se le solicita que firme el compromiso de confidencialidad.
- III *Preparación del aula e indicaciones.* El aplicador organiza el mobiliario del aula, ubica a los estudiantes según el orden del registro de estudiantes, entrega los cuadernillos de las pruebas y las fichas de respuestas, y da lectura a las indicaciones de acuerdo con el guion de aplicación.
- IV *Orden y duración de las sesiones de aplicación.* El aplicador administra los cuadernillos siguiendo la secuencia y los tiempos indicados en la tabla 3.8

Tabla 3.8 Orden y duración de las sesiones de aplicación de la ECE de 2.º grado de secundaria

Grado y operativo	Primer día	Segundo día	Tercer día
2.º grado de secundaria	Lectura (60 minutos)	Matemática (70 minutos)	Ciencia y Tecnología (90 minutos)
	Descanso (10 minutos)	Descanso (10 minutos)	Descanso (10 minutos)
	Matemática (70 minutos)	Lectura (60 minutos)	Ciencias Sociales (70 minutos)
		Cuestionario para el estudiante (30 minutos)	

- V *Procedimiento al culminar un día de aplicación.* En un ambiente privado, el aplicador revisa las fichas de respuestas y, en caso sea necesario, completa la información de los estudiantes con la nómina de matrícula. Luego, en la dirección, el aplicador cuenta la totalidad de cuadernillos y procede a lacrarlos utilizando una bolsa de seguridad. El director presencia el conteo y firma un acta

de conformidad respecto de las condiciones de aplicación, dando fe de la cantidad de cuadernillos de retorno. Asimismo, el último día de aplicación el director o personal responsable de la IE da conformidad a las condiciones de la aplicación en una declaración jurada de aplicación, la cual es firmada para cada sección evaluada. Para los casos en los que no se concretó la aplicación, se cuenta con una constancia de no aplicación, la cual detalla el motivo de la no aplicación, que es corroborado con una constancia de la UGEL.

VI *Procedimiento al culminar el trabajo en la IE.* El aplicador se dirige a la sede donde el asistente de procesos verifica el material devuelto teniendo en cuenta la cantidad y la calidad de la información recogida. Luego de este proceso, los asistentes de procesos, y los supervisores de almacén y sistemas hacen un inventario de los instrumentos. Las fichas ópticas son retiradas de sus cajas, y clasificadas y ordenadas según su correlativo y tipo de instrumento. Los paquetes de cuadernillos retornan en sus mismas cajas.

Sexto grado de primaria (EM)

- I *Recepción de materiales.* Los aplicadores de las IE que requieren pernocte se desplazan con el tiempo necesario para iniciar la aplicación en la fecha indicada (19 de noviembre). En las IE que no requieren pernocte, el aplicador acude a la sede el mismo día de aplicación para recoger los instrumentos de evaluación.
- II *Presentación ante el director de la IE y organización de la aplicación.* El aplicador llega temprano a la IE, se presenta ante el director, le entrega los documentos enviados por el Minedu, y le solicita la nómina de matrícula del grado a evaluar y la lista auxiliar de asistencia del docente. El docente firma el compromiso de confidencialidad.
- III *Preparación del aula e indicaciones.* El aplicador organiza el mobiliario del aula, ubica a los estudiantes según el orden del registro de estudiantes, entrega los cuadernillos de las pruebas y da lectura a las indicaciones de acuerdo con el guion de aplicación.
- IV *Orden y duración de las sesiones de aplicación.* El aplicador administra los cuadernillos siguiendo la secuencia y los tiempos indicados en la tabla 3.9.

Tabla 3.9 Orden y duración de las sesiones de aplicación de la EM de 6.º grado de primaria

Grado y operativo	Único día
6.º grado de primaria	Ciudadanía (90 minutos)
	Descanso (20 minutos)
	Cuestionario al estudiante
	Sección I (25 minutos)
	Sección II (25 minutos)

V *Procedimiento al culminar un día de aplicación.* En un ambiente privado, el aplicador completa el registro de estudiantes y la ficha del aplicador. Además, en caso sea necesario, completa la información de los estudiantes con la nómina de matrícula. Luego, en la dirección, el aplicador cuenta la totalidad de cuadernillos y procede a lacrarlos utilizando una bolsa de seguridad. El director presencia el conteo y firma un acta de conformidad respecto de las condiciones de aplicación, dando fe de la cantidad de cuadernillos de retorno. Asimismo, el director o personal responsable de la IE da conformidad a las condiciones de la aplicación en una declaración jurada de aplicación, la cual es firmada para cada sección evaluada. Antes de retirarse de la IE, el aplicador se asegura de no dejar ningún instrumento en la IE. Para los casos en los que no se concretó la aplicación, se cuenta con una constancia de no aplicación, la cual detalla el motivo de la no aplicación, que es corroborado con una constancia de la UGEL.

VI *Procedimiento al culminar el trabajo en la IE.* El aplicador se dirige a la sede donde el asistente de procesos verifica el material devuelto teniendo en cuenta la cantidad y la calidad de la información recogida. Luego de este proceso, los supervisores de aplicación, y los supervisores de almacén y sistemas hacen un inventario de los instrumentos. Todos los instrumentos y los paquetes de cuadernillos retornan en sus mismas cajas.

Procesamiento de datos y análisis psicométrico

Capítulo 4

La UMC ejecuta un protocolo que contempla los lineamientos establecidos por la American Psychological Association (APA), la American Educational Research Association (AERA) y el National Council on Measurement in Education (NCME). Ello permite garantizar la calidad de la información obtenida mediante la aplicación de sus evaluaciones. En este capítulo, se detalla cada uno de estos procesos.

4.1 Gestión y depuración de datos

Este proceso tiene como finalidad convertir la información contenida en los documentos físicos (las fichas ópticas) en información digital (base de datos) consistente y confiable. Para ello, la UMC elabora especificaciones técnicas que detallan los procedimientos que se requieren para un correcto procesamiento de la información. Este proceso de depuración puede dividirse en tres etapas principales: gestión física, captura de datos e imágenes y depuración de bases de datos.

Luego, se contrata a una empresa que cuente con experiencia en este rubro y que siga los procedimientos bajo la atenta supervisión de la UMC. Dicha empresa provee todo el personal necesario, el cual es evaluado, seleccionado y capacitado por los representantes de la UMC.

4.1.1 Gestión física

Este proceso tiene como objetivo asegurar que toda la documentación física a procesar esté completa y lista para la captura de datos. En este momento, se realizan los siguientes pasos:

- **Recepción e inventario:** la empresa de captura de datos recibe todos los documentos físicos y realiza un inventario uno a uno escaneando el código de barras impreso en cada documento.
- **Clasificación:** se clasifican todos los documentos físicos según su tipo, los cuales pueden ser fichas ópticas de respuesta, cuestionarios u otros.
- **Preparación:** consiste en agrupar los documentos físicos en lotes pequeños y quitar las grapas de estos para una mejor gestión en la siguiente etapa.
- **Almacenamiento:** todos los documentos físicos son almacenados por la empresa de captura de datos de forma ordenada para poder acceder a estos en caso se requiera durante la siguiente etapa.

Para asegurar una correcta gestión física de los documentos, la empresa de captura de datos proporciona un local con el espacio suficiente para todas las etapas,

incluyendo un espacio exclusivo para el almacenamiento de todos los documentos debido a su confidencialidad.

4.1.2 Captura de datos

Este proceso tiene como objetivo digitalizar los documentos físicos y capturar los datos consignados en ellos.

- Digitalización y captura de datos: Consiste en pasar por un escáner especial todos los documentos físicos ya preparados. El escáner utilizado es capaz de digitalizar el documento físico en una imagen con buena calidad, así como capturar los datos consignados en este utilizando tres tecnologías: OMR (reconocimiento de marcas tipo burbujas), OCR (reconocimiento de caracteres impresos) e ICR (reconocimiento de caracteres escritos a mano).
- Control de calidad: Esta tarea se realiza cuando el software de reconocimiento de datos presenta un bajo nivel de confiabilidad de la data capturada. Estos casos son enviados a un grupo de personas dedicadas a revisar y asegurar que el dato capturado corresponda con el dato consignado en la imagen del documento.

4.1.3 Depuración de datos

En este proceso, se verifica que todos los campos de las bases de datos contengan respuestas que estén dentro de los márgenes permitidos y que exista coherencia entre los campos y la información adicional que se maneja en la evaluación. Para este proceso, la UMC elabora un manual del depurador, el cual contiene todas las revisiones a realizar en cada campo de todas las bases de datos. Además de este manual, los depuradores son capacitados para que realicen de la forma más eficiente posible dicho proceso. Durante esta etapa de depuración, algunas de las revisiones realizadas son las siguientes:

- Revisiones iniciales:
 - Correspondencia de campos según “diccionario de variables” y la frecuencia de cada campo
 - Correspondencia de registros según el reporte de inventario de instrumentos
 - Registros únicos en los campos de identificación del instrumento como el código de barras
 - Correspondencia entre código modular y correlativo Minedu (según padrón de IE)
 - Hojas faltantes en cada ficha óptica

- Revisiones principales (por tipo de documento):
 - Caracteres no válidos en los DNI de los estudiantes
 - Duplicados en los DNI de los estudiantes
 - Caracteres no válidos en los nombres y apellidos de los estudiantes
 - Duplicados en los apellidos y nombres de los estudiantes
 - Caracteres no válidos en el correlativo del estudiante
 - Duplicados en el correlativo del estudiante
 - Caracteres no válidos en el sexo del estudiante
 - Caracteres no válidos en la lengua materna del estudiante
 - Caracteres no válidos en la discapacidad del estudiante
 - Caracteres no válidos en el número de forma del cuadernillo
 - Caracteres no válidos en las respuestas del estudiante
 - Inconsistencia en las respuestas y la “asistencia” del estudiante
- Revisiones finales:
 - Repetir todas las revisiones iniciales
 - Contrastar con el listado de IE aplicadas proporcionado por el equipo encargado del trabajo de campo de la UMC
 - Identificar valores fuera del rango permitido

4.2 Procesamiento psicométrico

4.2.1 Modelo Rasch para ítems dicotómicos

El análisis psicométrico de las pruebas aplicadas en la ECE se basó en el modelo Rasch para ítems dicotómicos (Bond y Fox, 2015; Wright y Stone, 1998). Este modelo se centra en el análisis de cada ítem, concretamente en la interacción entre un ítem y una persona. Establece la probabilidad de respuesta de una persona ante un ítem en términos de la diferencia entre la medida de rasgo o habilidad latente de la persona (θ) y la medida del ítem utilizado en términos de su dificultad (β)⁸. Por este motivo, se le denomina usualmente modelo de un parámetro (Hambleton, Swaminathan y Rogers, 1991). Es importante señalar que, en este reporte, se utilizará $\hat{\theta}$ para referirse a la cantidad del rasgo latente estimado para una persona y b para la dificultad estimada de un ítem en una muestra.

En los modelos Rasch, la habilidad de las personas y las dificultades de los ítems se ubican en la misma métrica. Al respecto, Smith y Kramer (1989) afirman que la existencia de una métrica común permite combinar la habilidad de la persona y la

⁸En algunos textos, se utiliza la letra B para referirse a la habilidad de las personas y la letra D para referirse a la dificultad de los ítems.

dificultad del ítem para predecir el desempeño de esta persona en un ítem cualquiera e identificar respuestas inesperadas. El objetivo central del análisis Rasch es construir una escala conformada por los ítems ordenados según su dificultad. Ello implica que, a mayor habilidad, la persona presentará mayor probabilidad de acertar y, por lo tanto, obtendrá un mayor número de respuestas correctas. Es muy importante tener en cuenta que la medida estimada de la persona no es igual al puntaje directo que posee (número de ítems correctos); este puntaje es solo un insumo a partir del cual se construye la medida Rasch.

La relación entre la habilidad y la dificultad puede graficarse por medio de las curvas características del ítem (CCI), que brindan información concreta sobre la probabilidad de respuesta de una persona ante dicho ítem. Al trazar estas curvas, se establecen las siguientes relaciones en caso de tener ítems dicotómicos:

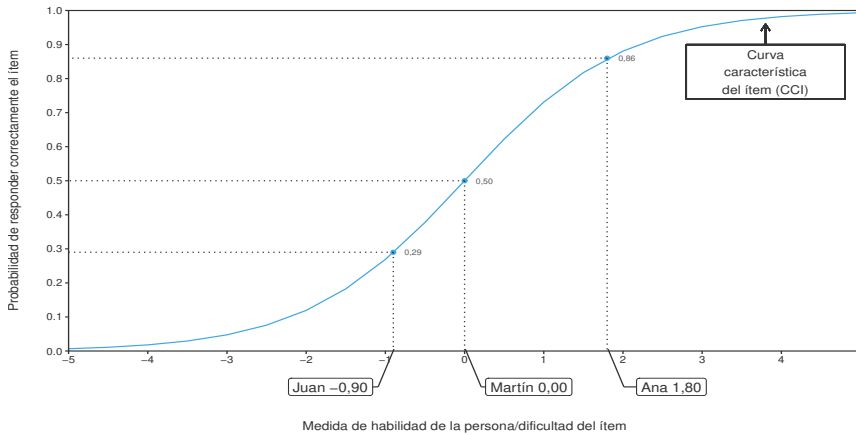
$$\begin{aligned}\theta > \beta; & \quad p(X = 1 \mid \theta, \beta) \in]0,5; 1,0] \\ \theta < \beta; & \quad p(X = 1 \mid \theta, \beta) \in [0,0; 0,5[\\ \theta = \beta; & \quad p(X = 1 \mid \theta, \beta) = 0,5\end{aligned}$$

En el primer caso, si la habilidad de la persona es mayor que la dificultad del ítem, la probabilidad de responder correctamente dicho ítem es mayor que 0,5 (50 %). En el segundo caso, si la habilidad de la persona es menor que la dificultad del ítem, la probabilidad de responder correctamente dicho ítem es menor que 0,5 (50 %). Finalmente, en el tercer caso, si la habilidad de la persona es igual a la dificultad del ítem, la probabilidad de responder correctamente dicho ítem es igual a 0,5 (50 %). Como señala Ingebo (1997), de esta manera se puede comprobar empíricamente la teoría de que los estudiantes con mayores conocimientos tienen mayores probabilidades de responder correctamente una pregunta, a diferencia de los estudiantes con menores conocimientos. Matemáticamente, la CCI se grafica con la siguiente función:

$$P(X_{ni} = 1 \mid \theta_n, \beta_i) = \frac{e^{\theta_n - \beta_i}}{1 + e^{\theta_n - \beta_i}},$$

donde θ representa el rasgo latente de cada persona n y β , la medida de dificultad de cada ítem i . Esta relación entre la habilidad de una persona y la dificultad de un ítem se presenta en la siguiente figura:

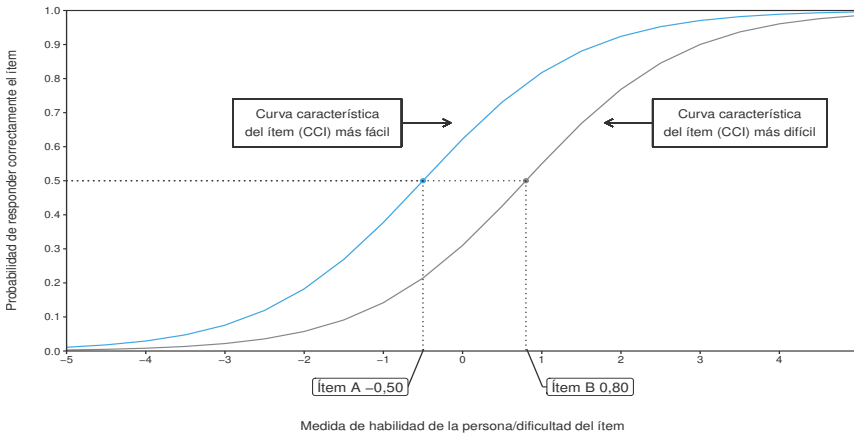
Figura 4.1 Curva característica del ítem y probabilidades de tres personas de responder correctamente dicho ítem



En la figura 4.1, se observa que, al enfrentarse a este ítem, Juan, cuya habilidad es de $-0,90$, tiene una probabilidad de $0,29$ de acertarlo; es decir, lo más probable es que lo falle y obtenga 0 puntos. En cambio, Ana, cuya habilidad se ha estimado en $1,80$, tiene una probabilidad de $0,86$ de acertar este ítem; por lo tanto, es más probable que lo responda correctamente y reciba 1 punto. Finalmente, Martín tiene una habilidad igual a la dificultad del ítem; por eso, se afirma que tiene iguales posibilidades de acertar o fallar el ítem.

Al comparar dos o más curvas características de ítems, se puede decir que, mientras más hacia la derecha se desplaza una curva, más difícil es el ítem. Además, se expresa numéricamente la dificultad de un ítem como aquel valor en la escala de habilidad que posee una probabilidad del 50 % de acertar dicho ítem. En la figura que aparece a continuación, el ítem más fácil (A) tiene una dificultad de $-0,50$ y el más difícil (B), de $0,80$:

Figura 4.2 Curva característica de dos ítems y comparación de la dificultad de ambos



Como ya se señaló, θ representa la medida verdadera del rasgo latente de una persona n . Es decir, sería su medida de habilidad si se le pudiese evaluar en condiciones óptimas y con todos los ítems que potencialmente pueden usarse para medir dicho rasgo latente. Pero como ello no es posible, en términos empíricos, lo que se tiene es una estimación de dicha habilidad, representada por $\hat{\theta}$. Por su parte, β representa el parámetro de dificultad de un ítem i a nivel poblacional, por lo cual a nivel de muestra se le representa como b . De esta manera, el modelo Rasch puede ser expresado del siguiente modo:

$$P_{ni} = P(X_{ni} = 1 | \hat{\theta}_n, b_i) = \frac{e^{\hat{\theta}_n - b_i}}{1 + e^{\hat{\theta}_n - b_i}}$$

La dificultad del ítem (b) es un estimado, pues está sujeta a un grado de incertidumbre, el cual se expresa mediante el error estándar de la dificultad del ítem (Wilson, 2005). Este error estándar puede ser calculado mediante esta fórmula:

$$e.e.(b_i) = \frac{1}{\sqrt{\sum_n^{I-1} N_n P_{ni}(1 - P_{ni})}}$$

4.3 Modelo Rasch de crédito parcial

Masters (1982) introdujo el modelo de crédito parcial para trabajar ítems politómicos de categorías ordenadas. Se deriva de los test de opción múltiple: algunas respuestas incorrectas reciben crédito o puntaje parcial porque indican algún conocimiento sobre la respuesta apropiada (Wright, 1999). Además, este modelo especifica que cada ítem tiene su propia estructura de calificación. Según Fox (1999), es una generalización del modelo Rasch para ítems dicotómicos que puede aplicarse en test cuyos ítems

pueden variar en el número de alternativas correctas y en la cantidad de opciones de respuesta.

Este modelo supone que el proceso de resolución de un ítem involucra la aplicación secuencial de un conjunto de pasos. Los créditos parciales dados equivalen al número de pasos que deben completarse antes de obtener la respuesta final del problema (Masters, 1982). Por ejemplo, supongamos que un estudiante debe hallar el valor de x en el siguiente ítem matemático:

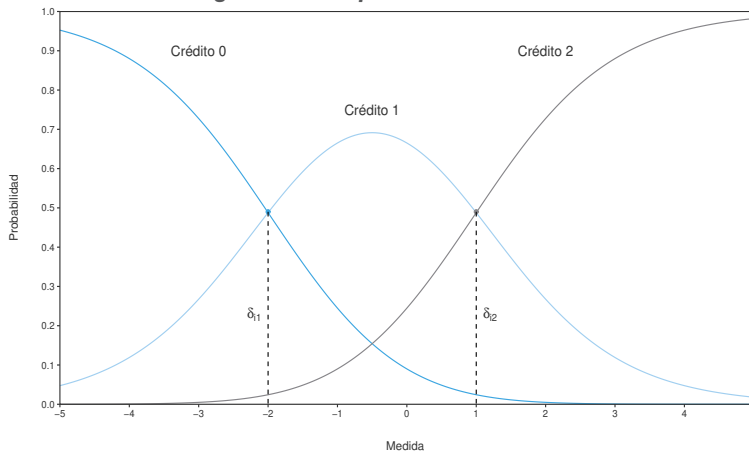
$$\sqrt{\frac{196}{4}} = x$$

Existen dos pasos para la solución de este ítem: (1) se divide 196 entre 4, lo que da como resultado 49, y (2) se saca la raíz cuadrada de 49 para obtener el resultado final, que es 7. Si el estudiante no puede resolver siquiera la división, recibe 0 puntos. Si es capaz de resolver la división, pero no la raíz cuadrada, recibe 1 punto. Si además de resolver la división obtiene la raíz cuadrada, recibe 2 puntos.

El modelo asume que puntajes o créditos superiores son más probables de ser alcanzados en ítems con baja dificultad que en ítems con alta dificultad. Además, los estudiantes con mayores niveles del rasgo latente tienen mayor probabilidad de recibir el crédito total para el ítem que los estudiantes con menores niveles del rasgo latente. Ambas propiedades deben mantenerse para todas las personas, todos los ítems y todas las categorías de respuesta. Masters (1982) indica que, cuando estas propiedades se cumplen, los puntajes directos obtenidos pueden ser transformados para lograr una escala de intervalo.

Una manera de graficar el modelo de crédito parcial es mediante las curvas de categorías de respuesta (Embretson y Reise, 2000). Estas curvas muestran la relación entre la cantidad del rasgo latente evaluado y la probabilidad de obtener puntajes parciales. Además, indican en qué lugar del continuo del rasgo latente las respuestas a una categoría son más probables que las respuestas a otra categoría. La figura 4.3 presenta un ejemplo de este tipo de curvas:

Figura 4.3 Curva de categorías de respuestas con umbrales Rasch-Andrich



Los parámetros δ_{ij} se encuentran en la intersección de las curvas de las categorías de respuesta. Se les denomina umbrales Rasch-Andrich (Linacre, 2001) y representan aquel valor en la escala del rasgo latente en el cual es igualmente probable recibir un puntaje o el inmediatamente adyacente. En algunos casos (Linacre, 2001), δ_{ij} es descompuesto en dos términos, D_{ij} y F_{ij} , de tal manera que $\delta_{i1} = D_i + F_{i1}$ y $\delta_{i2} = D_i + F_{i2}$.

En este ejemplo, el primer umbral con un valor de $-1,74$ representa la medida en la cual es igualmente probable recibir 0 y 1 punto en el ítem. A partir de la figura 4.3, se puede señalar además que, si una persona tiene una medida del rasgo latente menor a $-1,74$, lo más probable es que haya recibido 0 puntos (crédito 0) en este ítem. Si su medida se encuentra entre $-1,71$ y $0,52$, lo más probable es que reciba 1 punto (crédito 1). Finalmente, si su medida de habilidad es mayor que $0,52$, lo más probable es que reciba 2 puntos (crédito 2).

Para expresar matemáticamente el modelo de crédito parcial, se debe partir del supuesto de que los ítems pueden recibir puntuaciones desde 0 hasta un máximo (2, 3, etc.). Ello se expresa del siguiente modo:

$$X_{ni} = x \in (0, 1, \dots, m_i),$$

donde:

n : identificador (número) de persona

i : identificador (número) de ítem

X_{ni} : puntaje obtenido por una persona n en un ítem i

x : puntaje observado

m_i : crédito máximo posible en un ítem i

La probabilidad de obtener cierto puntaje en un ítem viene dada por el modelo de crédito parcial, el cual se representa usando una adaptación de la notación original propuesta por Masters (1982):

$$P(X_{ni} = x) = \frac{\exp \sum_{j=0}^x (\theta_n - \delta_{ij})}{\sum_{k=0}^{m_i} \exp \sum_{j=0}^k (\theta_n - \delta_{ij})},$$

donde:

j : número del umbral (0 a $m_i - 1$)

δ_{ij} : valor del umbral j , del ítem i

Es importante considerar que por conveniencia computacional el valor δ_{i0} se elige de tal manera que $\sum_{j=0}^k (\theta_n - \delta_{ij}) = 0$.

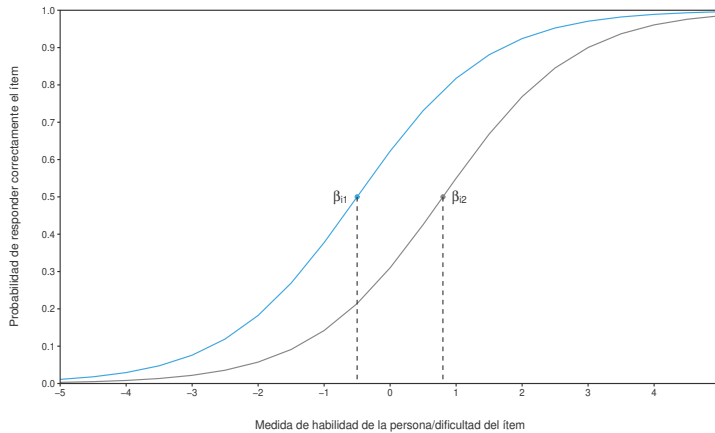
Supongamos que un ítem i tiene tres puntajes posibles, es decir, $X_{ni} = x \in (0,1,2)$. A partir de ello, se puede observar que $m_i = 2$. Por lo tanto, las probabilidades de obtener 0, 1 o 2 puntos se pueden expresar de la siguiente manera:

$$P(X_{ni} = 0) = \frac{\exp(0)}{\exp(0) + \exp(0 + (\theta_n - \delta_{i1})) + \exp(0 + (\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2}))}$$

$$P(X_{ni} = 1) = \frac{\exp(0 + (\theta_n - \delta_{i1}))}{\exp(0) + \exp(0 + (\theta_n - \delta_{i1})) + \exp(0 + (\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2}))}$$

$$P(X_{ni} = 2) = \frac{\exp(0 + (\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2}))}{\exp(0) + \exp(0 + (\theta_n - \delta_{i1})) + \exp(0 + (\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2}))}$$

Como ya se señaló, el modelo de crédito parcial intenta captar diferentes niveles de conocimiento frente a un mismo ítem. Por ejemplo, en un ítem de dos pasos, se modela la probabilidad de obtener más de 0 puntos y, luego, la probabilidad de obtener más de 1 punto. Este proceso se puede representar con dos curvas características del ítem, una para cada paso, como se muestra a continuación:

Figura 4.4 Curva de categorías de respuestas con umbrales Rasch-Thurstone

La curva de la izquierda se refiere a la probabilidad de recibir 1 o 2 puntos en lugar de 0. La curva de la derecha indica la probabilidad de recibir 2 puntos frente a 1 o 0 puntos. Es decir, una persona con una medida de rasgo latente de $-1,83$ tiene una probabilidad igual al 50 % de superar el primer paso (tener más de 0), mientras que una persona con una medida de $0,61$ tiene una probabilidad del 50 % de superar el segundo paso (más de 1 punto). A dichos valores ($-1,83$ y $0,61$) se les denomina umbrales Rasch-Thurstone (Linacre, 2001).

4.4 Análisis de ítems

El buen ajuste a un modelo es una parte importante en cualquier análisis estadístico, pero en el caso del análisis Rasch es una parte esencial (Wilson, 2005). Las relaciones datos-modelo son muy diferentes de lo que usualmente se aplica en el análisis estadístico tradicional, en el cual el desajuste de los datos con respecto al modelo suele dar lugar a la desestimación del modelo estadístico. Sin embargo, en el caso del análisis Rasch, los valores que no ajustan (misfit) conducen al rechazo de los datos, ya que no cumplen con los requisitos del modelo Rasch. Por lo tanto, se debe entender el modelo Rasch como un modelo prescriptivo en el que se investiga cómo los datos se ajustan al modelo, en lugar de cómo el modelo se ajusta a los datos (Bond y Fox, 2015).

Como señala Wilson (2005), el ajuste de los ítems al modelo Rasch se fundamenta en dos estadísticos: el *outfit* e *infit*, cuyo cálculo se basa en los residuos cuadráticos estandarizados del modelo (Z_{ni}^2). A partir de las propuestas de diversos autores (Bond y Fox, 2015; Wilson, 2005; Wright y Linacre, 1994), se exponen brevemente el significado y el cálculo de ambos indicadores de ajuste.

El *outfit* (*outlier sensitive fit statistic*) es más sensible al comportamiento inesperado alejado de la medida de una persona. Se basa en la suma de cuadrados de los

residuales estandarizados (Z_{ni}^2), de tal manera que es como sigue:

$$Outfit_i = \frac{1}{N} \sum_{n=1}^N Z_{ni}^2$$

El *infit* (*information weighted fit statistic*) está afectado básicamente por respuestas no esperadas a ítems calibrados cerca de la medida de la persona. Su cálculo se basa en la siguiente fórmula:

$$Infit_i = \frac{\sum_{n=1}^N R_{ni}^2}{\sum_{n=1}^N VAR(P_{ni})}$$

Como $R_{ni}^2 = VAR(P_{ni})Z_{ni}^2$, se puede redefinir el *infit* como una media cuadrática ponderada de residuales (Z_{ni}^2) que es sensible a patrones de respuesta irregulares:

$$Infit_i = \sum_{n=1}^N \omega_{ni} Z_{ni}^2$$

En la fórmula anterior, los residuales están ponderados por sus varianzas individuales:

$$\omega_{ni} = \frac{VAR(P_{ni})}{\sum_{n=1}^N VAR(P_{ni})}$$

En cuanto a la interpretación de los valores *infit* y *outfit*, puede señalarse como ejemplo que un valor de ajuste de 1,25 indica que hay un 25 % más de “ruido” en los datos que el modelado.

Además de los índices antes mencionados, se puede calcular la correlación de Pearson entre la respuesta al ítem y la medida estimada para cada persona (*ptme*). A continuación, se presenta la fórmula para calcularlo (Linacre, 2019):

$$ptme = \frac{\sum_{n=1}^N [(X_n - \sum_{m=1}^N X_m / N)(\hat{\theta}_n - \sum_{m=1}^N \hat{\theta}_m / N)]}{\sqrt{\sum_{n=1}^N [(X_n - \sum_{m=1}^N X_m / N)]^2 \sum_{n=1}^N [(\hat{\theta}_n - \sum_{m=1}^N \hat{\theta}_m / N)]^2}}$$

donde X_1, \dots, X_I son las respuestas a los ítems, y $\hat{\theta}_1, \dots, \hat{\theta}_N$ son las medidas estimadas para cada persona.

Siguiendo las recomendaciones de Linacre (2019), se espera que los ítems con un ajuste adecuado al modelo Rasch tengan valores *infit* y *outfit* entre 0,50 y 1,50; sin embargo, dentro de este grupo, se prefieren los valores entre 0,70 y 1,30 (Wright y Linacre, 1994). Algunos autores (Schulz, 1990) señalan que debe prestársele mayor atención al *infit*, ya que es menos sensible a las variaciones en el tamaño de la

muestra utilizada para calibrar los ítems. Además, se esperan correlaciones ítem medida (ptme) positivas (Linacre, 2019).

Finalmente, se calcula la medida de dificultad de un ítem, utilizada por la teoría clásica de los test (p), que corresponde a la tasa de acierto. Sus valores fluctúan entre 0 y 1, y los valores más bajos representan ítems más difíciles. Este índice para un ítem i se define como el cociente entre el número de personas que lo han acertado (A_i) y el número total de personas que lo han intentado resolver (N_i), tal como se muestra a continuación:

$$D_i = \frac{A_i}{N_i}$$

4.4.1 Calibración de ítems

A continuación, el detalle de las características psicométricas de los ítems ya mencionadas en el acápite anterior.

Tabla 4.1 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura de 4.º grado de primaria

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	1	-0,285	0,003	1,06	1,09	0,40	0,57	3
B01	2	-2,360	0,004	0,94	0,92	0,38	0,85	1
B01	3	-1,931	0,004	0,97	0,97	0,39	0,81	2
B01	4	0,308	0,003	1,18	1,30	0,31	0,47	4
B01	5	-3,700	0,006	0,88	0,64	0,33	0,94	1
B01	6	0,180	0,028	1,16	1,20	0,33	0,49	4
B02	1	-2,355	0,004	0,84	0,64	0,47	0,85	1
B02	2	-2,048	0,004	0,91	0,80	0,44	0,82	2
B02	3	-0,974	0,003	0,94	0,90	0,48	0,68	2
B02	4	-1,300	0,003	0,87	0,76	0,52	0,73	2
B02	5	-0,500	0,003	1,03	1,04	0,42	0,61	3
B02	6	-0,368	0,003	1,09	1,13	0,37	0,58	3
B03	1	-2,329	0,004	0,94	0,85	0,40	0,85	1
B03	2	-1,170	0,003	0,94	0,87	0,47	0,71	2
B03	3	-2,720	0,005	0,94	0,89	0,36	0,89	1
B03	4	-1,209	0,003	0,93	0,88	0,47	0,72	2
B03	5	-1,239	0,003	0,99	0,98	0,43	0,72	2
B03	6	0,502	0,003	0,94	0,98	0,49	0,43	4
B04	1	-1,142	0,003	0,97	0,91	0,45	0,70	2
B04	2	-1,659	0,004	0,94	0,84	0,45	0,78	2
B04	3	0,101	0,003	1,21	1,31	0,28	0,50	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B04	4	0,598	0,003	1,14	1,25	0,33	0,42	4
B04	5	-0,216	0,003	1,03	1,05	0,43	0,55	3
B04	6	-0,407	0,028	0,97	0,98	0,47	0,59	3
B05	1	-1,244	0,003	0,88	0,78	0,51	0,72	2
B05	2	0,531	0,003	0,99	1,03	0,45	0,43	4
B05	3	-1,090	0,003	1,02	1,04	0,40	0,70	2
B05	4	0,142	0,003	0,95	0,93	0,49	0,49	3
B05	5	-0,619	0,003	0,98	0,93	0,46	0,62	3
B05	6	0,444	0,003	1,14	1,22	0,34	0,44	4
B06	1	-2,147	0,004	0,97	0,94	0,38	0,83	1
B06	2	-1,780	0,004	0,94	0,90	0,43	0,79	2
B06	3	-0,578	0,003	1,03	1,04	0,42	0,62	3
B06	4	-2,414	0,004	0,88	0,69	0,44	0,86	1
B06	5	0,169	0,003	1,12	1,20	0,35	0,49	4
B06	6	-2,107	0,004	0,93	0,90	0,41	0,83	1
B07	1	0,152	0,005	1,11	1,17	0,35	0,48	4
B07	2	-0,356	0,005	0,95	0,92	0,48	0,57	3
B07	3	-0,884	0,006	1,23	1,30	0,24	0,66	3
B07	4	-0,428	0,005	0,96	0,92	0,48	0,58	3
B07	5	0,799	0,006	1,12	1,22	0,34	0,37	4
B07	6	0,776	0,006	1,11	1,21	0,34	0,38	4
B08	1	0,927	0,006	1,00	1,12	0,42	0,36	4
B08	2	0,195	0,005	1,03	1,06	0,42	0,48	4
B08	3	-0,650	0,006	1,20	1,35	0,27	0,63	3
B08	4	-0,114	0,005	1,04	1,06	0,41	0,54	3
B08	5	-0,079	0,005	1,02	1,03	0,43	0,53	3
B08	6	-	-	-	-	-	-	-
B09	1	0,143	0,005	0,97	0,98	0,48	0,51	4
B09	2	-0,287	0,006	0,94	0,92	0,50	0,58	3
B09	3	0,087	0,005	0,86	0,83	0,57	0,52	3
B09	4	0,064	0,005	0,89	0,87	0,54	0,52	3
B09	5	-1,842	0,007	0,91	0,86	0,45	0,81	2
B09	6	-1,020	0,006	0,88	0,78	0,52	0,70	2
B10	1	-0,219	0,005	0,94	0,92	0,49	0,54	3
B10	2	-1,644	0,006	1,01	1,04	0,37	0,77	2
B10	3	1,257	0,006	1,15	1,33	0,30	0,30	4
B10	4	-0,903	0,006	0,99	0,95	0,44	0,66	3
B10	5	-0,930	0,006	0,95	0,91	0,47	0,66	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B10	6	-0,836	0,006	1,09	1,11	0,36	0,65	3
B11	1	-2,991	0,009	0,87	0,64	0,41	0,90	1
B11	2	-1,668	0,006	0,95	0,90	0,44	0,77	2
B11	3	-1,911	0,007	0,83	0,72	0,51	0,80	2
B11	4	-1,804	0,006	0,99	0,92	0,40	0,79	2
B11	5	-1,826	0,006	1,06	1,19	0,33	0,79	2
B11	6	0,096	0,005	1,14	1,20	0,34	0,50	3
B12	1	0,749	0,005	1,11	1,24	0,36	0,41	4
B12	2	-1,460	0,006	0,88	0,77	0,50	0,76	2
B12	3	0,614	0,005	1,15	1,26	0,34	0,43	4
B12	4	0,620	0,005	1,02	1,09	0,44	0,43	4
B12	5	-1,293	0,006	0,97	0,92	0,44	0,74	2
B12	6	-1,016	0,006	1,02	1,09	0,40	0,70	2

Tabla 4.2 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Matemática de 4.º grado de primaria

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	1	-2,407	0,009	0,98	1,05	0,35	0,90	1
B01	2	0,755	0,006	1,22	1,41	0,38	0,41	4
B01	3	-0,571	0,006	1,02	0,99	0,49	0,66	3
B01	4	-2,456	0,009	1,04	1,17	0,31	0,91	1
B01	5	-	-	-	-	-	-	-
B01	6	-	-	-	-	-	-	-
B01	7	-	-	-	-	-	-	-
B01	8	-1,102	0,006	1,04	1,03	0,44	0,75	2
B01	9	-1,006	0,006	1,16	1,25	0,37	0,74	3
B01	10	-1,231	0,007	0,84	0,75	0,55	0,77	2
B02	1	-1,729	0,007	1,05	1,15	0,36	0,83	2
B02	2	-	-	-	-	-	-	-
B02	3	0,181	0,006	1,14	1,23	0,40	0,51	3
B02	4	1,318	0,006	1,12	1,47	0,37	0,29	4
B02	5	-1,447	0,007	1,05	1,09	0,39	0,80	2
B02	6	0,612	0,006	1,05	1,16	0,45	0,42	4
B02	7	0,372	0,006	1,14	1,21	0,40	0,47	4
B02	8	-1,034	0,006	0,89	0,82	0,53	0,73	3
B02	9	-0,624	0,006	1,24	1,34	0,32	0,66	3
B02	10	-0,036	0,006	1,22	1,34	0,35	0,55	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	1	-0,951	0,006	1,03	1,06	0,45	0,71	3
B03	2	-1,546	0,007	0,93	0,84	0,47	0,80	2
B03	3	-	-	-	-	-	-	-
B03	4	-2,973	0,011	0,94	0,91	0,31	0,94	1
B03	5	-	-	-	-	-	-	-
B03	6	-2,406	0,009	1,07	1,49	0,27	0,90	1
B03	7	0,994	0,006	0,97	1,06	0,53	0,35	4
B03	8	-0,591	0,006	1,24	1,31	0,36	0,65	3
B03	9	0,720	0,006	1,09	1,22	0,47	0,40	4
B03	10	0,309	0,006	1,00	1,06	0,52	0,48	3
B04	1	-	-	-	-	-	-	-
B04	2	-	-	-	-	-	-	-
B04	3	-0,146	0,006	1,05	1,07	0,48	0,58	3
B04	4	0,700	0,006	1,11	1,35	0,44	0,42	4
B04	5	-1,010	0,006	1,08	1,20	0,42	0,73	3
B04	6	-0,007	0,006	0,90	0,86	0,58	0,55	3
B04	7	-2,143	0,008	0,87	0,71	0,45	0,88	1
B04	8	-	-	-	-	-	-	-
B04	9	-	-	-	-	-	-	-
B04	10	-2,381	0,009	1,08	1,47	0,28	0,90	1
B05	1	-1,775	0,007	0,92	0,89	0,45	0,84	2
B05	2	-0,262	0,006	1,14	1,22	0,40	0,59	3
B05	3	0,607	0,006	0,95	1,07	0,51	0,42	4
B05	4	-0,889	0,006	0,92	0,87	0,52	0,71	3
B05	5	0,306	0,006	1,15	1,28	0,39	0,48	3
B05	6	-0,473	0,006	0,85	0,79	0,58	0,63	3
B05	7	-0,250	0,006	1,04	1,04	0,47	0,59	3
B05	8	-	-	-	-	-	-	-
B05	9	-0,669	0,006	0,88	0,79	0,56	0,67	3
B05	10	-2,724	0,010	0,89	0,77	0,38	0,92	1
B06	1	-1,891	0,007	0,99	1,04	0,39	0,85	2
B06	2	-	-	-	-	-	-	-
B06	3	0,672	0,006	1,10	1,33	0,45	0,41	4
B06	4	-0,462	0,006	0,99	0,98	0,51	0,63	3
B06	5	-2,086	0,008	0,96	1,00	0,39	0,87	1
B06	6	0,864	0,006	1,14	1,39	0,43	0,38	4
B06	7	-	-	-	-	-	-	-
B06	8	0,318	0,006	0,85	0,83	0,61	0,48	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B06	9	-0,586	0,006	1,08	1,12	0,45	0,65	3
B06	10	-1,402	0,007	0,97	0,91	0,46	0,78	2
BC1	1	2,045	0,037	0,92	0,75	0,52	0,20	4
BC1	2	-1,889	0,004	0,96	0,88	0,42	0,85	2
BC1	3	-	-	-	-	-	-	-
BC1	4	-1,624	0,004	0,85	0,68	0,51	0,82	2
BC1	5	0,196	0,003	1,02	1,05	0,50	0,51	3
BC1	6	-1,240	0,004	0,85	0,70	0,54	0,77	2
BC1	7	-0,814	0,003	1,03	0,99	0,47	0,70	3
BC1	8	-0,747	0,003	1,12	1,11	0,41	0,69	3
BC1	9	-	-	-	-	-	-	-
BC1	10	-1,351	0,004	0,96	0,85	0,47	0,78	2
BC1	11	-0,682	0,003	0,90	0,82	0,55	0,67	3
BC1	12	-1,495	0,004	0,92	0,91	0,47	0,80	2
BC1	13	-1,616	0,004	1,01	1,13	0,40	0,82	2
BC1	14	-2,225	0,005	0,91	0,75	0,42	0,88	1
BC1	15	-0,140	0,003	1,00	1,02	0,50	0,57	3
BC2	1	-	-	-	-	-	-	-
BC2	2	-2,791	0,006	0,99	1,39	0,28	0,93	1
BC2	3	-	-	-	-	-	-	-
BC2	4	-0,586	0,003	0,82	0,74	0,60	0,66	3
BC2	5	-1,506	0,004	1,08	1,07	0,38	0,80	2
BC2	6	-0,816	0,003	0,90	0,89	0,53	0,69	3
BC2	7	-1,593	0,004	1,07	1,04	0,38	0,81	2
BC2	8	-3,051	0,006	0,90	0,84	0,32	0,94	1
BC2	9	0,303	0,003	1,29	1,49	0,33	0,48	3
BC2	10	-1,529	0,004	0,87	0,77	0,50	0,80	2
BC2	11	-0,681	0,003	0,95	0,91	0,52	0,67	3
BC2	12	-1,856	0,004	0,99	1,05	0,39	0,84	2
BC2	13	-1,246	0,004	1,00	0,93	0,46	0,76	2
BC2	14	-2,216	0,005	0,90	0,80	0,42	0,88	1
BC2	15	-1,747	0,004	0,88	0,74	0,48	0,83	2

Tabla 4.3 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura en aimara de 4.º grado de primaria EIB

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	1	-0,929	0,159	1,01	1,00	0,20	0,77	1
BC1	2	-1,261	0,179	0,97	0,90	0,27	0,83	1
BC1	3	-1,481	0,196	0,96	0,83	0,29	0,87	1
BC1	4	-1,481	0,196	0,94	0,83	0,31	0,87	1
BC1	5	-1,165	0,172	0,94	0,89	0,31	0,82	1
BC2	1	0,821	0,147	0,99	0,96	0,26	0,29	2
BC2	2	0,804	0,147	1,04	1,05	0,15	0,30	2
BC2	3	0,548	0,139	0,96	0,96	0,31	0,36	2
BC2	4	0,282	0,135	1,00	1,01	0,24	0,44	2
BC2	5	0,610	0,141	1,03	1,04	0,17	0,35	2
BC3	1	-2,573	0,342	0,96	0,70	0,23	0,96	1
BC3	2	-2,406	0,311	0,97	0,71	0,23	0,95	1
BC3	3	-1,732	0,218	0,93	0,74	0,34	0,90	1
BC3	4	-0,514	0,141	0,95	0,92	0,33	0,67	1
BC3	5	0,250	0,134	1,02	1,02	0,21	0,45	2
BC4	1	-0,514	0,141	0,92	0,90	0,39	0,67	1
BC4	2	0,678	0,141	1,00	1,00	0,23	0,33	2
BC4	3	0,293	0,134	1,00	0,99	0,25	0,44	2
BC4	4	0,465	0,137	0,97	0,96	0,30	0,39	2
BC4	5	0,824	0,146	1,01	1,05	0,20	0,29	2
B01	1	0,463	0,192	0,94	0,93	0,36	0,39	2
B01	2	-0,238	0,191	0,94	0,92	0,36	0,60	2
B01	3	0,672	0,198	1,04	1,05	0,19	0,34	2
B01	4	0,014	0,188	0,91	0,89	0,42	0,52	2
B01	5	0,377	0,190	1,05	1,06	0,17	0,42	2
B01	6	0,703	0,199	1,10	1,12	0,08	0,33	2
B02	1	0,865	0,207	1,04	1,05	0,17	0,29	2
B02	2	0,463	0,192	0,91	0,90	0,41	0,39	2
B02	3	0,641	0,197	1,05	1,05	0,17	0,34	2
B02	4	0,611	0,196	0,97	0,97	0,31	0,35	2
B02	5	0,098	0,187	0,93	0,93	0,37	0,50	2
B02	6	0,236	0,188	0,96	0,95	0,33	0,46	2
B03	1	0,608	0,202	0,97	0,95	0,29	0,34	2
B03	2	0,272	0,194	1,00	0,99	0,24	0,44	2
B03	3	0,674	0,205	1,06	1,19	0,03	0,32	2
B03	4	0,577	0,201	1,04	1,06	0,12	0,35	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	5	0,452	0,197	1,05	1,06	0,11	0,39	2
B03	6	0,845	0,213	1,04	1,11	0,10	0,28	2
B04	1	1,033	0,225	1,08	1,12	0,01	0,24	2
B04	2	0,391	0,196	1,02	1,03	0,18	0,40	2
B04	3	0,845	0,213	1,00	1,08	0,18	0,28	2
B04	4	0,918	0,217	0,97	0,96	0,26	0,26	2
B04	5	0,741	0,208	1,00	1,04	0,20	0,31	2
B04	6	0,674	0,205	0,97	0,96	0,27	0,32	2
B05	1	0,463	0,192	0,99	0,99	0,28	0,39	2
B05	2	0,002	0,189	0,98	0,96	0,30	0,53	2
B05	3	0,483	0,193	1,02	1,03	0,22	0,39	2
B05	4	0,734	0,201	1,08	1,11	0,11	0,32	2
B05	5	0,349	0,190	0,97	0,98	0,30	0,43	2
B05	6	0,641	0,197	1,06	1,06	0,16	0,34	2
B06	1	0,581	0,195	1,06	1,05	0,16	0,36	2
B07	2	0,703	0,199	1,04	1,04	0,18	0,33	2
B06	3	0,767	0,202	1,08	1,11	0,10	0,31	2
B06	4	0,581	0,195	1,04	1,03	0,19	0,36	2
B06	5	0,581	0,195	1,08	1,11	0,11	0,36	2
B06	6	0,826	0,205	1,07	1,11	0,11	0,30	2
B07	1	0,468	0,195	1,04	1,05	0,14	0,38	2
B07	2	0,498	0,195	0,96	0,97	0,30	0,37	2
B07	3	0,815	0,208	1,04	1,06	0,12	0,29	2
B07	4	0,066	0,189	1,09	1,09	0,06	0,50	2
B07	5	0,095	0,189	1,06	1,06	0,12	0,49	2
B07	6	0,815	0,208	1,08	1,14	0,03	0,29	2
B08	1	-0,133	0,191	0,92	0,91	0,40	0,56	2
B08	2	0,684	0,202	1,05	1,10	0,11	0,32	2
B08	3	0,235	0,190	0,98	0,96	0,29	0,45	2
B08	4	0,528	0,196	0,98	1,00	0,25	0,36	2
B08	5	0,468	0,195	1,07	1,08	0,08	0,38	2
B08	6	0,409	0,193	0,98	0,99	0,26	0,40	2

Tabla 4.4 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura en ashaninka de 4.º grado de primaria EIB

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	1	-2,139	0,084	0,93	0,76	0,36	0,87	1
BC1	2	-1,612	0,072	0,93	0,87	0,36	0,81	1
BC1	3	-1,243	0,066	0,93	0,86	0,38	0,75	1
BC1	4	-0,996	0,062	0,90	0,85	0,43	0,70	1
BC1	5	-0,908	0,062	0,90	0,85	0,43	0,69	1
BC2	1	-1,262	0,066	0,86	0,77	0,48	0,75	1
BC2	2	-0,354	0,058	0,94	0,92	0,38	0,57	2
BC2	3	0,665	0,060	1,02	1,02	0,25	0,35	2
BC2	4	-0,503	0,059	0,91	0,88	0,43	0,60	1
BC2	5	0,550	0,059	1,05	1,08	0,20	0,38	2
BC3	1	-2,561	0,097	0,96	0,76	0,29	0,91	1
BC3	2	-1,470	0,069	0,92	0,83	0,39	0,78	1
BC3	3	-1,149	0,064	0,87	0,79	0,48	0,73	1
BC3	4	0,255	0,058	0,98	0,98	0,32	0,44	2
BC3	5	-0,553	0,059	1,02	1,03	0,24	0,62	1
BC3	1	-0,085	0,057	1,01	1,01	0,27	0,51	2
BC4	2	1,322	0,067	1,10	1,17	0,09	0,23	3
BC4	3	0,789	0,061	1,02	1,03	0,24	0,33	3
BC4	4	0,345	0,058	1,09	1,10	0,15	0,42	2
BC4	5	0,872	0,062	1,05	1,08	0,18	0,31	3
B01	1	0,246	0,082	0,98	0,97	0,33	0,44	2
B01	2	-0,198	0,082	0,90	0,88	0,44	0,54	2
B01	3	1,306	0,095	1,10	1,25	0,08	0,24	3
B01	4	0,432	0,083	0,95	0,94	0,36	0,40	2
B01	5	1,064	0,090	1,10	1,17	0,11	0,28	3
B01	6	0,243	0,082	1,13	1,18	0,10	0,44	2
B02	1	0,201	0,082	0,99	0,98	0,31	0,45	2
B02	2	0,240	0,082	1,05	1,07	0,21	0,44	2
B02	3	0,633	0,085	1,03	1,04	0,24	0,36	2
B02	4	0,084	0,082	0,96	0,95	0,35	0,48	2
B02	5	1,067	0,090	1,06	1,10	0,18	0,28	3
B02	6	0,672	0,085	1,07	1,09	0,18	0,35	2
B03	1	0,248	0,081	0,95	0,97	0,35	0,44	2
B03	2	0,282	0,081	1,05	1,05	0,20	0,43	2
B03	3	0,979	0,089	1,13	1,18	0,05	0,29	3
B03	4	0,356	0,082	0,99	0,99	0,29	0,41	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	5	0,758	0,086	1,03	1,04	0,21	0,33	2
B03	6	0,708	0,085	1,07	1,08	0,16	0,34	2
B04	1	0,784	0,086	0,96	0,96	0,33	0,32	3
B04	2	0,928	0,088	1,01	1,02	0,25	0,30	3
B04	3	0,065	0,081	0,96	0,95	0,34	0,48	2
B04	4	0,728	0,085	1,07	1,09	0,16	0,34	2
B04	5	0,912	0,088	1,03	1,06	0,20	0,30	3
B04	6	0,969	0,089	1,16	1,21	0,00	0,29	3
B05	1	0,308	0,082	0,91	0,90	0,42	0,43	2
B05	2	0,558	0,084	0,93	0,96	0,38	0,38	2
B05	3	1,259	0,094	1,12	1,27	0,07	0,24	3
B05	4	-0,518	0,083	0,93	0,91	0,39	0,61	1
B05	5	0,739	0,086	1,18	1,28	0,00	0,34	2
B05	6	0,770	0,086	1,01	1,02	0,27	0,33	3
B06	1	0,069	0,082	1,06	1,07	0,21	0,48	2
B06	2	0,368	0,083	0,96	0,96	0,35	0,42	2
B06	3	0,870	0,087	1,10	1,14	0,13	0,31	3
B06	4	0,510	0,084	1,05	1,05	0,23	0,39	2
B06	5	0,744	0,086	1,01	1,02	0,26	0,34	2
B06	6	0,935	0,088	0,97	0,97	0,32	0,30	3
B07	1	0,531	0,083	0,99	1,00	0,28	0,38	2
B07	2	0,624	0,084	0,99	1,01	0,27	0,36	2
B07	3	0,524	0,083	0,98	1,01	0,29	0,38	2
B07	4	0,153	0,081	1,00	1,01	0,27	0,46	2
B07	5	0,314	0,081	0,97	0,97	0,33	0,42	2
B07	6	0,241	0,081	1,03	1,03	0,23	0,44	2
B08	1	0,917	0,088	1,03	1,06	0,21	0,30	3
B08	2	-0,034	0,080	1,07	1,07	0,17	0,50	2
B08	3	0,721	0,085	1,06	1,11	0,15	0,34	2
B08	4	0,819	0,086	0,95	0,94	0,34	0,32	3
B08	5	0,943	0,088	1,14	1,19	0,03	0,29	3
B08	6	0,691	0,084	0,97	0,98	0,31	0,34	2

Tabla 4.5 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura en awajún de 4.º grado de primaria EIB

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	1	-2,354	0,051	0,96	0,90	0,30	0,82	1
BC1	2	-1,217	0,043	0,91	0,85	0,40	0,71	1
BC1	3	-0,505	0,041	0,92	0,87	0,42	0,62	1
BC1	4	0,458	0,040	0,94	0,90	0,42	0,49	2
BC1	5	1,103	0,041	1,01	1,01	0,34	0,41	3
BC2	1	-0,070	0,040	0,87	0,83	0,48	0,57	2
BC2	2	1,290	0,041	0,98	0,97	0,37	0,39	3
BC2	3	0,443	0,040	0,91	0,88	0,44	0,50	2
BC2	4	1,328	0,041	0,91	0,90	0,45	0,38	3
BC2	5	0,728	0,040	1,00	0,99	0,34	0,46	2
BC3	1	-2,005	0,047	0,93	0,87	0,36	0,79	1
BC3	2	-1,090	0,042	0,93	0,85	0,39	0,69	1
BC3	3	-0,788	0,041	0,90	0,86	0,42	0,66	1
BC3	4	-0,355	0,040	0,92	0,88	0,42	0,60	1
BC3	5	0,577	0,040	1,14	1,17	0,17	0,48	2
BC4	1	0,134	0,040	0,93	0,90	0,42	0,54	2
BC4	2	0,143	0,040	0,90	0,86	0,45	0,53	2
BC4	3	0,071	0,040	0,99	1,00	0,34	0,54	2
BC4	4	0,880	0,040	0,96	0,94	0,39	0,44	2
BC4	5	0,165	0,040	0,98	0,95	0,36	0,53	2
B01	1	1,651	0,060	1,02	1,03	0,34	0,34	3
B01	2	1,294	0,059	1,12	1,14	0,24	0,39	3
B01	3	1,371	0,059	1,10	1,14	0,25	0,38	3
B01	4	1,140	0,058	0,96	0,95	0,41	0,41	3
B01	5	1,266	0,058	0,96	0,93	0,41	0,39	3
B01	6	0,704	0,057	1,09	1,11	0,26	0,46	2
B02	1	1,779	0,061	0,94	0,92	0,43	0,33	3
B02	2	0,094	0,057	0,88	0,84	0,48	0,54	2
B02	3	0,647	0,057	1,02	1,01	0,34	0,47	2
B02	4	2,070	0,063	1,25	1,31	0,11	0,30	3
B02	5	1,398	0,059	0,98	0,96	0,39	0,37	3
B02	6	1,779	0,061	1,07	1,07	0,30	0,33	3
B03	1	1,128	0,058	1,18	1,21	0,11	0,41	3
B03	2	1,322	0,058	1,03	1,03	0,29	0,38	3
B03	3	0,716	0,057	1,07	1,07	0,24	0,46	2
B03	4	0,872	0,057	1,08	1,08	0,23	0,44	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	5	2,340	0,064	1,08	1,11	0,22	0,26	3
B03	6	2,338	0,064	1,05	1,07	0,27	0,26	3
B04	1	0,764	0,057	0,98	0,98	0,35	0,45	2
B04	2	0,746	0,057	1,02	1,02	0,30	0,46	2
B04	3	2,346	0,064	1,11	1,13	0,19	0,26	3
B04	4	2,551	0,066	1,11	1,16	0,19	0,24	3
B04	5	2,075	0,062	0,98	0,98	0,35	0,29	3
B04	6	1,818	0,061	0,94	0,91	0,41	0,32	3
B05	1	1,736	0,060	0,95	0,95	0,42	0,33	3
B05	2	0,580	0,056	0,93	0,89	0,44	0,48	2
B05	3	1,179	0,057	1,13	1,14	0,23	0,40	3
B05	4	1,055	0,057	1,05	1,13	0,29	0,41	3
B05	5	1,717	0,060	1,09	1,10	0,28	0,33	3
B05	6	1,616	0,059	1,03	1,07	0,32	0,35	3
B06	1	0,650	0,056	1,15	1,15	0,20	0,47	2
B06	2	1,083	0,057	1,01	0,99	0,36	0,41	3
B06	3	1,636	0,059	1,17	1,20	0,19	0,34	3
B06	4	1,280	0,058	1,00	0,99	0,37	0,39	3
B06	5	0,903	0,057	1,08	1,11	0,28	0,43	2
B06	6	2,121	0,062	1,13	1,20	0,22	0,29	3
B07	1	1,179	0,057	1,03	1,04	0,29	0,40	3
B07	2	0,983	0,057	1,11	1,13	0,19	0,42	3
B07	3	1,213	0,057	1,04	1,04	0,28	0,39	3
B07	4	1,227	0,057	1,10	1,12	0,20	0,39	3
B07	5	1,186	0,057	1,02	1,12	0,28	0,40	3
B07	6	1,754	0,060	1,02	1,04	0,30	0,33	3
B08	1	0,935	0,057	1,07	1,08	0,23	0,43	3
B08	2	1,638	0,059	0,98	0,96	0,35	0,34	3
B08	3	1,659	0,059	1,05	1,05	0,27	0,34	3
B08	4	2,042	0,062	1,02	1,05	0,29	0,29	3
B08	5	1,014	0,057	1,08	1,09	0,22	0,42	3
B08	6	1,812	0,060	0,95	0,93	0,39	0,32	3

Tabla 4.6 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura en quechua chanca de 4.º grado de primaria EIB

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	1	-1,978	0,067	0,96	0,76	0,32	0,90	1
BC1	2	-1,557	0,058	0,93	0,79	0,36	0,87	1
BC1	3	-1,146	0,051	0,94	0,85	0,34	0,82	1
BC1	4	-1,323	0,054	0,94	0,82	0,34	0,84	1
BC1	5	0,803	0,040	1,09	1,11	0,14	0,45	2
BC2	1	0,244	0,040	0,93	0,91	0,38	0,57	2
BC2	2	1,359	0,042	1,03	1,03	0,23	0,33	3
BC2	3	1,446	0,043	1,08	1,10	0,15	0,32	3
BC2	4	0,977	0,040	1,02	1,03	0,24	0,41	2
BC2	5	0,695	0,040	1,03	1,03	0,24	0,47	2
BC3	1	-2,527	0,082	0,95	0,72	0,30	0,94	1
BC3	2	-1,920	0,065	0,95	0,81	0,31	0,90	1
BC3	3	0,562	0,040	0,92	0,90	0,40	0,50	2
BC3	4	-0,085	0,041	0,90	0,87	0,42	0,64	2
BC3	5	0,057	0,041	0,93	0,91	0,38	0,61	2
BC4	1	1,510	0,043	1,16	1,21	0,03	0,31	3
BC4	2	1,313	0,042	1,05	1,06	0,20	0,34	2
BC4	3	1,192	0,041	1,03	1,06	0,22	0,37	2
BC4	4	0,782	0,040	1,03	1,02	0,24	0,45	2
BC4	5	1,349	0,042	1,01	1,02	0,25	0,34	2
B01	1	0,305	0,057	0,99	1,00	0,29	0,55	2
B01	2	1,440	0,060	1,11	1,14	0,11	0,32	3
B01	3	0,719	0,057	1,01	1,01	0,27	0,46	2
B01	4	0,166	0,057	1,12	1,21	0,08	0,58	2
B01	5	1,224	0,059	0,97	0,97	0,32	0,36	2
B01	6	1,628	0,062	1,19	1,31	-0,02	0,28	3
B02	1	0,726	0,057	0,98	0,98	0,31	0,46	2
B02	2	0,975	0,057	0,94	0,93	0,37	0,41	2
B02	3	1,733	0,064	0,95	0,94	0,33	0,26	3
B02	4	0,872	0,057	0,93	0,91	0,39	0,43	2
B02	5	1,832	0,065	1,08	1,14	0,13	0,25	3
B02	6	1,647	0,063	1,05	1,09	0,18	0,28	3
B03	1	0,008	0,058	0,89	0,85	0,44	0,62	2
B03	2	0,259	0,057	0,91	0,89	0,42	0,56	2
B03	3	0,958	0,057	0,94	0,94	0,37	0,41	2
B03	4	0,349	0,056	0,96	0,96	0,34	0,54	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	5	0,497	0,056	0,94	0,93	0,37	0,51	2
B03	6	1,535	0,061	1,19	1,27	-0,02	0,30	3
B04	1	0,631	0,056	0,95	0,94	0,36	0,48	2
B04	2	0,456	0,056	0,94	0,93	0,37	0,52	2
B04	3	1,419	0,060	0,91	0,90	0,40	0,32	3
B04	4	1,393	0,060	1,07	1,12	0,16	0,32	3
B04	5	0,674	0,056	0,97	0,96	0,33	0,47	2
B04	6	1,023	0,057	1,01	1,01	0,27	0,40	2
B05	1	0,267	0,057	0,96	0,94	0,33	0,57	2
B05	2	-0,243	0,060	0,91	0,86	0,40	0,67	1
B05	3	1,211	0,058	1,04	1,06	0,20	0,37	2
B05	4	0,418	0,056	0,92	0,90	0,39	0,54	2
B05	5	0,891	0,057	0,99	1,00	0,28	0,43	2
B05	6	1,727	0,063	1,06	1,08	0,17	0,27	3
B06	1	1,150	0,058	1,10	1,12	0,12	0,38	2
B06	2	1,432	0,060	1,05	1,06	0,18	0,32	3
B06	3	1,272	0,059	1,00	1,00	0,27	0,36	2
B06	4	0,952	0,057	0,98	0,98	0,29	0,42	2
B06	5	0,681	0,056	1,00	1,00	0,27	0,48	2
B06	6	1,643	0,062	1,03	1,03	0,20	0,29	3
B07	1	0,678	0,056	0,96	0,96	0,32	0,48	2
B07	2	1,165	0,058	1,01	1,02	0,24	0,38	2
B07	3	1,559	0,061	1,12	1,14	0,09	0,30	3
B07	4	1,125	0,058	1,11	1,13	0,11	0,39	2
B07	5	1,061	0,057	1,01	1,02	0,25	0,40	2
B07	6	0,906	0,057	0,95	0,94	0,34	0,43	2
B08	1	0,222	0,057	0,98	0,97	0,30	0,58	2
B08	2	0,508	0,056	0,95	0,94	0,34	0,52	2
B08	3	1,009	0,057	0,96	0,95	0,32	0,41	2
B08	4	1,534	0,061	0,99	1,00	0,27	0,31	3
B08	5	0,785	0,056	1,03	1,03	0,23	0,46	2
B08	6	1,020	0,057	1,09	1,11	0,14	0,41	2

Tabla 4.7 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura en quechua cusco collao de 4.º grado de primaria EIB

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	1	-2,985	0,073	0,94	0,79	0,27	0,94	1
BC1	2	-2,272	0,059	0,93	0,72	0,32	0,90	1
BC1	3	-1,332	0,046	0,95	0,87	0,32	0,82	1
BC1	4	-0,657	0,041	0,93	0,88	0,37	0,74	1
BC1	5	1,988	0,039	1,05	1,09	0,24	0,32	3
BC2	1	1,467	0,037	1,06	1,07	0,23	0,40	3
BC2	2	0,841	0,036	0,98	0,97	0,34	0,50	2
BC2	3	1,405	0,037	1,12	1,17	0,16	0,41	3
BC2	4	1,780	0,038	1,07	1,09	0,22	0,35	3
BC2	5	2,267	0,040	1,13	1,22	0,11	0,28	3
BC3	1	-3,013	0,073	0,93	0,63	0,31	0,94	1
BC3	2	-2,608	0,064	0,94	0,71	0,31	0,92	1
BC3	3	0,823	0,036	1,00	1,00	0,31	0,51	2
BC3	4	1,389	0,037	1,01	1,01	0,30	0,41	3
BC3	5	0,287	0,037	1,00	1,00	0,30	0,60	2
BC4	1	-0,487	0,039	0,93	0,88	0,38	0,72	1
BC4	2	0,622	0,036	0,95	0,93	0,37	0,54	2
BC4	3	1,166	0,036	0,94	0,94	0,39	0,45	2
BC4	4	0,544	0,036	0,99	0,98	0,33	0,55	2
BC4	5	1,324	0,036	1,08	1,10	0,21	0,42	3
B01	1	0,324	0,051	1,00	0,99	0,29	0,59	2
B01	2	1,554	0,052	1,02	1,03	0,28	0,38	3
B01	3	1,469	0,052	1,05	1,06	0,23	0,39	3
B01	4	1,178	0,051	1,03	1,04	0,26	0,44	2
B01	5	1,370	0,052	1,02	1,02	0,27	0,41	3
B01	6	2,238	0,057	1,09	1,14	0,17	0,28	3
B02	1	1,580	0,052	1,04	1,04	0,25	0,38	3
B02	2	0,833	0,051	0,97	0,95	0,35	0,50	2
B02	3	0,731	0,051	1,02	1,00	0,28	0,52	2
B02	4	1,634	0,053	1,07	1,09	0,21	0,37	3
B02	5	1,493	0,052	0,95	0,93	0,38	0,39	3
B02	6	2,369	0,058	1,07	1,11	0,18	0,26	3
B03	1	-1,362	0,066	0,91	0,76	0,39	0,82	1
B03	2	-0,466	0,056	0,88	0,81	0,45	0,71	1
B03	3	0,317	0,052	0,96	0,95	0,37	0,59	2
B03	4	-0,512	0,057	0,90	0,82	0,42	0,72	1

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	5	-0,232	0,055	0,90	0,84	0,43	0,68	1
B03	6	0,723	0,051	0,97	0,96	0,36	0,53	2
B04	1	1,087	0,051	1,11	1,14	0,18	0,46	2
B04	2	1,953	0,055	1,11	1,14	0,18	0,32	3
B04	3	1,192	0,052	1,01	1,02	0,31	0,45	3
B04	4	1,643	0,053	1,01	1,02	0,31	0,37	3
B04	5	1,094	0,052	1,00	1,00	0,32	0,46	2
B04	6	2,489	0,059	1,11	1,38	0,15	0,25	3
B05	1	0,769	0,051	1,01	1,01	0,28	0,51	2
B05	2	0,056	0,052	0,92	0,88	0,40	0,63	2
B05	3	1,193	0,051	0,92	0,90	0,41	0,44	3
B05	4	0,295	0,051	0,87	0,83	0,47	0,59	2
B05	5	1,183	0,051	0,95	0,95	0,36	0,44	2
B05	6	2,061	0,055	1,01	1,03	0,28	0,30	3
B06	1	1,052	0,051	0,92	0,90	0,41	0,46	2
B06	2	1,656	0,053	1,05	1,08	0,23	0,36	3
B06	3	1,569	0,052	0,98	0,99	0,32	0,38	3
B06	4	1,498	0,052	1,04	1,04	0,25	0,39	3
B06	5	1,986	0,054	1,05	1,11	0,21	0,31	3
B06	6	1,119	0,051	1,10	1,13	0,16	0,45	2
B07	1	0,885	0,051	0,93	0,92	0,41	0,50	2
B07	2	1,067	0,051	0,97	0,97	0,36	0,47	2
B07	3	0,936	0,051	0,94	0,92	0,41	0,49	2
B07	4	1,385	0,052	0,95	0,95	0,38	0,41	3
B07	5	0,816	0,051	0,88	0,86	0,47	0,51	2
B07	6	1,917	0,054	1,09	1,13	0,20	0,33	3
B08	1	1,071	0,051	0,98	0,97	0,36	0,47	2
B08	2	1,217	0,052	1,01	1,01	0,31	0,44	3
B08	3	0,575	0,051	0,91	0,88	0,44	0,55	2
B08	4	2,032	0,055	1,19	1,28	0,07	0,31	3
B08	5	1,439	0,052	1,02	1,03	0,30	0,40	3
B08	6	2,096	0,056	1,14	1,19	0,13	0,30	3

Tabla 4.8 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura en shipibo-konibo de 4.º grado de primaria EIB

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	1	-1,698	0,078	1,00	1,01	0,32	0,77	1
BC1	2	-0,986	0,070	0,97	0,89	0,40	0,64	1
BC1	3	-0,870	0,069	0,96	0,90	0,41	0,62	1
BC1	4	-0,114	0,068	1,07	1,28	0,32	0,47	1
BC1	5	0,067	0,069	1,08	1,14	0,33	0,43	2
BC2	1	0,137	0,069	0,87	0,82	0,53	0,42	2
BC2	2	0,262	0,070	0,89	0,85	0,51	0,39	2
BC2	3	0,239	0,070	0,92	0,88	0,48	0,40	2
BC2	4	0,306	0,070	1,04	1,05	0,37	0,38	2
BC2	5	0,813	0,075	0,98	0,96	0,43	0,29	3
BC3	1	-1,885	0,081	0,97	0,98	0,35	0,80	1
BC3	2	-1,459	0,074	0,90	0,80	0,44	0,73	1
BC3	3	-1,321	0,073	0,96	0,86	0,39	0,71	1
BC3	4	0,041	0,068	0,94	0,93	0,45	0,44	2
BC3	5	-0,524	0,067	0,89	0,83	0,49	0,55	1
BC3	1	-0,261	0,067	0,96	0,93	0,43	0,50	1
BC4	2	-0,620	0,068	0,97	0,92	0,42	0,57	1
BC4	3	-0,182	0,068	0,90	0,88	0,48	0,48	1
BC4	4	-0,499	0,067	0,88	0,83	0,50	0,55	1
BC4	5	-0,266	0,067	0,89	0,86	0,50	0,50	1
B01	1	0,472	0,100	1,02	1,02	0,37	0,36	2
B01	2	0,526	0,100	1,05	1,05	0,34	0,34	2
B01	3	0,002	0,096	1,01	0,99	0,38	0,45	2
B01	4	0,049	0,096	1,07	1,07	0,31	0,44	2
B01	5	0,839	0,105	1,22	1,29	0,18	0,29	3
B01	6	1,421	0,119	1,15	1,21	0,24	0,20	3
B02	1	1,350	0,117	0,93	0,87	0,45	0,21	3
B02	2	0,850	0,106	1,16	1,22	0,23	0,29	3
B02	3	0,858	0,106	1,00	1,02	0,38	0,29	3
B02	4	1,171	0,113	1,02	1,01	0,37	0,24	3
B02	5	0,620	0,102	1,08	1,14	0,30	0,33	2
B02	6	1,174	0,113	1,10	1,11	0,29	0,24	3
B03	1	0,260	0,100	1,05	1,15	0,38	0,39	2
B03	2	-0,294	0,097	1,05	1,17	0,36	0,50	1
B03	3	0,445	0,102	1,09	1,08	0,37	0,36	2
B03	4	0,022	0,098	1,02	0,95	0,42	0,44	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	5	0,081	0,098	1,03	1,02	0,41	0,43	2
B03	6	-0,228	0,097	0,95	0,94	0,46	0,49	1
B04	1	0,300	0,100	1,04	1,01	0,41	0,38	2
B04	2	0,719	0,106	0,90	0,91	0,51	0,31	3
B04	3	0,968	0,110	1,17	1,16	0,31	0,27	3
B04	4	0,703	0,106	1,00	0,99	0,44	0,31	3
B04	5	1,823	0,133	1,04	1,06	0,37	0,16	3
B04	6	0,453	0,102	1,01	1,02	0,43	0,36	2
B05	1	0,367	0,098	0,90	0,89	0,47	0,37	2
B05	2	0,189	0,096	1,08	1,10	0,31	0,41	2
B05	3	0,180	0,096	1,11	1,13	0,28	0,41	2
B05	4	0,898	0,106	1,19	1,23	0,22	0,28	3
B05	5	-0,062	0,095	1,00	0,97	0,38	0,46	2
B05	6	0,067	0,096	1,09	1,08	0,30	0,43	2
B06	1	0,492	0,100	0,96	0,96	0,42	0,35	2
B06	2	-	-	-	-	-	-	-
B06	3	0,170	0,096	1,07	1,07	0,32	0,41	2
B06	4	0,564	0,101	1,01	0,98	0,38	0,33	2
B06	5	0,199	0,097	1,16	1,19	0,23	0,40	2
B06	6	0,161	0,096	1,00	1,02	0,38	0,41	2
B07	1	0,424	0,101	0,92	0,88	0,50	0,36	2
B07	2	0,112	0,098	1,10	1,16	0,35	0,42	2
B07	3	0,025	0,097	1,11	1,13	0,34	0,44	2
B07	4	0,034	0,097	1,16	1,18	0,30	0,44	2
B07	5	0,816	0,107	0,91	0,87	0,51	0,30	3
B07	6	-0,167	0,096	1,21	1,47	0,22	0,48	1
B08	1	0,128	0,098	0,87	0,82	0,54	0,42	2
B08	2	0,329	0,100	1,05	1,04	0,39	0,38	2
B08	3	0,080	0,098	1,06	1,04	0,39	0,43	2
B08	4	0,178	0,099	0,95	0,92	0,47	0,41	2
B08	5	-	-	-	-	-	-	-
B08	6	-0,258	0,096	1,09	1,05	0,36	0,50	1

Tabla 4.9 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura en castellano como segunda lengua de 4.º grado de primaria EIB

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	1	-1,570	0,021	0,93	0,79	0,37	0,84	1
BC1	2	-1,252	0,019	0,86	0,70	0,45	0,79	1
BC1	3	-1,440	0,020	0,86	0,68	0,43	0,82	1
BC1	4	-0,622	0,017	0,97	1,01	0,36	0,68	1
BC1	5	-0,282	0,016	0,92	0,87	0,44	0,62	1
BC2	1	0,823	0,017	1,09	1,13	0,32	0,38	2
BC2	2	0,034	0,016	0,92	0,87	0,46	0,55	2
BC2	3	-0,183	0,016	0,92	0,87	0,45	0,59	1
BC2	4	0,795	0,017	1,10	1,15	0,31	0,39	2
BC2	5	0,715	0,017	1,04	1,06	0,37	0,40	2
BC3	1	-2,138	0,025	0,92	0,75	0,34	0,90	1
BC3	2	-1,563	0,021	0,88	0,67	0,41	0,84	1
BC3	3	-1,351	0,020	0,87	0,69	0,44	0,81	1
BC3	4	-1,464	0,020	0,88	0,70	0,42	0,82	1
BC3	5	0,438	0,016	0,94	0,93	0,45	0,46	2
BC4	1	0,405	0,016	0,93	0,91	0,46	0,47	2
BC4	2	0,309	0,016	0,99	0,97	0,41	0,49	2
BC4	3	0,942	0,017	1,06	1,11	0,34	0,36	2
BC4	4	0,835	0,017	1,09	1,14	0,32	0,38	2
BC4	5	1,193	0,018	1,24	1,35	0,19	0,31	3
B01	1	-0,061	0,023	1,10	1,23	0,30	0,57	1
B01	2	1,023	0,024	0,97	0,99	0,45	0,34	2
B01	3	0,324	0,023	1,06	1,09	0,36	0,49	2
B01	4	0,444	0,023	0,95	0,94	0,46	0,46	2
B01	5	0,349	0,023	1,03	1,02	0,39	0,48	2
B01	6	0,805	0,024	1,11	1,17	0,33	0,39	2
B02	1	0,122	0,023	0,89	0,84	0,50	0,53	2
B02	2	0,010	0,023	0,93	0,90	0,45	0,55	2
B02	3	0,642	0,023	1,00	1,00	0,42	0,42	2
B02	4	0,766	0,024	1,06	1,07	0,37	0,40	2
B02	5	0,517	0,023	1,13	1,16	0,31	0,45	2
B02	6	0,635	0,023	0,96	0,94	0,45	0,42	2
B03	1	1,916	0,029	1,02	1,07	0,34	0,19	3
B03	2	-0,138	0,023	0,99	0,95	0,38	0,59	1
B03	3	0,068	0,023	0,90	0,86	0,47	0,54	2
B03	4	0,361	0,023	0,99	0,98	0,39	0,48	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	5	0,635	0,023	1,05	1,06	0,33	0,42	2
B03	6	0,983	0,024	1,21	1,29	0,17	0,34	2
B04	1	0,710	0,023	1,06	1,11	0,32	0,40	2
B04	2	0,396	0,023	0,97	0,95	0,41	0,47	2
B04	3	0,538	0,023	0,93	0,91	0,45	0,44	2
B04	4	0,992	0,024	1,04	1,07	0,34	0,34	2
B04	5	0,400	0,023	1,07	1,06	0,31	0,47	2
B04	6	0,325	0,023	1,05	1,08	0,32	0,48	2
B05	1	-0,169	0,023	0,93	0,86	0,45	0,59	1
B05	2	0,126	0,023	0,94	0,91	0,45	0,53	2
B05	3	-0,149	0,023	0,89	0,82	0,48	0,59	1
B05	4	0,602	0,023	1,01	1,03	0,41	0,43	2
B05	5	-0,010	0,023	1,03	1,04	0,38	0,56	2
B05	6	0,890	0,024	1,15	1,22	0,30	0,37	2
B06	1	-0,242	0,023	0,97	0,92	0,41	0,61	1
B06	2	0,600	0,023	1,12	1,21	0,31	0,43	2
B06	3	0,382	0,023	0,95	0,94	0,45	0,47	2
B06	4	0,334	0,023	1,01	1,01	0,40	0,48	2
B06	5	-	-	-	-	-	-	-
B06	6	0,390	0,023	1,01	1,00	0,41	0,47	2
B07	1	0,680	0,023	1,15	1,20	0,23	0,41	2
B07	2	0,498	0,023	1,08	1,09	0,30	0,45	2
B07	3	0,598	0,023	0,96	0,96	0,42	0,42	2
B07	4	-0,309	0,023	0,97	0,94	0,39	0,62	1
B07	5	0,682	0,023	0,93	0,92	0,45	0,41	2
B07	6	1,407	0,026	1,29	1,45	0,09	0,27	3
B08	1	-0,485	0,023	0,94	0,91	0,40	0,66	1
B08	2	0,543	0,023	1,03	1,06	0,35	0,44	2
B08	3	0,877	0,024	1,00	1,02	0,38	0,36	2
B08	4	0,924	0,024	1,12	1,18	0,26	0,36	2
B08	5	0,663	0,023	1,11	1,15	0,27	0,41	2
B08	6	0,338	0,023	0,98	0,98	0,40	0,48	2

Tabla 4.10 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Lectura de 2.º grado de secundaria

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	1	0,064	0,006	0,92	0,81	0,46	0,76	2
B01	2	1,677	0,005	1,15	1,23	0,29	0,42	3
B01	3	0,137	0,006	0,99	0,94	0,40	0,74	2
B01	4	0,528	0,005	0,93	0,86	0,48	0,67	2
B01	5	1,481	0,005	1,25	1,32	0,22	0,46	3
B01	6	1,129	0,005	0,92	0,89	0,50	0,54	3
B02	1	1,611	0,005	1,22	1,31	0,24	0,44	3
B02	2	0,457	0,005	0,91	0,82	0,49	0,68	2
B02	3	1,521	0,005	0,88	0,85	0,54	0,46	3
B02	4	-0,384	0,006	0,92	0,79	0,42	0,83	1
B02	5	1,001	0,005	0,97	0,97	0,45	0,57	3
B02	6	0,645	0,005	0,83	0,76	0,56	0,65	2
B03	1	2,397	0,005	0,99	1,07	0,39	0,27	4
B03	2_1	-0,323	0,005	1,13	1,11	0,17	0,71	1
B03	2_2	3,496	0,005	1,13	1,11	0,12	0,11	4
B03	3	1,315	0,005	1,04	1,06	0,39	0,50	3
B03	4	0,581	0,005	0,91	0,86	0,49	0,66	2
B03	5	1,012	0,005	0,86	0,81	0,55	0,56	3
B03	6	0,332	0,005	0,92	0,89	0,47	0,71	2
B04	1	1,011	0,005	0,96	0,93	0,46	0,57	3
B04	2	0,232	0,005	0,95	0,97	0,42	0,73	2
B04	3	1,679	0,005	0,98	0,99	0,44	0,42	3
B04	4	0,795	0,005	1,16	1,21	0,28	0,62	2
B04	5	-0,078	0,006	0,94	0,93	0,41	0,78	2
B04	6	0,922	0,005	1,07	1,10	0,36	0,59	2
B05	1	1,139	0,005	1,14	1,18	0,31	0,54	3
B05	2	-0,585	0,007	0,88	0,68	0,44	0,85	1
B05	3	1,026	0,005	0,86	0,83	0,54	0,56	3
B05	4	1,286	0,005	0,98	0,96	0,45	0,51	3
B05	5	1,790	0,005	0,94	0,99	0,47	0,40	4
B05	6	1,575	0,005	0,93	0,91	0,50	0,44	3
B06	1	-0,032	0,006	0,97	0,89	0,41	0,77	2
B06	2	1,758	0,005	1,12	1,19	0,32	0,40	3
B06	3	1,022	0,005	1,06	1,08	0,38	0,57	3
B06	4	1,257	0,005	1,13	1,19	0,31	0,51	3
B06	5	2,285	0,005	0,95	0,99	0,44	0,30	4

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B06	6	1,374	0,005	0,96	0,96	0,46	0,49	3
B07	1	-1,360	0,009	0,95	0,84	0,30	0,92	1
B07	2	-0,495	0,006	0,95	0,92	0,37	0,84	1
B07	3	-0,429	0,006	0,89	0,75	0,44	0,83	1
B07	4	0,219	0,005	0,92	0,85	0,46	0,73	2
B07	5	-0,147	0,006	0,96	0,91	0,40	0,79	1
B07	6	0,677	0,005	0,95	0,93	0,46	0,64	2
B08	1	-1,505	0,009	0,95	0,78	0,29	0,93	1
B08	2	-0,039	0,006	0,99	0,94	0,38	0,77	2
B08	3	0,960	0,005	1,00	1,00	0,43	0,57	3
B08	4	1,976	0,005	1,10	1,18	0,33	0,35	4
B08	5	1,318	0,005	0,95	0,93	0,48	0,49	3
B08	6	0,314	0,005	0,94	0,87	0,45	0,71	2
B09	1	0,729	0,005	1,23	1,43	0,19	0,62	2
B09	2	0,039	0,006	0,88	0,78	0,48	0,76	2
B09	3	1,181	0,005	1,04	1,06	0,39	0,52	3
B09	4	0,760	0,005	0,88	0,84	0,52	0,62	2
B09	5	1,146	0,005	0,94	0,92	0,48	0,53	3
B09	6	0,687	0,005	0,91	0,87	0,49	0,63	2
B10	1	2,140	0,005	1,20	1,35	0,25	0,32	4
B10	2	2,722	0,006	1,01	1,26	0,36	0,22	4
B10	3	1,984	0,005	1,05	1,17	0,38	0,35	4
B10	4	2,041	0,005	0,90	0,97	0,49	0,34	4
B10	5	2,251	0,005	1,02	1,10	0,39	0,30	4
B10	6	2,037	0,005	0,96	0,96	0,46	0,34	4
BC1	1	-1,807	0,007	0,91	0,81	0,28	0,95	1
BC1	2	0,438	0,003	0,92	0,87	0,47	0,69	2
BC1	3	-0,993	0,005	0,93	0,79	0,35	0,89	1
BC1	4	-0,747	0,004	0,89	0,76	0,41	0,87	1
BC1	5	0,866	0,003	1,06	1,07	0,37	0,60	2
BC1	6	1,649	0,003	0,89	0,88	0,52	0,43	3
BC1	7	2,701	0,033	0,81	0,65	0,56	0,23	4
BC1	8	0,248	0,003	0,96	0,89	0,43	0,72	2
BC1	9	1,616	0,003	1,08	1,12	0,36	0,43	3
BC1	10	0,080	0,004	0,96	0,91	0,41	0,75	2
BC1	11	0,432	0,003	0,89	0,81	0,50	0,69	2
BC1	12	0,622	0,003	0,80	0,72	0,58	0,65	2
BC1	13	1,150	0,028	1,21	1,30	0,24	0,54	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC2	1_1	-1,280	0,003	1,07	1,15	-0,15	0,37	1
BC2	1_2	0,963	0,003	1,07	1,15	0,46	0,57	3
BC2	2	0,892	0,003	1,26	1,41	0,19	0,59	2
BC2	3	2,580	0,004	1,04	1,29	0,33	0,24	4
BC2	4	1,848	0,003	1,15	1,25	0,29	0,38	4
BC2	5	0,555	0,003	1,12	1,16	0,31	0,66	2
BC2	6	1,796	0,003	1,08	1,12	0,36	0,39	4
BC2	7	1,229	0,028	1,03	1,04	0,40	0,52	3
BC2	8	2,663	0,004	1,16	1,42	0,23	0,23	4
BC2	9	0,041	0,004	0,90	0,78	0,47	0,76	2
BC2	10	0,895	0,003	0,97	0,94	0,45	0,59	2
BC2	11	0,437	0,003	0,97	0,93	0,43	0,69	2
BC2	12	0,073	0,003	0,98	0,93	0,40	0,75	2
BC2	13	1,392	0,028	1,01	1,01	0,43	0,48	3

Tabla 4.11 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Matemática de 2.º grado de secundaria

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	1	0,226	0,008	0,88	0,81	0,50	0,71	1
B01	2	1,144	0,007	0,96	0,94	0,48	0,53	2
B01	3	1,947	0,007	0,96	0,98	0,48	0,37	3
B01	4	1,717	0,007	1,09	1,14	0,39	0,42	3
B01	5	0,978	0,007	1,07	1,09	0,39	0,57	2
B02	1	2,062	0,007	0,91	0,88	0,53	0,35	4
B02	2	0,654	0,007	1,02	0,99	0,41	0,63	2
B02	3	1,917	0,007	1,13	1,23	0,35	0,38	3
B02	4	1,385	0,007	1,08	1,11	0,39	0,48	3
B02	5	3,176	0,009	1,02	1,31	0,38	0,18	4
B03	1	-0,019	0,008	0,84	0,77	0,50	0,75	1
B03	2	1,261	0,007	1,00	1,01	0,44	0,51	2
B03	3	0,483	0,007	0,85	0,80	0,53	0,66	2
B03	4	2,175	0,007	1,14	1,33	0,33	0,33	4
B03	5	0,445	0,007	1,19	1,36	0,25	0,67	2
B04	1	-0,424	0,009	0,89	0,80	0,44	0,81	1
B04	2	1,798	0,007	0,96	0,97	0,50	0,41	3
B04	3	0,438	0,007	1,12	1,30	0,33	0,67	2
B04	4	2,631	0,008	0,89	0,90	0,52	0,26	4

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B04	5	0,699	0,007	0,94	0,92	0,48	0,62	2
B05	1	-0,550	0,009	1,00	1,04	0,32	0,83	1
B05	2	2,551	0,008	0,92	1,02	0,49	0,26	4
B05	3	1,003	0,007	0,89	0,86	0,51	0,56	2
B05	4	0,657	0,007	1,24	1,46	0,21	0,63	2
B05	5	2,244	0,008	1,05	1,17	0,39	0,32	4
B06	1	1,765	0,007	0,91	0,91	0,52	0,41	3
B06	2	0,940	0,007	0,84	0,78	0,56	0,57	2
B06	3	2,864	0,008	0,95	1,21	0,44	0,22	4
B06	4	0,719	0,007	0,99	0,98	0,44	0,62	2
B06	5	1,664	0,007	1,28	1,46	0,24	0,43	3
B07	1	2,021	0,007	0,86	0,87	0,56	0,36	3
B07	2	1,277	0,007	0,89	0,85	0,54	0,50	2
B07	3	1,997	0,007	1,02	1,07	0,44	0,36	3
B07	4	-	-	-	-	-	-	-
B07	5	1,130	0,007	1,03	1,04	0,42	0,53	2
B08	1	-	-	-	-	-	-	-
B08	2	-0,694	0,009	0,94	0,88	0,38	0,84	1
B08	3	1,015	0,007	1,28	1,40	0,22	0,55	2
B08	4	-0,661	0,009	0,97	1,05	0,35	0,84	1
B08	5	-	-	-	-	-	-	-
B09	1	2,026	0,007	1,04	1,03	0,45	0,36	4
B09	2	-0,416	0,009	0,86	0,75	0,47	0,81	1
B09	3	0,719	0,007	1,01	0,98	0,44	0,62	2
B09	4	2,526	0,008	0,81	0,76	0,59	0,27	4
B09	5	-0,422	0,009	0,83	0,67	0,49	0,81	1
B10	1	0,201	0,008	0,93	0,97	0,43	0,71	1
B10	2	1,369	0,007	1,07	1,08	0,39	0,48	3
B10	3	2,045	0,007	0,91	0,88	0,52	0,35	4
B10	4	2,242	0,007	1,22	1,36	0,27	0,31	4
B10	5	1,969	0,007	1,04	1,15	0,40	0,36	3
BC1	1	-0,275	0,004	0,81	0,68	0,51	0,79	1
BC1	2	-0,137	0,004	0,92	0,94	0,42	0,77	1
BC1	3	3,984	0,047	0,85	0,57	0,50	0,10	4
BC1	4	-1,394	0,005	0,94	1,03	0,30	0,91	1
BC1	5	3,116	0,004	0,97	1,34	0,40	0,19	4
BC1	6	0,789	0,003	0,89	0,83	0,52	0,60	2
BC1	7	1,695	0,003	0,91	0,90	0,52	0,42	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
BC1	8	1,549	0,003	1,06	1,10	0,41	0,45	3
BC1	9	1,921	0,003	0,92	0,96	0,51	0,38	3
BC1	10	1,455	0,003	1,22	1,38	0,27	0,47	3
BC1	11	4,314	0,052	0,92	0,78	0,42	0,08	4
BC1	12	-0,488	0,004	1,01	1,13	0,33	0,82	1
BC1	13	2,180	0,003	1,23	1,46	0,27	0,33	4
BC1	14	-1,316	0,005	0,92	1,01	0,32	0,91	1
BC1	15	0,569	0,003	0,96	0,91	0,46	0,65	2
BC1	16	-	-	-	-	-	-	-
BC1	17	-0,033	0,004	0,89	0,79	0,47	0,75	1
BC1	18	0,853	0,003	0,94	0,90	0,48	0,59	2
BC1	19	2,694	0,004	1,17	1,50	0,29	0,24	4
BC1	20	0,726	0,003	0,93	0,91	0,48	0,62	2
BC2	1	-0,394	0,004	0,84	0,72	0,48	0,80	1
BC2	2	-1,050	0,005	1,01	1,21	0,28	0,88	1
BC2	3	4,870	0,063	0,85	0,58	0,43	0,05	4
BC2	4	0,735	0,003	0,99	0,97	0,44	0,61	2
BC2	5	0,290	0,003	1,18	1,35	0,26	0,70	2
BC2	6	2,358	0,003	1,02	1,10	0,43	0,30	4
BC2	7	1,025	0,003	1,05	1,09	0,40	0,55	2
BC2	8	1,563	0,003	1,00	1,01	0,46	0,44	3
BC2	9	-	-	-	-	-	-	-
BC2	10	0,726	0,003	1,15	1,26	0,31	0,61	2
BC2	11	3,072	0,085	0,88	0,68	0,54	0,21	4
BC2	12	0,154	0,003	1,02	1,08	0,37	0,72	1
BC2	13	0,282	0,003	0,93	0,87	0,46	0,70	2
BC2	14	1,740	0,003	0,98	1,04	0,46	0,41	3
BC2	15	1,823	0,003	1,08	1,18	0,39	0,39	3
BC2	16	-0,039	0,004	0,84	0,76	0,50	0,75	1
BC2	17	1,193	0,003	0,95	0,95	0,48	0,52	2
BC2	18	0,752	0,003	0,95	0,92	0,47	0,61	2
BC2	19	1,228	0,003	0,97	0,96	0,47	0,51	2
BC2	20	1,126	0,003	0,98	0,96	0,47	0,53	2

Tabla 4.12 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Ciencia y Tecnología de 2.º grado de secundaria

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	1	-0,819	0,005	0,87	0,81	0,49	0,68	2
B01	2	1,398	0,006	1,04	1,18	0,21	0,23	4
B01	3	-0,341	0,005	1,20	1,25	0,09	0,59	3
B01	4	-	-	-	-	-	-	-
B01	5	0,987	0,006	1,09	1,17	0,19	0,30	4
B01	6	-2,316	0,008	0,91	0,74	0,37	0,89	1
B01	7	-1,035	0,006	0,94	0,89	0,40	0,72	2
B01	8	0,844	0,005	0,98	0,98	0,35	0,33	4
B01	9	-1,871	0,007	0,91	0,78	0,40	0,85	1
B01	10	-1,456	0,006	0,91	0,83	0,41	0,79	1
B01	11	-0,122	0,005	0,87	0,85	0,49	0,54	3
B01	12	-1,059	0,006	0,90	0,83	0,45	0,73	2
B01	13	-0,289	0,005	0,95	0,94	0,40	0,57	3
B01	14	-0,067	0,005	0,92	0,90	0,44	0,53	3
B02	1	1,513	0,006	1,09	1,25	0,08	0,20	4
B02	2	0,454	0,005	1,12	1,14	0,13	0,39	3
B02	3	-0,609	0,005	0,97	0,95	0,35	0,63	2
B02	4	1,348	0,006	1,05	1,14	0,17	0,22	4
B02	5	1,208	0,006	1,10	1,23	0,09	0,24	4
B02	6	-0,357	0,005	1,04	1,04	0,25	0,57	3
B02	7	-1,414	0,006	0,91	0,83	0,41	0,78	1
B02	8	0,117	0,005	0,93	0,93	0,40	0,47	3
B02	9	-0,307	0,005	0,91	0,89	0,43	0,56	3
B02	10	1,177	0,006	1,15	1,32	0,01	0,25	4
B02	11	0,221	0,005	1,00	1,01	0,30	0,44	3
B02	12	0,074	0,005	1,08	1,09	0,19	0,48	3
B02	13	0,089	0,005	1,09	1,10	0,18	0,47	3
B02	14	0,143	0,005	1,00	1,00	0,31	0,46	3
B03	1	-0,407	0,005	0,96	0,95	0,36	0,58	2
B03	2	0,509	0,005	0,92	0,91	0,41	0,38	4
B03	3	0,705	0,005	0,99	0,99	0,31	0,34	4
B03	4	-0,254	0,005	0,97	0,97	0,35	0,55	3
B03	5	0,062	0,005	1,02	1,03	0,28	0,47	3
B03	6	0,083	0,005	0,92	0,91	0,42	0,47	3
B03	7	-0,295	0,005	0,98	0,97	0,34	0,55	3
B03	8	0,772	0,005	0,97	0,97	0,34	0,32	4

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	9	0,412	0,005	1,05	1,06	0,24	0,40	3
B03	10	0,407	0,005	0,97	0,97	0,34	0,40	3
B03	11	1,559	0,006	1,12	1,32	0,05	0,19	4
B03	12	-0,112	0,005	0,87	0,85	0,49	0,51	3
B03	13	-0,329	0,005	1,08	1,10	0,20	0,56	3
B03	14	0,053	0,005	0,98	0,97	0,35	0,48	3
B04	1	-2,578	0,009	0,95	0,78	0,30	0,92	1
B04	2	1,513	0,006	1,06	1,15	0,18	0,21	4
B04	3	0,748	0,005	1,09	1,13	0,19	0,34	4
B04	4	-1,217	0,006	0,96	0,91	0,35	0,75	2
B04	5	-0,274	0,005	1,05	1,08	0,25	0,57	3
B04	6	0,250	0,005	1,01	1,01	0,32	0,45	3
B04	7	0,056	0,005	1,03	1,04	0,28	0,49	3
B04	8	-0,105	0,005	0,98	0,97	0,35	0,53	3
B04	9	0,666	0,005	1,07	1,10	0,22	0,36	4
B04	10	-0,339	0,005	1,00	1,02	0,32	0,58	3
B04	11	0,640	0,005	1,01	1,04	0,29	0,36	4
B04	12	0,692	0,005	1,00	1,01	0,32	0,35	4
B04	13	-0,351	0,005	0,93	0,91	0,41	0,58	3
B04	14	-0,846	0,005	0,90	0,84	0,45	0,68	2
B05	1	-1,179	0,006	0,95	0,92	0,36	0,76	2
B05	2	1,058	0,006	1,03	1,08	0,25	0,30	4
B05	3	-1,164	0,006	0,93	0,86	0,39	0,76	2
B05	4	-0,118	0,005	1,01	1,00	0,31	0,55	3
B05	5	0,963	0,005	1,08	1,15	0,19	0,31	4
B05	6	-1,081	0,006	0,96	0,92	0,35	0,74	2
B05	7	-0,418	0,005	0,94	0,91	0,40	0,61	2
B05	8	-0,032	0,005	1,00	1,00	0,33	0,53	3
B05	9	0,461	0,005	0,96	0,97	0,36	0,42	3
B05	10	-2,229	0,008	0,92	0,77	0,35	0,89	1
B05	11	-1,681	0,007	0,97	0,99	0,28	0,83	1
B05	12	0,404	0,005	0,98	0,98	0,35	0,43	3
B05	13	0,314	0,005	1,10	1,12	0,19	0,45	3
B05	14	-0,097	0,005	1,10	1,11	0,19	0,54	3
B06	1	1,725	0,006	1,19	1,50	0,00	0,19	4
B06	2	0,707	0,005	1,13	1,21	0,16	0,37	4
B06	3	-	-	-	-	-	-	-
B06	4	-0,163	0,005	0,93	0,92	0,42	0,56	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B06	5	-1,386	0,006	1,01	1,02	0,28	0,79	2
B06	6	0,354	0,005	1,05	1,08	0,27	0,44	3
B06	7	-0,201	0,005	0,97	0,96	0,37	0,56	3
B06	8	1,072	0,006	0,93	0,93	0,39	0,30	4
B06	9	-1,672	0,007	0,94	0,85	0,36	0,83	1
B06	10	0,535	0,005	1,06	1,08	0,26	0,40	4
B06	11	0,336	0,005	1,07	1,09	0,26	0,45	3
B06	12	-2,125	0,008	0,96	0,87	0,31	0,88	1
B06	13	-0,591	0,005	0,90	0,86	0,46	0,65	2
B06	14	2,707	0,009	1,00	1,15	0,19	0,09	4
B07	1	0,480	0,047	0,94	0,94	0,41	0,41	3
B08	1	0,906	0,050	0,94	0,91	0,38	0,30	4
B09	1_1	1,351	0,050	0,91	0,86	0,36	0,23	4
B09	1_2	3,817	0,050	0,91	0,86	0,18	0,01	4

Tabla 4.13 Dificultad y ajuste de los ítems al modelo Rasch de la prueba censal de Ciencias Sociales de 2.º grado de secundaria

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	1	1,524	0,006	1,11	1,23	0,20	0,25	4
B01	2	1,221	0,006	1,03	1,09	0,31	0,30	4
B01	3	1,035	0,005	1,14	1,21	0,21	0,33	4
B01	4	0,502	0,005	1,06	1,08	0,32	0,44	3
B01	5	0,646	0,005	1,09	1,13	0,27	0,41	3
B01	6	0,749	0,005	1,03	1,05	0,33	0,39	3
B01	7	-2,516	0,009	0,91	0,69	0,36	0,91	1
B01	8	-0,533	0,005	0,87	0,80	0,51	0,65	2
B01	9	0,591	0,005	1,00	1,02	0,37	0,42	3
B01	10	-1,903	0,007	0,95	0,94	0,34	0,86	1
B01	11	0,118	0,005	0,92	0,91	0,46	0,52	3
B01	12	0,114	0,005	0,96	0,94	0,42	0,52	3
B01	13	-0,275	0,005	0,91	0,87	0,47	0,60	2
B01	14	-1,394	0,006	0,85	0,70	0,49	0,80	1
B02	1	0,226	0,005	0,90	0,88	0,47	0,48	3
B02	2	1,770	0,006	1,12	1,42	0,10	0,19	4
B02	3	0,871	0,005	0,92	0,91	0,43	0,35	4
B02	4	0,686	0,005	1,10	1,14	0,22	0,38	3
B02	5	1,844	0,006	1,08	1,26	0,17	0,18	4

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B02	6	0,514	0,005	0,96	0,96	0,39	0,42	3
B02	7	-0,630	0,005	1,00	0,99	0,34	0,66	2
B02	8	0,142	0,005	0,90	0,89	0,46	0,50	3
B02	9	0,818	0,005	1,04	1,06	0,29	0,36	4
B02	10	0,293	0,005	0,97	0,97	0,38	0,47	3
B02	11	0,511	0,005	1,02	1,02	0,33	0,42	3
B02	12	1,794	0,006	1,17	1,46	0,05	0,19	4
B02	13	1,409	0,006	1,09	1,20	0,19	0,25	4
B02	14	0,156	0,005	1,02	1,03	0,33	0,50	3
B03	1	1,420	0,006	1,01	1,08	0,31	0,25	4
B03	2	-2,103	0,008	0,91	0,76	0,37	0,87	1
B03	3	-0,314	0,005	1,01	1,02	0,34	0,59	2
B03	4	-1,897	0,007	0,90	0,76	0,41	0,85	1
B03	5	0,476	0,005	1,03	1,04	0,33	0,43	3
B03	6	0,849	0,005	1,03	1,06	0,32	0,35	4
B03	7	0,869	0,005	1,09	1,12	0,26	0,35	4
B03	8	0,600	0,005	1,06	1,11	0,28	0,40	3
B03	9	0,356	0,005	1,09	1,12	0,26	0,45	3
B03	10	-2,892	0,010	0,92	0,67	0,33	0,93	1
B03	11	-0,571	0,005	0,97	0,96	0,38	0,64	2
B03	12	-0,049	0,005	0,98	0,97	0,39	0,54	3
B03	13	0,326	0,005	0,96	0,97	0,40	0,46	3
B03	14	-1,061	0,006	0,97	0,95	0,37	0,73	1
B04	1	0,733	0,005	0,93	0,92	0,46	0,38	3
B04	2	-0,368	0,005	0,95	0,95	0,43	0,61	2
B04	3	-0,367	0,005	0,96	0,92	0,43	0,61	2
B04	4	1,466	0,006	1,13	1,29	0,23	0,25	4
B04	5	0,138	0,005	1,08	1,09	0,32	0,50	3
B04	6	0,873	0,005	1,17	1,27	0,22	0,36	4
B04	7	-0,337	0,005	1,00	1,00	0,39	0,60	2
B04	8	-1,277	0,006	0,91	0,84	0,43	0,77	1
B04	9	-0,966	0,006	0,90	0,83	0,46	0,72	2
B04	10	-0,195	0,005	0,92	0,90	0,47	0,57	3
B04	11	0,058	0,005	0,96	0,96	0,43	0,52	3
B04	12	-0,671	0,005	1,02	1,08	0,34	0,67	2
B04	13	0,874	0,005	1,11	1,20	0,28	0,36	4
B04	14	0,013	0,005	0,96	0,95	0,44	0,53	3
B05	1	-1,365	0,006	0,90	0,83	0,44	0,79	1

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B05	2	-1,048	0,006	0,96	0,97	0,40	0,74	1
B05	3	0,346	0,005	1,02	1,04	0,40	0,48	3
B05	4	-0,180	0,005	0,99	0,99	0,43	0,58	3
B05	5	0,129	0,005	1,02	1,02	0,41	0,52	3
B05	6	-0,821	0,006	0,96	0,95	0,42	0,70	2
B05	7	-0,464	0,005	0,93	0,88	0,47	0,64	2
B05	8	-0,685	0,006	0,92	0,88	0,47	0,68	2
B05	9	0,380	0,005	1,07	1,09	0,37	0,47	3
B05	10	-	-	-	-	-	-	-
B05	11	0,405	0,005	1,19	1,24	0,26	0,47	3
B05	12	-0,698	0,006	0,87	0,78	0,51	0,68	2
B05	13	0,368	0,005	0,92	0,91	0,49	0,47	3
B05	14	-0,390	0,005	0,94	0,88	0,47	0,62	2
B06	1	-0,202	0,005	1,03	1,04	0,38	0,60	3
B06	2	-0,441	0,005	1,03	1,06	0,36	0,64	2
B06	3	0,689	0,005	1,01	1,04	0,40	0,42	3
B06	4	1,930	0,006	1,10	1,38	0,24	0,20	4
B06	5	-0,160	0,005	0,97	0,95	0,44	0,59	3
B06	6	0,748	0,005	1,16	1,23	0,26	0,40	3
B06	7	0,151	0,005	1,06	1,09	0,36	0,52	3
B06	8	-0,032	0,005	1,01	1,01	0,40	0,56	3
B06	9	-	-	-	-	-	-	-
B06	10	-0,205	0,005	0,96	0,92	0,45	0,60	3
B06	11	-1,702	0,007	0,90	0,76	0,42	0,84	1
B06	12	-0,243	0,005	0,90	0,87	0,49	0,60	3
B06	13	-0,229	0,005	0,91	0,88	0,49	0,60	3
B06	14	0,644	0,005	0,95	0,96	0,46	0,43	3
B07	1	1,651	0,057	0,90	0,81	0,45	0,22	4
B08	1	0,666	0,048	0,98	0,96	0,40	0,39	3
B09	1	-1,006	0,054	0,86	0,74	0,51	0,74	2

Tabla 4.14 Dificultad y ajuste de los ítems al modelo Rasch de la prueba muestral de Ciudadanía de 6.º grado de primaria

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	1	0,382	0,045	0,97	0,94	0,49	0,58	3
B01	2	-0,269	0,048	1,11	1,14	0,35	0,71	3
B01	3	0,082	0,046	1,09	1,15	0,37	0,64	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B01	4	-0,864	0,054	0,84	0,63	0,52	0,81	2
B01	5	1,963	0,050	0,94	0,91	0,46	0,26	4
B01	6	-0,250	0,048	1,05	1,05	0,39	0,71	3
B01	7	0,155	0,045	0,94	0,90	0,50	0,63	3
B01	8	-1,854	0,074	0,87	0,59	0,40	0,92	1
B01	9	0,028	0,046	0,99	0,97	0,45	0,65	3
B01	10	-0,344	0,049	0,87	0,79	0,52	0,73	2
B01	11	0,463	0,045	1,01	1,01	0,45	0,56	3
B01	12	1,751	0,048	1,06	1,24	0,37	0,29	4
B01	13	-0,122	0,047	0,90	0,81	0,52	0,68	3
B01	14	0,001	0,047	1,01	0,95	0,45	0,66	3
B01	15	-1,144	0,059	0,92	0,76	0,43	0,85	2
B02	1	-0,717	0,053	0,88	0,75	0,49	0,79	2
B02	2	0,105	0,046	0,93	0,86	0,50	0,64	3
B02	3	-0,358	0,049	0,99	0,93	0,42	0,73	2
B02	4	0,738	0,044	1,01	1,02	0,44	0,51	4
B02	5	2,430	0,055	0,90	0,76	0,46	0,19	4
B02	6	0,085	0,046	0,87	0,78	0,55	0,65	3
B02	7	-0,357	0,049	1,03	1,10	0,38	0,73	2
B02	8	1,209	0,045	1,24	1,40	0,25	0,41	4
B02	9	1,099	0,045	1,08	1,12	0,39	0,43	4
B02	10	1,189	0,045	0,95	0,91	0,49	0,41	4
B02	11	0,339	0,045	1,00	0,98	0,45	0,60	3
B02	12	0,500	0,045	0,99	0,95	0,47	0,56	3
B02	13	1,372	0,046	1,27	1,46	0,21	0,37	4
B02	14	-0,732	0,053	0,84	0,68	0,52	0,79	2
B02	15	0,211	0,046	1,00	0,96	0,45	0,62	3
B03	1	0,103	0,046	1,23	1,34	0,24	0,66	3
B03	2	-0,216	0,048	0,94	0,93	0,46	0,72	3
B03	3	1,023	0,044	1,12	1,19	0,34	0,46	4
B03	4	0,948	0,044	1,16	1,21	0,31	0,47	4
B03	5	0,871	0,044	1,10	1,12	0,36	0,49	4
B03	6	-0,437	0,050	0,86	0,77	0,51	0,76	2
B03	7	-	-	-	-	-	-	-
B03	8	-0,286	0,049	0,83	0,71	0,55	0,73	3
B03	9	0,431	0,045	0,95	0,94	0,48	0,59	3
B03	11	0,983	0,044	1,07	1,15	0,37	0,47	4
B03	12	-0,522	0,052	0,90	0,73	0,49	0,77	2

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B03	13	0,109	0,046	0,88	0,81	0,53	0,66	3
B03	14	0,102	0,046	1,00	1,03	0,42	0,66	3
B03	15	0,295	0,045	1,09	1,10	0,36	0,62	3
B04	1	0,650	0,044	1,11	1,14	0,34	0,54	3
B04	2	0,005	0,047	1,01	1,01	0,42	0,67	3
B04	3	-	-	-	-	-	-	-
B04	4	0,467	0,044	0,97	0,95	0,47	0,58	3
B04	5_1	-0,782	0,036	1,09	1,09	-0,03	0,60	2
B04	5_2	2,109	0,036	1,09	1,09	0,35	0,22	4
B04	6	0,137	0,046	0,94	0,90	0,48	0,65	3
B04	7	0,234	0,045	0,88	0,81	0,54	0,63	3
B04	8	-1,529	0,069	0,87	0,58	0,44	0,90	1
B04	9	-0,209	0,048	0,97	0,99	0,43	0,71	3
B04	10	0,682	0,044	1,05	1,06	0,40	0,53	4
B04	11	0,306	0,045	1,00	1,00	0,44	0,61	3
B04	12	1,416	0,045	1,19	1,41	0,24	0,37	4
B04	13	-0,301	0,049	0,94	0,88	0,46	0,73	3
B04	14	-0,301	0,049	0,95	0,88	0,46	0,73	3
B04	15	0,678	0,044	1,02	1,03	0,42	0,53	4
B05	1	0,733	0,044	1,19	1,27	0,29	0,52	4
B05	2	0,623	0,044	1,14	1,20	0,34	0,55	3
B05	3	0,432	0,045	1,17	1,19	0,32	0,59	3
B05	4	-0,761	0,055	0,95	0,79	0,45	0,81	2
B05	5_1_1	-0,523	0,032	1,02	1,01	-0,11	0,38	2
B05	5_1_2	1,107	0,032	1,02	1,01	0,49	0,43	4
B05	5_2_1	0,533	0,045	0,86	0,80	-0,11	0,13	3
B05	5_2_2	0,533	0,045	0,86	0,80	0,59	0,57	3
B05	6	0,298	0,046	1,01	1,00	0,44	0,62	3
B05	7	-0,667	0,054	0,79	0,63	0,57	0,79	2
B05	8	-0,297	0,050	0,94	0,92	0,47	0,73	3
B05	9	-0,917	0,057	0,83	0,66	0,52	0,83	2
B05	10_1	-0,971	0,038	1,28	1,28	0,24	0,63	2
B05	10_2	2,161	0,038	1,28	1,28	0,09	0,22	4
B05	11	0,551	0,045	0,98	0,99	0,47	0,56	3
B05	12	-0,800	0,056	0,88	0,69	0,50	0,81	2
B05	13	-0,097	0,048	0,95	0,93	0,47	0,70	3
B05	14	0,584	0,045	1,07	1,10	0,39	0,56	3
B05	15	0,294	0,046	0,99	0,96	0,46	0,62	3

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B06	1	-1,023	0,058	0,97	0,83	0,40	0,84	2
B06	2	-	-	-	-	-	-	
B06	3	-1,783	0,075	0,89	0,60	0,40	0,92	1
B06	4	-0,616	0,053	1,03	1,07	0,36	0,79	2
B06	5	1,156	0,044	0,91	0,85	0,52	0,43	4
B06	6	-2,334	0,094	0,88	0,51	0,35	0,95	1
B06	7	-0,289	0,049	0,95	0,93	0,45	0,73	3
B06	8	-1,329	0,065	0,90	0,69	0,43	0,88	1
B06	9	-1,241	0,063	0,88	0,70	0,45	0,87	2
B06	10	-0,283	0,050	0,96	0,89	0,45	0,73	3
B06	11	-0,461	0,051	1,20	1,47	0,22	0,76	2
B06	12	-0,968	0,058	0,94	0,93	0,41	0,84	2
B06	13	-0,682	0,054	1,03	1,00	0,37	0,80	2
B06	14	0,353	0,045	0,98	0,96	0,46	0,61	3
B06	15	0,162	0,046	0,87	0,80	0,54	0,65	3
B07	1	1,361	0,045	1,15	1,32	0,28	0,38	4
B07	2	-0,389	0,050	0,92	0,81	0,47	0,75	2
B07	3	-1,081	0,060	0,94	0,80	0,39	0,85	2
B07	4	0,188	0,046	0,91	0,88	0,50	0,64	3
B07	5	0,613	0,044	1,09	1,10	0,36	0,55	3
B07	6	1,169	0,044	1,17	1,24	0,28	0,43	4
B07	7	1,851	0,048	0,98	1,14	0,40	0,29	4
B07	8	-0,314	0,049	0,95	0,85	0,45	0,74	3
B07	9	-1,091	0,060	0,87	0,66	0,45	0,86	2
B07	10_1	0,267	0,029	1,06	1,06	0,02	0,31	3
B07	10_2	1,366	0,029	1,06	1,06	0,47	0,35	4
B07	11	1,194	0,044	1,16	1,27	0,28	0,42	4
B07	12	1,984	0,049	0,96	1,10	0,41	0,26	4
B07	13	-0,426	0,051	0,93	0,85	0,46	0,76	2
B07	14	1,060	0,044	0,97	1,00	0,45	0,45	4
B07	15	0,586	0,044	1,14	1,22	0,31	0,56	3
B08	1	-0,238	0,049	0,95	0,92	0,45	0,72	3
B08	2	0,598	0,045	1,02	1,01	0,43	0,55	3
B08	3	-	-	-	-	-	-	-
B08	4	-1,196	0,062	0,79	0,54	0,51	0,86	2
B08	5_1	-1,380	0,039	1,10	1,12	0,09	0,65	1
B08	5_2	2,033	0,039	1,10	1,12	0,23	0,25	4
B08	6	1,030	0,044	0,96	0,95	0,48	0,46	4

Bloque	Posición en el bloque	Medida	Error	Infit	Outfit	ptme	p	Nivel
B08	7	0,068	0,047	1,10	1,11	0,35	0,67	3
B08	8	0,369	0,045	1,08	1,08	0,38	0,60	3
B08	9	-0,702	0,054	1,04	1,05	0,35	0,80	2
B08	10	0,142	0,046	0,83	0,77	0,57	0,65	3
B08	11	-0,588	0,053	0,99	0,95	0,40	0,78	2
B08	12	1,173	0,045	1,01	1,08	0,43	0,43	4
B08	13	0,176	0,046	1,04	1,04	0,41	0,64	3
B08	14	0,332	0,045	1,07	1,08	0,39	0,61	3
B08	15	0,158	0,046	1,02	1,00	0,43	0,65	3

Tabla 4.15 Dificultad y ajuste de los ítems al modelo Rasch de la prueba muestral de Escritura de 2.º grado de secundaria

Posición en la forma	Medida	Error	Infit	Outfit	ptme	p	Nivel
1	-2,892	0,057	0,97	0,89	0,39	0,93	1
2	-3,005	0,059	0,92	0,67	0,44	0,94	1
3_1	-3,843	0,027	1,19	1,18	-0,14	0,40	1
3_2	-0,261	0,027	1,19	1,18	0,33	0,56	2
4_1	-1,829	0,022	1,08	1,09	-0,01	0,46	1
4_2	0,609	0,022	1,08	1,09	0,31	0,41	3
5_1	-1,770	0,017	1,11	1,12	-0,19	0,15	1
5_2	-0,749	0,017	1,11	1,12	-0,03	0,25	2
5_3	0,152	0,017	1,11	1,12	0,44	0,50	2
6	-	-	-	-	-	-	-
7	-1,355	0,034	0,84	0,80	0,38	0,83	1
8_1	-3,168	0,043	0,99	0,92	-0,11	0,07	1
8_2	-2,223	0,043	0,99	0,92	0,36	0,91	1
9_1	-4,633	0,032	1,01	0,96	-0,15	0,18	1
9_2	-1,237	0,032	1,01	0,96	0,32	0,80	1
10_1	-1,635	0,021	1,03	1,04	-0,04	0,47	1
10_2	0,475	0,021	1,03	1,04	0,34	0,38	3
11_1	-2,071	0,018	1,09	1,06	-0,21	0,14	1
11_2	-0,983	0,018	1,09	1,06	-0,07	0,24	1
11_3	-0,150	0,018	1,09	1,06	0,43	0,56	2
12_1	-2,842	0,019	1,22	1,22	-0,13	0,12	1
12_2	-1,196	0,019	1,22	1,22	-0,10	0,29	1
12_3	-0,159	0,019	1,22	1,22	0,33	0,57	2
13	2,267	0,039	0,87	0,88	0,17	0,11	3

Posición en la forma	Medida	Error	Infit	Outfit	ptme	p	Nivel
14_1	-1,050	0,020	0,92	1,09	-0,03	0,25	1
14_2	-0,565	0,020	0,92	1,09	0,34	0,62	2
15	0,766	0,029	0,99	0,98	0,33	0,38	3
16	-1,397	0,035	0,99	1,01	0,32	0,81	1
17_1	-3,105	0,029	0,88	0,81	-0,14	0,16	1
17_2	-0,991	0,029	0,88	0,81	0,38	0,80	1
18_1	-4,087	0,029	1,17	1,22	-0,13	0,59	1
18_2	1,167	0,029	1,17	1,22	0,34	0,38	3
19_1	-0,880	0,021	0,98	0,98	0,10	0,50	2
19_2	1,324	0,021	0,98	0,98	0,32	0,25	3
20_1	-1,988	0,018	0,86	0,75	-0,20	0,10	1
20_2	-0,944	0,018	0,86	0,75	-0,12	0,25	2
20_3	0,029	0,018	0,86	0,75	0,50	0,58	2
21_1	-2,400	0,018	1,12	1,12	-0,13	0,21	1
21_2	-0,627	0,018	1,12	1,12	0,02	0,41	2
21_3	0,793	0,018	1,12	1,12	0,33	0,32	3
22	4,960	0,122	1,01	0,81	0,08	0,01	3
23_1	-0,974	0,019	0,93	0,93	-0,01	0,36	1
23_2	0,392	0,019	0,93	0,93	0,33	0,49	2
24_1	0,327	0,019	1,15	1,17	0,17	0,22	2
24_2	1,363	0,019	1,15	1,17	0,28	0,21	3
25	-0,362	0,029	1,16	1,22	0,20	0,62	2
26_1	-2,431	0,028	0,87	0,79	-0,17	0,18	1
26_2	-1,079	0,028	0,87	0,79	0,42	0,77	1
27_1	-5,701	0,029	1,25	1,28	-0,15	0,37	1
27_2	-0,427	0,029	1,25	1,28	0,35	0,59	2
28_1	-1,429	0,022	1,01	1,02	0,11	0,54	1
28_2	1,015	0,022	1,01	1,02	0,25	0,28	3
29_1	-2,326	0,019	0,94	0,82	-0,21	0,10	1
29_2	-1,141	0,019	0,94	0,82	-0,13	0,24	1
29_3	-0,102	0,019	0,94	0,82	0,49	0,61	2
30_1	-2,103	0,017	1,09	1,08	-0,17	0,22	1
30_2	-0,664	0,017	1,09	1,08	0,01	0,36	2
30_3	0,600	0,017	1,09	1,08	0,40	0,34	3
31	2,294	0,040	1,21	1,41	0,15	0,15	3
32	-1,473	0,035	0,94	0,91	0,41	0,81	1
33	4,163	0,084	1,00	0,96	0,11	0,03	3

4.4.2 Evidencias de validez vinculadas a la estructura interna de las pruebas

Wilson (2005) señala que la evidencia de validez basada en la estructura interna requiere, en primer lugar, que se tenga claridad conceptual sobre dicha estructura, dado el constructo de interés. En este sentido, un tipo de evidencia está relacionado con la unidimensionalidad de las medidas derivadas de la aplicación del test. El concepto de unidimensionalidad implica que un solo rasgo latente o constructo se encuentra en la base de un conjunto de ítems (Hattie, 1985). En otras palabras, un instrumento será unidimensional si las respuestas son producidas sobre la base de un único atributo. Wright y Linacre (1989) señalan que, en la práctica, ningún instrumento puede ser perfectamente unidimensional. Lo que se busca es tener instrumentos que, en esencia, muestren unidimensionalidad. Por ejemplo, muchos factores como la motivación, la ansiedad y la velocidad de respuesta tienen un impacto sobre el desempeño de una persona en un conjunto de ítems (Hambleton et ál., 1991). Lo importante es que un instrumento de medida represente con sus puntuaciones un solo factor dominante. Con esto, lo que se quiere lograr es que la mayor cantidad de la varianza observada en las respuestas a los ítems sea explicada por un solo atributo latente (Embretson y Reise, 2000). Esto se suele analizar aplicando el análisis factorial, que debe ser diferenciado del análisis de componentes principales (Preacher y MacCallum, 2003). El primero de ellos trata de describir la estructura latente de un conjunto de variables, usando factores que expliquen la varianza común; el segundo trata de reducir las variables a un conjunto menor de componentes enfocándose en la varianza total (Park, Dailey y Lemus, 2002; Preacher y MacCallum, 2003).

En el caso de las pruebas aplicadas en la ECE, se realizó un análisis de componentes principales de los residuos estandarizados luego de ajustar el modelo Rasch unidimensional (Linacre, 2019). El análisis factorial de los residuos sirve para detectar cualquier varianza que queda en los datos luego de extraer la dimensión principal (Bond y Fox, 2015); es decir, trata de modelar la varianza que no es explicada por el constructo de interés que pretende medir un test.

Linacre (1998) ha demostrado que este tipo de residuos sirve para detectar mejor la presencia de dimensiones secundarias que otros tipos de residuos, como los no estandarizados o los logarítmicos. Además, este mismo autor señala que los resultados obtenidos con el análisis de componentes principales son similares a los del análisis factorial, pero prefirió el análisis de componentes principales por su rigurosa base matemática.

Tabla 4.16 Análisis de unidimensionalidad de las medidas derivadas de la aplicación de las pruebas censales en primaria

Prueba	Varianza de la dimensión principal (%)	Varianza modelada (%)	Primer autovalor	Primer autovalor (%)
Lectura, 4.º grado de primaria	27,7	28,0	1,74	1,8
Matemática, 4.º grado de primaria	32,3	32,5	1,67	1,5
Lectura en aimara, 4.º grado de primaria EIB	21,8	22,0	2,72	3,1
Lectura en awajún, 4.º grado de primaria EIB	17,7	17,8	2,55	3,1
Lectura en ashaninka, 4.º grado de primaria EIB	19,7	19,8	2,73	3,2
Lectura en quechua cusco collao, 4.º grado de primaria EIB	21,7	21,9	2,26	2,6
Lectura en quechua chanca, 4.º grado de primaria EIB	20,0	20,2	2,69	3,2
Lectura en shipibo-konibo, 4.º grado de primaria EIB	22,8	22,9	2,80	3,3
Lectura en castellano como L2, 4.º grado de primaria EIB	24,3	24,6	2,70	3,0

Tabla 4.17 Análisis de unidimensionalidad de las medidas derivadas de la aplicación de las pruebas censales en secundaria

Prueba	Varianza de la dimensión principal (%)	Varianza modelada (%)	Primer autovalor	Primer autovalor (%)
Lectura, 2.º grado de secundaria	29,6	29,8	1,83	1,5
Matemática, 2.º grado de secundaria	31,6	31,3	1,83	1,5
Ciencias Sociales, 2.º grado de secundaria	25,3	25,1	1,61	1,4
Ciencia y Tecnología, 2.º grado de secundaria	23,1	23,0	1,86	1,7

Tabla 4.18 Análisis de unidimensionalidad de las medidas derivadas de la aplicación de la prueba muestral en primaria

Prueba	Varianza de la dimensión principal (%)	Varianza modelada (%)	Primer autovalor	Primer autovalor (%)
Ciudadanía, 6.º grado de primaria	27,8	28,4	1,60	1,0

Tabla 4.19 Análisis de unidimensionalidad de las medidas derivadas de la aplicación de la prueba muestral en secundaria

Prueba	Varianza de la dimensión principal (%)	Varianza modelada (%)	Primer autovalor	Primer autovalor (%)
Escritura, 2.º grado de secundaria	41,4	42,6	2,01	3,7

Linacre (2019) afirma que, si bien no existen parámetros absolutos para interpretar los resultados del análisis de componentes principales de los residuos, si el primer autovalor contiene menos del 5 % de varianza o si su valor es menor que 3, no hay serias evidencias que atenten contra el supuesto de unidimensionalidad y, por lo tanto, los datos se pueden analizar adecuadamente utilizando modelos Rasch. Además, se está frente a datos unidimensionales cuando la cantidad de varianza empírica es similar a la cantidad de varianza predicha por el modelo.

Todas estas condiciones se cumplen para las pruebas analizadas; es decir, no hay evidencias contundentes en contra del supuesto de unidimensionalidad del conjunto de ítems incluidos en estos instrumentos de medición.

4.4.3 Confiabilidad y consistencia de la clasificación

El índice de confiabilidad de la separación de las personas (R_p) es el coeficiente utilizado para las pruebas aplicadas de la UMC, el cual es análogo al de consistencia interna alpha de Cronbach, pero produce mejores estimaciones, pues los valores numéricos son lineales si los datos se ajustan al modelo Rasch aplicado. Además, utiliza la varianza de error promedio de la muestra en lugar de la varianza de error de una persona promedio (Schumacker y Smith, 2007). El coeficiente sirve para indicar la capacidad de las medidas de un test para diferenciar las cantidades de rasgo latente que poseen los evaluados (Wright y Masters, 1982). En ese sentido, indica la replicabilidad del ordenamiento de las personas según su medida de habilidad si se les da otro conjunto de ítems que miden el mismo constructo (Bond y Fox, 2015).

Un índice menor a 0,50 indica que las diferencias entre las medidas son producidas principalmente por el error de medición (Fisher, 1992). Sobre los valores mínimos aceptables de los coeficientes de confiabilidad, Charter (2003) ha realizado una revisión de numerosas investigaciones que proponen diferentes niveles mínimos. En ese estudio, encontró bastante variabilidad; asimismo, observó valores propuestos con los diversos métodos para obtener la confiabilidad, que oscilan entre 0,60 y 0,95. A pesar de esta gran variabilidad, un estándar mínimo aceptable que aparece con frecuencia en la literatura es el de 0,70, señalado por Nunnally y Bernstein (n.d.).

Estos coeficientes también pueden ser expresados como índices de separación de personas, que se refieren a la dispersión de los datos medidos como el número de errores estándar que separan a las personas (Schumacker y Smith, 2007). El índice de separación de personas (G_p) representa la variabilidad ajustada de las personas dividida entre el error estándar de medición promedio.

Tabla 4.20 Análisis de confiabilidad de las medidas derivadas de la aplicación de las pruebas censales en primaria

Prueba	<i>Rp</i>	<i>Gp</i>
Lectura, 4.º grado de primaria	0,87	2,64
Matemática, 4.º grado de primaria	0,85	2,41
Lectura en aimara, 4.º grado de primaria EIB	0,58	1,17
Lectura en awajún, 4.º grado de primaria EIB	0,82	2,11
Lectura en ashaninka, 4.º grado de primaria EIB	0,71	1,56
Lectura en quechua cusco collao, 4.º grado de primaria EIB	0,78	1,86
Lectura en quechua chanca, 4.º grado de primaria EIB	0,74	1,67
Lectura en shipibo-konibo, 4.º grado de primaria EIB	0,86	2,50
Lectura en castellano como L2, 4.º grado de primaria EIB	0,86	2,51

Tabla 4.21 Análisis de confiabilidad de las medidas derivadas de la aplicación de las pruebas censales en secundaria

Prueba	<i>Rp</i>	<i>Gp</i>
Lectura, 2.º grado de secundaria	0,87	2,64
Matemática, 2.º grado de secundaria	0,88	2,73
Ciencias Sociales, 2.º grado de secundaria	0,76	1,80
Ciencia y Tecnología, 2.º grado de secundaria	0,65	1,35

Considerando el valor de *Rp* de las pruebas censales aplicadas en el 2018, se puede apreciar que la varianza de error como máximo es de 42 % en la de Lectura en lengua originaria aimara, en 4.º grado de primaria EIB y como mínimo de 12 % en Matemática de 2.º grado de secundaria; por lo cual es posible afirmar que las medidas derivadas de aplicar dichas pruebas poseen adecuadas evidencias de confiabilidad.

Tabla 4.22 Análisis de confiabilidad de las medidas derivadas de la aplicación de la prueba muestral en primaria

Prueba	<i>Rp</i>	<i>Gp</i>
Ciudadanía, 6.º grado de primaria	0,82	2,13

Tabla 4.23 Análisis de confiabilidad de las medidas derivadas de la aplicación de la prueba muestral en secundaria

Prueba	<i>Rp</i>	<i>Gp</i>
Escritura, 2.º grado de secundaria	0,87	2,63

Del mismo modo, en las pruebas muestrales aplicadas en el 2018, se puede apreciar que la varianza de error es de 13 % en la prueba de Escritura en 2.º grado de secundaria y 18 % en la de Ciudadanía de 6.º grado de primaria, por lo cual es

posible afirmar que las medidas derivadas de aplicar dichas pruebas poseen adecuadas evidencias de confiabilidad.

Otros indicadores importantes para el análisis son los de precisión y consistencia de clasificación. La consistencia de la clasificación es el grado de acuerdo en dos administraciones independientes o paralelas de un instrumento de medición. Se espera que el alumno sea clasificado en la misma categoría al repetirse la evaluación. Por otro lado, la precisión de la clasificación implica el grado en el cual la clasificación observada coincide con la clasificación verdadera (Kim, Choi, Um y Kim, 2006).

En la práctica, es difícil de lograrlo; por ello, se han propuesto métodos para estimarla a partir de una sola aplicación. En la EM 2018, se utilizó el método de Rudner, que es una aproximación de tipo individual, pues calcula la consistencia de clasificación para cada persona y luego las promedia (Lee, 2010). Además, asume que los errores de estimación se distribuyen normalmente. A medida que aumenta el número de ítems, dadas las propiedades del estimador $\hat{\theta}$ basado en un método de máxima verosimilitud, este supuesto es más plausible.

Tabla 4.24 Indicadores del análisis de precisión y consistencia de la clasificación de las pruebas censales de Lectura y Matemática de 4.º grado de primaria

	Lectura		Matemática	
	Precisión	Consistencia	Precisión	Consistencia
Nacional	0,80	0,73	0,78	0,69
Hombre	0,80	0,73	0,78	0,69
Mujer	0,81	0,73	0,78	0,69
Urbano	0,80	0,73	0,78	0,69
Rural	0,81	0,73	0,78	0,70
Estatatal	0,80	0,72	0,78	0,69
No estatal	0,81	0,74	0,78	0,69
Polidocente completo	0,80	0,73	0,78	0,69
Unidocente / Multigrado	0,81	0,73	0,78	0,69

Tabla 4.25 Indicadores del análisis de precisión y consistencia de la clasificación de la prueba censal de Lectura de 4.º grado de primaria EIB

	Precisión	Consistencia
Aimara, 4.º grado de primaria EIB	0,78	0,69
Awajún, 4.º grado de primaria EIB	0,87	0,81
Ashaninka, 4.º grado de primaria EIB	0,82	0,74
Quechua cusco collao, 4.º grado de primaria EIB	0,82	0,75
Quechua chanca, 4.º grado de primaria EIB	0,81	0,74
Shipibo-konibo, 4.º grado de primaria EIB	0,91	0,87
Castellano como L2, 4.º grado de primaria EIB	0,86	0,80

Tabla 4.26 Indicadores del análisis de precisión y consistencia de la clasificación de las pruebas censales de Lectura, Matemática, Ciencias Sociales (CCSS) y Ciencia y Tecnología (CyT) de 2.º grado de secundaria

	Lectura		Matemática		CCSS		CyT	
	Pr	Cons	Pr	Cons	Pr	Cons	Pr	Cons
Nacional	0,74	0,64	0,77	0,68	0,68	0,58	0,66	0,55
Hombre	0,74	0,64	0,76	0,68	0,69	0,58	0,66	0,55
Mujer	0,74	0,65	0,77	0,68	0,68	0,58	0,66	0,55
Urbano	0,74	0,64	0,76	0,67	0,68	0,58	0,66	0,55
Rural	0,79	0,71	0,83	0,76	0,70	0,60	0,68	0,57
Estatad	0,74	0,65	0,77	0,69	0,68	0,58	0,66	0,55
No estatal	0,73	0,64	0,74	0,65	0,69	0,59	0,66	0,55

Tabla 4.27 Indicadores del análisis de precisión y consistencia de la clasificación de la prueba muestral de Ciudadanía de 6.º grado de primaria

	Precisión	Consistencia
Nacional	0,73	0,63
Hombre	0,72	0,61
Mujer	0,74	0,64
Urbano	0,73	0,63
Rural	0,70	0,59
Estatad	0,72	0,62
No estatal	0,76	0,67
Polidocente completo	0,73	0,63
Unidocente / Multigrado	0,70	0,59

Tabla 4.28 Indicadores del análisis de precisión y consistencia de la clasificación de la prueba muestral de Escritura de 2.º grado de secundaria

	Precisión	Consistencia
Nacional	0,82	0,75
Hombre	0,86	0,77
Mujer	0,82	0,75
Urbano	0,82	0,75
Rural	0,84	0,78
Estatal	0,82	0,75
No estatal	0,84	0,78
Polidocente completo	0,82	0,75
Unidocente / Multigrado	0,81	0,74

En general, las pruebas mantienen buenos indicadores de precisión y consistencia de la clasificación de los niveles de logro.

4.4.4 Equiparación de medidas

En el marco del análisis Rasch, las puntuaciones del rasgo latente tienen un origen y una escala de medición arbitrarios (Bond y Fox, 2015). Por ello, dos instrumentos que miden lo mismo, aplicados a distintas muestras de personas, no se encontrarán en la misma métrica. A fin de establecer las comparaciones necesarias, se deben poner ambas pruebas en la misma escala o equipararlas. Esta conversión, en ocasiones, implica el uso de una constante aditiva y, en otras, el uso de una constante multiplicativa, además de la aditiva (Livingston, 2004).

Diversas síntesis teóricas que tratan sobre la equiparación de puntuaciones (Kolen y Brennan, 2004; Navas, 1996) citan los trabajos de Angoff y Lord (ambos en la primera mitad de los años ochenta) como los pioneros en la equiparación. Estos autores entienden la equiparación como el desarrollo de un sistema de conversión de las unidades de una prueba a las unidades de otra, de manera que sus resultados sean comparables o equivalentes.

La finalidad es tener una métrica común para dos o más medidas de un mismo rasgo, de modo que se puedan comparar los resultados de personas a las que se les aplican diferentes instrumentos que miden el mismo rasgo. Por ejemplo, se tienen dos test (X e Y) y se desea equiparar el test X con el test Y . Por lo tanto, el objetivo es encontrar un sistema o función para convertir la métrica de X en Y .

Revisando la literatura sobre el tema (Ho y Osborn, 2005; Kolen y Brennan, 2004; Livingston, 2004; Navas, 1996; Zhu, 1998), se puede señalar que hay tres tipos de diseño de equiparación: de un solo grupo, de grupos equivalentes y de grupos no equivalentes con ítems comunes. Este último diseño fue utilizado para colocar las

medidas de las pruebas de las distintas evaluaciones censales de segundo grado de primaria en la misma métrica que las pruebas aplicadas en el 2007 por la UMC. En este diseño, se trabaja con dos muestras de personas que no necesariamente han sido extraídas de la misma población. En cada muestra, se aplica una única forma del test. Lo esencial de este método es que en cada grupo se administra un conjunto de ítems o un test común que permite establecer la equivalencia entre los test que se quieren equiparar.

Una vez obtenidos los datos, se utilizó la transformación lineal: se consideran equivalentes las puntuaciones directas que corresponden a la misma puntuación típica. En el contexto de la teoría de respuesta al ítem (TRI), se denomina a este método *mean/sigma* (Kolen y Brennan, 2004), e implica el cálculo de dos constantes de equiparación a y b , de tal manera que surge la siguiente fórmula:

$$medida_{equiparadaECE} = b + a \times medida_{original},$$

donde:

$$a = \frac{S_{medidaECE}}{S_{medida_{original}}}$$

$$b = \bar{X}_{medidaECE} - a \times \bar{X}_{medida_{original}}$$

Un requisito importante para realizar la comparación del rendimiento entre dos o más grupos es asegurar la invarianza de la dificultad de los ítems de la prueba a lo largo del tiempo (Kolen y Brennan, 2004). Esto supone que el grado de dificultad de cada ítem debe mantenerse constante⁹, independientemente del grupo poblacional en el que se le aplique. En caso de no cumplirse con este supuesto, se dice que el ítem presenta un funcionamiento diferencial (FDI) a lo largo del tiempo. Por ello, todos los ítems utilizados para el proceso de equiparación son sometidos a este tipo de análisis según la metodología propuesta por Bond y Fox (2015). De esta manera, solo se utilizan en la equiparación aquellos ítems sin funcionamiento diferencial a lo largo del tiempo.

En el caso de la prueba de Escritura, se utilizaron los mismos estímulos que en la evaluación del 2016. Por ello, se aplicó un método de equiparación con todos los ítems disponibles, anclando sus medidas de dificultad según los datos obtenidos en el 2016. Además, se analizó su FDI a lo largo del tiempo considerando la medida de desplazamiento, cuyo valor se esperaba fuese menor a 0,5 logits (Linacre, 2019).

A continuación, se presentan las tablas con los resultados de la equiparación por ítems comunes de las pruebas utilizadas en las evaluaciones 2018.

⁹En realidad, la dificultad del ítem debe mantenerse dentro de ciertos márgenes de error al comparar su dificultad en dos poblaciones diferentes.

Tabla 4.29 Análisis FDI de las medidas de la prueba censal de Lectura de 4.º grado de primaria usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
B03	1	-2,199	0,004	-1,372	0,004	-2,329	0,130	0,006
B08	5	-0,420	0,006	0,607	0,005	-0,079	-0,341	0,008
B06	1	-2,249	0,004	-1,212	0,004	-2,147	-0,102	0,006
B06	2	-2,120	0,004	-0,889	0,004	-1,780	-0,340	0,006
B06	3	-0,699	0,003	0,168	0,003	-0,578	-0,121	0,004
B06	4	-2,070	0,004	-1,447	0,004	-2,414	0,344	0,006
B06	5	0,142	0,003	0,825	0,003	0,169	-0,027	0,004
B06	6	-2,107	0,004	-1,177	0,004	-2,107	0,000	0,006
B03	3	-2,861	0,005	-1,716	0,005	-2,720	-0,141	0,007
B03	5	-0,925	0,003	-0,413	0,003	-1,239	0,314	0,004
B03	4	-0,911	0,003	-0,387	0,003	-1,209	0,298	0,004
B12	4	0,554	0,006	1,222	0,005	0,621	-0,067	0,008
B09	4	-0,123	0,006	0,733	0,005	0,065	-0,188	0,008
B05	6	0,578	0,003	1,067	0,003	0,445	0,133	0,004
B05	4	0,212	0,003	0,801	0,003	0,142	0,070	0,004
B05	5	-0,607	0,003	0,132	0,003	-0,619	0,012	0,004
B09	1	0,073	0,006	0,802	0,005	0,143	-0,070	0,008
B12	6	-1,131	0,006	-0,217	0,006	-1,016	-0,115	0,008
B03	6	0,714	0,003	1,118	0,003	0,503	0,211	0,004

Tabla 4.30 Análisis FDI de las medidas de la prueba censal de Matemática de 4.º grado de primaria usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
B02	1	-2,008	0,008	-0,931	0,007	-1,729	-0,279	0,011
BC2	11	-0,508	0,006	0,204	0,003	-0,681	0,173	0,007
BC1	14	-2,496	0,009	-1,467	0,005	-2,224	-0,272	0,010
BC1	10	-1,572	0,007	-0,522	0,004	-1,351	-0,221	0,008
B06	1	-1,987	0,004	-1,106	0,007	-1,891	-0,096	0,008
B06	10	-1,482	0,007	-0,577	0,007	-1,402	-0,080	0,010
BC2	8	-2,996	0,010	-2,361	0,006	-3,050	0,054	0,012
BC2	13	-1,240	0,006	-0,408	0,004	-1,246	0,006	0,007
B05	10	-2,547	0,009	-2,008	0,010	-2,724	0,177	0,013
B01	1	-2,314	0,005	-1,665	0,009	-2,407	0,093	0,010
BC2	14	-1,799	0,007	-1,458	0,005	-2,216	0,417	0,009
B04	10	-2,623	0,005	-1,636	0,009	-2,380	-0,243	0,010
BC1	7	-1,055	0,006	0,060	0,003	-0,814	-0,241	0,007
BC2	10	-1,116	0,004	-0,714	0,004	-1,529	0,413	0,006
B02	8	-1,187	0,004	-0,179	0,006	-1,035	-0,152	0,007
BC2	4	-0,397	0,006	0,306	0,003	-0,587	0,190	0,007
B03	2	-1,323	0,004	-0,733	0,007	-1,546	0,223	0,008
BC2	15	-1,388	0,007	-0,950	0,004	-1,747	0,359	0,008
B04	7	-2,038	0,004	-1,379	0,008	-2,143	0,105	0,009

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC1	4	-1,886	0,007	-0,817	0,004	-1,624	-0,262	0,008
B06	3	0,521	0,003	1,668	0,006	0,671	-0,150	0,007
B06	9	-0,765	0,006	0,306	0,006	-0,587	-0,178	0,008
BC2	9	0,248	0,006	1,268	0,003	0,302	-0,054	0,007
B05	6	-0,459	0,006	0,429	0,006	-0,473	0,014	0,008
B04	3	-0,503	0,003	0,783	0,006	-0,146	-0,357	0,007
BC1	2	-2,207	0,008	-1,104	0,004	-1,889	-0,318	0,009
B06	4	-0,486	0,006	0,441	0,006	-0,462	-0,024	0,008
B04	6	-0,014	0,003	0,933	0,006	-0,007	-0,007	0,007
B05	5	0,413	0,006	1,272	0,006	0,306	0,107	0,008
BC1	12	-1,647	0,007	-0,677	0,004	-1,495	-0,152	0,008
BC1	6	-0,965	0,006	-0,402	0,004	-1,241	0,276	0,007
B01	10	-1,115	0,006	-0,392	0,007	-1,231	0,116	0,009
B02	6	0,734	0,006	1,603	0,006	0,611	0,123	0,008
BC2	12	-1,613	0,007	-1,068	0,004	-1,856	0,243	0,008

Tabla 4.31 Análisis FDI de las medidas de la prueba censal de Lectura en aimara de 4.º grado de primaria EIB usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC01	4	-1,682	0,149	-2,177	0,196	-1,572	-0,110	0,246
BC1	5	-1,097	0,118	-1,776	0,172	-1,238	0,141	0,209
BC1	1	-1,235	0,126	-1,477	0,159	-0,988	-0,247	0,203
BC1	2	-1,796	0,158	-1,898	0,179	-1,340	-0,456	0,239
BC3	3	-1,733	0,154	-2,497	0,218	-1,839	0,106	0,267
BC2	1	1,357	0,096	0,744	0,147	0,865	0,492	0,176
BC2	3	0,889	0,090	0,398	0,139	0,576	0,313	0,166
BC3	1	-2,350	0,203	-3,564	0,342	-2,729	0,379	0,398
BC3	2	-2,215	0,190	-3,352	0,311	-2,553	0,338	0,364
BC1	3	-1,538	0,140	-2,177	0,196	-1,572	0,034	0,241
BC4	1	-1,038	0,116	-0,949	0,141	-0,548	-0,490	0,183
BC4	2	0,901	0,090	0,562	0,141	0,713	0,188	0,167
BC4	3	0,065	0,090	0,074	0,134	0,306	-0,241	0,161
B02	1	0,694	0,089	0,800	0,207	0,911	-0,217	0,225
B02	2	0,545	0,088	0,290	0,192	0,486	0,059	0,211
B02	4	0,646	0,088	0,478	0,196	0,643	0,003	0,215
B05	1	0,457	0,088	0,290	0,192	0,486	-0,029	0,211
B05	2	-0,026	0,091	-0,295	0,189	-0,002	-0,024	0,210
B05	5	0,491	0,088	0,144	0,190	0,364	0,127	0,209
B02	4	0,454	0,089	0,061	0,135	0,295	0,159	0,162
BC4	4	0,335	0,089	0,293	0,137	0,488	-0,153	0,163
BC4	5	0,964	0,090	0,748	0,146	0,868	0,096	0,172
B02	6	0,198	0,089	0,002	0,188	0,246	-0,048	0,208
B05	4	0,350	0,089	0,635	0,201	0,774	-0,424	0,220

Tabla 4.32 Análisis FDI de las medidas de la prueba censal de Lectura en ashaninka de 4.º grado de primaria EIB usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC3	1	-2,007	0,078	-2,715	0,097	-2,520	0,513	0,124
BC3	3	-0,978	0,059	-1,349	0,064	-1,155	0,177	0,087
BC1	4	-1,270	0,063	-1,201	0,062	-1,007	-0,263	0,088
BC1	5	-0,916	0,059	-1,116	0,062	-0,922	0,006	0,086
BC2	1	-1,810	0,074	-1,458	0,066	-1,264	-0,546	0,099
BC2	2	-0,886	0,059	-0,580	0,058	-0,387	-0,499	0,083
BC2	3	0,521	0,055	0,405	0,060	0,598	-0,077	0,081
BC1	1	-2,185	0,084	-2,307	0,084	-2,112	-0,073	0,119
BC3	2	-1,156	0,062	-1,660	0,069	-1,466	0,310	0,093
BC4	1	1,300	0,062	1,041	0,067	1,233	0,067	0,091
BC4	4	1,127	0,060	0,605	0,062	0,797	0,330	0,086
BR2	1	-0,002	0,053	-0,044	0,082	0,149	-0,151	0,098
BR2	5	0,438	0,054	0,794	0,090	0,986	-0,548	0,105
BC3	5	-0,260	0,054	-0,773	0,059	-0,580	0,320	0,080
BC2	4	-0,687	0,057	-0,724	0,059	-0,531	-0,156	0,082
BC2	5	0,358	0,054	0,294	0,059	0,487	-0,129	0,080
BC3	4	0,523	0,055	0,009	0,058	0,202	0,321	0,080
BC3	6	0,332	0,054	-0,320	0,057	-0,127	0,459	0,079
BC4	3	0,652	0,056	0,096	0,058	0,289	0,363	0,081
B02	2	-0,156	0,054	-0,006	0,082	0,187	-0,343	0,098
B02	3	0,485	0,054	0,374	0,085	0,567	-0,082	0,101

Tabla 4.33 Análisis FDI de las medidas de la prueba censal de Lectura en awajún de 4.º grado de primaria EIB usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC1	1	-2,145	0,132	-2,073	0,051	-2,029	-0,116	0,142
BC3	1	-1,713	0,085	-1,860	0,047	-1,753	0,040	0,097
BC3	2	-1,033	0,107	-1,307	0,042	-1,037	0,004	0,115
BC1	2	-1,354	0,080	-1,385	0,043	-1,138	-0,216	0,091
BC4	1	-0,485	0,100	-0,575	0,039	-0,089	-0,396	0,107
BC4	3	-0,178	0,097	-0,612	0,039	-0,137	-0,041	0,105
BC4	2	0,045	0,095	0,447	0,059	1,234	-1,189	0,112
BC4	4	0,775	0,094	-0,569	0,039	-0,081	0,856	0,102
B01	4	1,629	0,095	0,260	0,058	0,992	0,637	0,111
BC2	1	-0,587	0,101	-1,126	0,041	-0,803	0,216	0,109
B08	6	2,786	0,108	0,603	0,061	1,436	1,350	0,124
B03	3	1,269	0,093	1,202	0,071	2,212	-0,943	0,117
B02	1	1,902	0,097	-0,239	0,056	0,346	1,556	0,112
B02	6	1,245	0,093	1,814	0,085	3,004	-1,759	0,126

Tabla 4.34 Análisis FDI de las medidas de la prueba censal de Lectura en quechua chanca de 4.º grado de primaria EIB usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
B05	5	0,943	0,032	0,186	0,057	0,908	0,035	0,065
B05	4	0,474	0,033	-0,261	0,056	0,424	0,050	0,065
B05	3	1,083	0,032	0,489	0,058	1,236	-0,153	0,066
BC1	3	-1,045	0,043	-1,741	0,051	-1,179	0,134	0,067
BC1	4	-1,037	0,043	-1,908	0,054	-1,360	0,323	0,069
BC1	1	-2,207	0,065	-2,528	0,067	-2,031	-0,176	0,093
BC3	2	-2,091	0,062	-2,473	0,065	-1,972	-0,119	0,090
B08	1	-0,023	0,035	-0,447	0,057	0,222	-0,245	0,067
B08	2	-0,117	0,035	-0,176	0,056	0,516	-0,633	0,066
B08	3	0,658	0,032	0,298	0,057	1,029	-0,371	0,065
B08	5	0,753	0,032	0,086	0,056	0,800	-0,047	0,064
BC3	1	-2,245	0,065	-3,047	0,082	-2,593	0,348	0,105
BC1	2	-1,507	0,050	-2,130	0,058	-1,600	0,093	0,077
BC4	1	1,435	0,033	0,772	0,043	1,543	-0,108	0,054
BC4	2	1,619	0,033	0,586	0,042	1,341	0,278	0,053
BC4	3	1,269	0,032	0,471	0,041	1,217	0,052	0,052
BC2	1	0,183	0,034	-0,426	0,040	0,245	-0,062	0,052
BC2	2	1,928	0,034	0,629	0,042	1,388	0,540	0,054
BC2	3	1,704	0,033	0,711	0,043	1,477	0,227	0,054
B05	1	0,108	0,034	-0,404	0,057	0,269	-0,161	0,066
B05	2	-0,260	0,036	-0,886	0,060	-0,253	-0,007	0,070

Tabla 4.35 Análisis FDI de las medidas de la prueba censal de Lectura en quechua cusco collao de 4.º grado de primaria EIB usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BR03	2	-1,149	0,043	-0,994	0,056	-0,479	-0,670	0,071
B03	4	-1,075	0,043	-1,029	0,057	-0,526	-0,549	0,071
B03	3	0,098	0,037	-0,398	0,052	0,330	-0,232	0,064
B03	6	0,228	0,036	-0,089	0,051	0,749	-0,521	0,062
BC3	1	-2,547	0,059	-2,932	0,073	-3,108	0,561	0,094
BC3	2	-2,453	0,058	-2,624	0,064	-2,690	0,237	0,086
BC4	1	-0,742	0,040	-1,010	0,039	-0,500	-0,242	0,056
BC4	2	0,301	0,036	-0,166	0,036	0,645	-0,344	0,051
BC4	3	1,325	0,035	0,248	0,036	1,206	0,119	0,050
BC4	4	0,548	0,036	-0,225	0,036	0,564	-0,016	0,051
B06	1	1,171	0,035	0,161	0,051	1,088	0,083	0,062
B08	1	1,371	0,035	0,176	0,051	1,108	0,263	0,062
B08	2	2,095	0,036	0,287	0,052	1,259	0,836	0,063
B08	3	0,658	0,035	-0,202	0,051	0,596	0,062	0,062
BC1	1	-3,005	0,067	-2,911	0,073	-3,079	0,074	0,099
BC1	2	-2,314	0,056	-2,368	0,059	-2,343	0,029	0,081

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC1	3	-1,023	0,042	-1,653	0,046	-1,373	0,350	0,062
BC2	1	1,258	0,035	0,477	0,037	1,517	-0,259	0,051
BC2	2	1,021	0,035	0,001	0,036	0,871	0,150	0,050
BC4	5	1,837	0,035	0,368	0,036	1,369	0,468	0,050
B06	2	1,824	0,035	0,621	0,053	1,712	0,112	0,064
B06	5	1,788	0,035	0,872	0,054	2,053	-0,265	0,064
BC1	4	-0,550	0,039	-1,139	0,041	-0,675	0,125	0,057
B03	5	-0,575	0,039	-0,816	0,055	-0,237	-0,338	0,067
BC2	3	1,420	0,035	0,430	0,037	1,453	-0,033	0,051

Tabla 4.36 Análisis FDI de las medidas de la prueba censal de Lectura en shipibonibo de 4.º grado de primaria EIB usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC03	1	-1,769	0,092	-2,011	0,081	-1,802	0,033	0,123
BC1	2	-1,099	0,077	-1,128	0,070	-0,948	-0,151	0,104
B03	5	0,103	0,071	-0,080	0,098	0,066	0,037	0,121
BC2	1	0,042	0,071	-0,025	0,069	0,119	-0,077	0,099
BC2	5	0,256	0,072	0,639	0,075	0,761	-0,505	0,104
BC4	3	-0,284	0,070	-0,649	0,067	-0,485	0,201	0,097
BC4	4	-0,009	0,070	-0,420	0,067	-0,263	0,254	0,097
B05	2	0,544	0,074	0,027	0,096	0,169	0,375	0,121
B05	5	0,157	0,071	-0,220	0,095	-0,070	0,227	0,119
BC1	1	-1,807	0,094	-1,827	0,078	-1,624	-0,183	0,122
BC3	2	-1,192	0,078	-1,592	0,074	-1,397	0,205	0,108
B03	1	-0,108	0,070	0,096	0,100	0,236	-0,344	0,122
B03	2	-0,536	0,071	-0,448	0,097	-0,290	-0,246	0,120
B03	4	-0,240	0,070	-0,138	0,098	0,010	-0,250	0,120
B03	6	-0,606	0,072	-0,383	0,097	-0,227	-0,379	0,121
BC2	2	0,359	0,072	0,098	0,070	0,238	0,121	0,100
BC2	4	0,166	0,071	0,141	0,070	0,280	-0,114	0,100
BC4	1	0,021	0,070	-0,768	0,068	-0,600	0,621	0,098
BC4	2	0,076	0,071	-0,338	0,068	-0,184	0,260	0,098
B05	6	-0,030	0,070	-0,093	0,096	0,053	-0,083	0,119

Tabla 4.37 Análisis FDI de las medidas de la prueba censal de Lectura en castellano como segunda lengua de 4.º grado de primaria EIB usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC1	2	-1,319	0,021	-1,555	0,019	-1,252	-0,067	0,028
BC3	2	-1,609	0,023	-1,878	0,021	-1,563	-0,046	0,031
BC3	3	-1,410	0,021	-1,658	0,020	-1,351	-0,059	0,029
B05	1	-0,266	0,016	-0,432	0,023	-0,170	-0,096	0,028
B05	3	-0,283	0,016	-0,411	0,023	-0,149	-0,134	0,028
B04	3	0,186	0,016	0,302	0,023	0,538	-0,352	0,028
B04	5	0,334	0,016	0,159	0,023	0,400	-0,066	0,028
B04	4	0,827	0,016	0,773	0,024	0,992	-0,165	0,029
BC3	1	-2,094	0,027	-2,474	0,025	-2,138	0,044	0,037
BC1	3	-1,370	0,021	-1,750	0,020	-1,440	0,070	0,029
BC1	1	-1,547	0,022	-1,885	0,021	-1,570	0,023	0,030
BC3	4	-1,402	0,021	-1,775	0,020	-1,464	0,062	0,029
BC3	5	0,677	0,015	0,198	0,016	0,438	0,239	0,022
B07	1	0,495	0,015	0,449	0,023	0,679	-0,184	0,027
B07	4	-0,230	0,016	-0,577	0,023	-0,309	0,079	0,028
B05	4	0,681	0,015	0,368	0,023	0,601	0,080	0,027
B05	5	0,216	0,016	-0,267	0,023	-0,011	0,227	0,028
B07	3	0,799	0,015	0,364	0,023	0,598	0,201	0,027
B07	5	0,825	0,016	0,451	0,023	0,681	0,144	0,028

Tabla 4.38 Análisis FDI de las medidas de la prueba censal de Lectura de 2.º grado de secundaria usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
B09	1	0,664	0,005	-0,160	0,005	0,729	-0,065	0,007
B09	2	-0,093	0,006	-0,917	0,006	0,038	-0,131	0,008
B09	4	0,715	0,005	-0,126	0,005	0,760	-0,045	0,007
B09	5	0,995	0,005	0,297	0,005	1,146	-0,151	0,007
B03	1	2,240	0,005	1,669	0,005	2,398	-0,158	0,007
B03	4	0,663	0,005	-0,322	0,005	0,581	0,082	0,007
B03	3	1,579	0,005	0,483	0,005	1,315	0,264	0,007
B03	5	0,882	0,005	0,150	0,005	1,012	-0,130	0,007
B06	1	0,043	0,006	-0,994	0,006	-0,032	0,075	0,008
B06	2	1,887	0,005	0,968	0,005	1,758	0,129	0,007
B06	3	1,019	0,005	0,161	0,005	1,022	-0,003	0,007
B06	4	1,337	0,005	0,419	0,005	1,257	0,080	0,007
B06	6	1,431	0,005	0,547	0,005	1,374	0,057	0,007
B09	3	1,145	0,005	0,336	0,005	1,181	-0,036	0,007
B03	6	0,498	0,005	-0,595	0,005	0,332	0,166	0,007
B06	5	2,289	0,006	1,546	0,005	2,285	0,004	0,008
B08	6	0,190	0,005	-0,615	0,005	0,314	-0,124	0,007
BC2	2	0,900	0,003	0,019	0,003	0,892	0,008	0,004

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
BC2	3	2,476	0,004	1,870	0,004	2,581	-0,105	0,006
BC2	4	1,904	0,003	1,067	0,003	1,848	0,056	0,004
BC2	7	1,338	0,033	0,388	0,028	1,229	0,109	0,043
BC1	12	0,700	0,005	-0,277	0,003	0,622	0,078	0,006
BC2	6	1,825	0,003	1,010	0,003	1,796	0,029	0,004
B09	6	0,603	0,005	-0,206	0,005	0,687	-0,084	0,007
BC1	8	0,298	0,005	-0,687	0,003	0,248	0,050	0,006
BC1	10	0,251	0,005	-0,872	0,004	0,079	0,172	0,006
BC1	11	0,530	0,005	-0,486	0,003	0,431	0,099	0,006
B08	1	-1,574	0,009	-2,610	0,009	-1,507	-0,067	0,013
B08	2	-0,090	0,006	-1,002	0,006	-0,040	-0,050	0,008
B08	3	0,734	0,005	0,093	0,005	0,960	-0,226	0,007
B08	4	1,878	0,005	1,207	0,005	1,976	-0,098	0,007
B08	5	1,331	0,005	0,486	0,005	1,318	0,013	0,007

Tabla 4.39 Análisis FDI de las medidas de la prueba censal de Matemática de 2.º grado de secundaria usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
B02	3	1,980	0,007	0,760	0,007	1,917	0,063	0,010
B03	2	1,354	0,007	0,081	0,007	1,261	0,093	0,010
BC2	5	0,316	0,003	-0,924	0,003	0,290	0,026	0,004
B01	4	1,786	0,003	0,553	0,007	1,717	0,069	0,008
B10	5	2,028	0,003	0,814	0,007	1,969	0,059	0,008
B09	3	0,705	0,003	-0,480	0,007	0,719	-0,014	0,008
B04	4	2,790	0,009	1,499	0,008	2,630	0,160	0,012
B04	5	0,774	0,003	-0,501	0,007	0,699	0,075	0,008
BC1	1	-0,429	0,004	-1,509	0,004	-0,274	-0,155	0,006
BC1	2	-0,084	0,004	-1,366	0,004	-0,136	0,052	0,006
B09	1	1,706	0,007	0,873	0,007	2,026	-0,320	0,010
BC1	5	3,087	0,009	2,001	0,004	3,115	-0,028	0,010
B03	3	0,696	0,003	-0,725	0,007	0,483	0,213	0,008
B07	2	1,274	0,007	0,097	0,007	1,276	-0,002	0,010
BC1	18	0,779	0,003	-0,342	0,003	0,853	-0,074	0,004
BC2	12	0,080	0,008	-1,065	0,003	0,154	-0,074	0,009
BC2	1	-0,376	0,004	-1,633	0,004	-0,394	0,018	0,006
B02	2	0,573	0,007	-0,548	0,007	0,654	-0,081	0,010
BC1	10	1,180	0,007	0,282	0,003	1,455	-0,275	0,008
B06	4	0,608	0,003	-0,480	0,007	0,719	-0,111	0,008
B09	4	2,540	0,004	1,390	0,008	2,525	0,015	0,009
B10	1	0,029	0,008	-1,017	0,008	0,201	-0,172	0,011
B01	2	1,315	0,007	-0,040	0,007	1,144	0,171	0,010
B08	4	-0,375	0,009	-1,909	0,009	-0,661	0,286	0,013
BC1	6	0,830	0,007	-0,408	0,003	0,789	0,041	0,008
BC2	13	0,246	0,003	-0,933	0,003	0,282	-0,036	0,004

Tabla 4.40 Análisis FDI de las medidas de la prueba censal de Ciencias Sociales de 2.º grado de secundaria usadas en la equiparación de 2016 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
B03	2	-2,228	0,008	-2,122	0,008	-2,102	-0,126	0,011
B05	7	-0,428	0,006	-0,509	0,005	-0,464	0,036	0,008
B05	6	-0,806	0,006	-0,860	0,006	-0,820	0,014	0,008
B01	14	-1,442	0,007	-1,424	0,006	-1,393	-0,049	0,009
B02	7	-0,541	0,006	-0,672	0,005	-0,629	0,088	0,008
B04	10	-0,554	0,006	-0,244	0,005	-0,194	-0,360	0,008
B04	11	-0,114	0,005	0,005	0,005	0,058	-0,172	0,007
B06	11	-1,355	0,006	-1,727	0,007	-1,701	0,346	0,009
B06	12	-0,329	0,006	-0,291	0,005	-0,242	-0,087	0,008
B01	8	-0,246	0,006	-0,577	0,005	-0,533	0,287	0,008
B03	14	-0,850	0,006	-1,096	0,006	-1,060	0,210	0,008
B02	1	0,403	0,005	0,170	0,005	0,226	0,177	0,007
B04	9	-0,930	0,006	-1,003	0,006	-0,965	0,035	0,008
B01	6	0,731	0,006	0,685	0,005	0,749	-0,018	0,008
B05	4	0,200	0,005	-0,229	0,005	-0,179	0,379	0,007
B03	13	0,203	0,005	0,269	0,005	0,327	-0,124	0,007
B05	13	0,189	0,005	0,310	0,005	0,368	-0,179	0,007
B01	2	1,109	0,006	1,150	0,006	1,221	-0,112	0,008
B01	5	0,911	0,006	0,584	0,005	0,646	0,265	0,008
B03	5	0,639	0,005	0,416	0,005	0,476	0,163	0,007
B06	6	0,469	0,005	0,684	0,005	0,748	-0,279	0,007
B03	8	0,589	0,006	0,538	0,005	0,600	-0,011	0,008
B04	3	-0,419	0,006	-0,413	0,005	-0,366	-0,053	0,008
B05	12	-1,027	0,006	-0,739	0,006	-0,697	-0,330	0,008
B06	7	0,117	0,005	0,096	0,005	0,151	-0,034	0,007
B05	14	-0,660	0,006	-0,436	0,005	-0,389	-0,271	0,008
B04	5	0,343	0,005	0,084	0,005	0,139	0,204	0,007

Tabla 4.41 Constantes de equiparación para las pruebas aplicadas en las evaluaciones censales

Prueba	a	b
Lectura, 4.º grado de primaria	1,137	-0,769
Matemática, 4.º grado de primaria	0,924	-0,869
Lectura en aimara, 4.º grado de primaria EIB	0,787	0,237
Lectura en awajún, 4.º grado de primaria EIB	1,655	0,961
Lectura en ashaninka, 4.º grado de primaria EIB	1,034	0,246
Lectura en quechua cusco collao, 4.º grado de primaria EIB	1,314	0,840
Lectura en quechua chanca, 4.º grado de primaria EIB	1,057	0,694
Lectura en shipibo-konibo, 4.º grado de primaria EIB	1,018	0,162
Lectura en castellano como L2, 4.º grado de primaria EIB	0,964	0,247
Lectura, 2.º grado de secundaria	0,912	0,875
Matemática, 2.º grado de secundaria	0,966	1,183
Ciencias Sociales, 2.º grado de secundaria	1,016	0,053
Ciencia y Tecnología, 2.º grado de secundaria	1	0

No se realizó el análisis FDI de las medidas de la prueba de Ciencia y Tecnología, ya que es la primera vez que esta prueba ha sido evaluada.

Tabla 4.42 Análisis FDI de las medidas de la prueba muestral de Ciudadanía de 6.º grado de primaria usadas en la equiparación de 2013 con 2018

Bloque	Posición en el bloque	2013		2018		Eqp Medida Trans	FDI	
		Medida	Error	Medida	Error		Medida	Error
B06	8	-1,367	0,022	-1,615	0,065	-1,347	-0,020	0,069
B07	9	-1,310	0,022	-1,351	0,060	-1,123	-0,187	0,064
B02	6	0,241	0,018	-0,046	0,046	-0,015	0,256	0,049
B04	6	0,661	0,018	0,012	0,046	0,034	0,627	0,049
B04	7	0,300	0,018	0,120	0,045	0,126	0,174	0,048
B02	15	-0,427	0,019	0,094	0,046	0,104	-0,531	0,050
B08	11	-0,523	0,019	-0,792	0,053	-0,648	0,125	0,056
B03	13	0,133	0,018	-0,019	0,046	0,008	0,125	0,049
B01	3	-0,008	0,019	-0,049	0,046	-0,017	0,009	0,050
B04	15	0,960	0,018	0,613	0,044	0,545	0,415	0,048
B03	14	0,051	0,018	-0,027	0,046	0,001	0,050	0,049
B04	8	-1,385	0,023	-1,837	0,069	-1,536	0,151	0,073
B05	7	-0,389	0,019	-0,880	0,054	-0,723	0,334	0,057
B03	6	-0,592	0,020	-0,625	0,050	-0,507	-0,085	0,054
B06	3	-1,774	0,025	-2,119	0,075	-1,776	0,002	0,079
B02	8	0,130	0,019	1,202	0,045	1,045	-0,915	0,049
B06	4	-0,925	0,021	-0,824	0,053	-0,676	-0,249	0,057
B01	15	-1,285	0,022	-1,410	0,059	-1,173	-0,112	0,063
B08	13	-0,518	0,020	0,056	0,046	0,072	-0,590	0,050

Bloque	Posición en el bloque	2013		2018		Eqp	FDI	
		Medida	Error	Medida	Error	Medida Trans	Medida	Error
B08	12	1,099	0,019	1,162	0,045	1,011	0,088	0,049
B06	6	-2,228	0,027	-2,730	0,094	-2,294	0,066	0,098
B08	4	-1,094	0,021	-1,467	0,062	-1,222	0,128	0,065
B05	9	-0,642	0,020	-1,158	0,057	-0,959	0,317	0,060
B08	1	-0,486	0,018	-0,404	0,049	-0,319	-0,167	0,052
B01	2	-0,546	0,019	-0,438	0,048	-0,348	-0,198	0,052
B02	12	0,703	0,018	0,415	0,045	0,377	0,326	0,048
B07	14	0,950	0,018	1,037	0,044	0,905	0,045	0,048
B06	9	-0,957	0,020	-1,517	0,063	-1,264	0,307	0,066
B03	2	-0,467	0,018	-0,380	0,048	-0,299	-0,168	0,051
B04	14	-0,325	0,018	-0,474	0,049	-0,378	0,053	0,052
B04	4	-0,031	0,018	0,379	0,044	0,346	-0,377	0,048

Tabla 4.43 Constantes de equiparación para las pruebas aplicadas en las evaluaciones muestrales

Prueba	a	b
Ciudadanía, 6.º grado de primaria	0,901	0,126

4.5 Niveles de logro y presentación de resultados

El diseño de las evaluaciones permite estimar lo que saben y hacen los estudiantes a partir de su desempeño en las pruebas, con respecto a lo que deberían saber y hacer. De acuerdo con ello, la interpretación de los resultados de la ECE está referida a criterios; es decir, los resultados se interpretan considerando un criterio o estándar de logro (Muñiz, 1998). Para esto, es necesario establecer puntos de corte, lo que constituye un procedimiento estándar utilizado para delimitar numéricamente dos o más niveles de desempeño de una competencia (Cizek y Bunch, 2007). Estos niveles de desempeño, conceptualmente, corresponden a un juicio compartido por una comunidad de expertos (docentes, evaluadores, tomadores de decisiones) que determinan qué es lo que un estudiante mínimamente competente debe ser capaz de hacer para ser considerado parte de un determinado nivel en pruebas referidas a criterios (Cizek, 1993; Shepard, 1980).

Este procedimiento se realiza por única vez al inicio de un ciclo de evaluaciones. Ello tiene por intención sostener los mismos puntos de corte en las ediciones posteriores y asegurar que los resultados sean comparables en el tiempo. Este año, fue utilizado para determinar los niveles de la prueba de Ciencia y Tecnología.

Se utiliza el método *bookmark* (Cizek y Bunch, 2007; Karantonis y Sireci, 2006; Lewis, Mitzel y Green, 1996), que es consistente con el modelo de medición utilizado por la UMC en las distintas evaluaciones que conduce. Como consecuencia, el juicio sobre los cortes considera también los ítems de respuesta construida y con

créditos parciales, dado que estos se incluyen en la misma métrica que los de opción múltiple. El método *bookmark*, en términos generales, consiste en colocar marcas (tantas como cortes se hayan preestablecido) en un cuadernillo de ítems ordenado por dificultad (estando esta dificultad estimada mediante procedimientos Rasch o de teoría de respuesta al ítem). La pregunta típica que guía el establecimiento de cortes es la siguiente: “¿Hasta qué ítem debe ser capaz de resolver un estudiante, como mínimo, para ser considerado parte del nivel?”. El procedimiento establece que los jueces, organizados en grupos pequeños, determinan, en tres rondas, los cortes para cada nivel de desempeño. De no llegar a un acuerdo, se aplican procedimientos estadísticos para resolver la discrepancia.

En general, un taller se conduce en tres rondas (Cizek y Bunch, 2007). En la primera ronda, los participantes leen las descripciones de los niveles de logro elaboradas por el equipo de la UMC, resuelven todos los ítems de la prueba y analizan las razones por las cuales un ítem era más difícil que el anterior. La primera ronda concluye con un primer establecimiento individual de cortes. En la segunda ronda, los participantes exponen, en los subgrupos, las razones que los motivaron a colocar sus cortes. Asimismo, se les entrega un reporte de discrepancias en el que señalan qué tan distintos son sus juicios con respecto de los demás grupos. La segunda ronda finaliza con un segundo establecimiento individual de cortes. En la tercera ronda, los participantes tienen acceso al impacto de la data, es decir, a la distribución de personas en los distintos niveles de desempeño, si los resultados de la segunda ronda hubiesen sido los definitivos. Con esta información, los jueces emiten su tercer y último corte.

Como fue mencionado con anterioridad, el modelo Rasch estima de manera conjunta la dificultad de los ítems y las medidas de habilidad de las personas. Utilizando dicha información, es posible calcular la probabilidad de un estudiante con una habilidad específica de responder correctamente un ítem que posee una dificultad concreta (P_{ni}).

A partir de lo trabajado en la *Evaluación Nacional 2004* (Ministerio de Educación, 2005), se establece que un estudiante está en un nivel de logro si tiene una probabilidad mayor o igual a 0,62 para responder correctamente el ítem que marca el corte entre dos niveles de logro consecutivos. En este sentido, es importante considerar que la probabilidad de 0,62 para responder correctamente un ítem implica una medida de habilidad que sea medio *logit* más que la medida de dificultad de dicho ítem. Después de esta adición, estos valores son transformados a medida 500 (M500) y redondeados a dos decimales. Finalmente, se asigna a cada estudiante, de acuerdo a su M500, a un nivel de logro determinado.

Se presentan, a continuación, los puntos de corte (en M500) de las evaluaciones censales y muestrales:

Tabla 4.44 Medidas (en M500) que marcan los puntos de corte de los alumnos en las evaluaciones censales

Área	<Nivel 1 vs. Nivel 1	Nivel 1 vs. Nivel 2	Nivel 2 vs. Nivel 3
Lectura, 4.º grado de primaria	356,92	444,72	522,03
Matemática, 4.º grado de primaria	351,90	422,21	526,46
Lectura en aimara, 4.º grado de primaria EIB	494,18	686,00	—
Lectura en awajún, 4.º grado de primaria EIB	550,58	684,97	—
Lectura en ashaninka, 4.º grado de primaria EIB	510,90	620,47	—
Lectura en quechua cusco collao, 4.º grado de primaria EIB	482,01	611,07	—
Lectura en quechua chanca, 4.º grado de primaria EIB	443,16	597,48	—
Lectura en shipibo-konibo, 4.º grado de primaria EIB	603,61	703,65	—
Lectura en castellano como L2, 4.º grado de primaria EIB	514,03	604,77	—
Lectura, 2.º grado de secundaria	505,14	580,61	641,25
Matemática, 2.º grado de secundaria	519,67	595,96	649,38
Ciencias Sociales, 2.º grado de secundaria	425,93	500,05	607,28
Ciencia y Tecnología, 2.º grado de secundaria	374,60	509,58	628,48

Tabla 4.45 Medidas (en M500) que marcan los puntos de corte de los alumnos en las evaluaciones muestrales

Área	<Nivel 1 vs. Nivel 1	Nivel 1 vs. Nivel 2	Nivel 2 vs. Nivel 3
Ciudadanía, 6.º grado de primaria	405,45	478,42	553,67
Escritura, 2.º grado de secundaria	425,74	619,37	—

Finalmente, es importante señalar que las medidas de habilidad, así como los puntos de corte antes mencionados, de las personas fueron transformadas linealmente ($Y = a + bX$), de tal manera que, en la primera evaluación que se realizó en cada grado y área, su media aritmética fue 500 y su desviación estándar 100. Para ello, se usaron las constantes que aparecen a continuación:

Tabla 4.46 Valores de a y b utilizados para la transformación lineal de las medidas de las pruebas censales

Prueba	a	b
Lectura, 4.º grado de primaria	474,896	73,414
Matemática, 4.º grado de primaria	463,157	73,781
Lectura en aimara, 4.º grado de primaria EIB	467,700	137,900
Lectura en awajún, 4.º grado de primaria EIB	508,900	123,300
Lectura en ashaninka, 4.º grado de primaria EIB	506,688	89,664
Lectura en quechua cusco collao, 4.º grado de primaria EIB	439,400	101,700
Lectura en quechua chanca, 4.º grado de primaria EIB	410,500	100,800
Lectura en shipibo-konibo, 4.º grado de primaria EIB	550,100	131,800
Lectura en castellano como L2, 4.º grado de primaria EIB	472,600	85,600
Lectura, 2.º grado de secundaria	474,896	73,414
Matemática, 2.º grado de secundaria	463,157	73,781
Ciencias Sociales, 2.º grado de secundaria	476,741	100,028
Ciencia y Tecnología, 2.º grado de secundaria	496,120	134,513

Tabla 4.47 Valores de a y b utilizados para la transformación lineal de las medidas de las pruebas muestrales

Prueba	a	b
Ciudadanía, 6.º grado de primaria	465,498	76,007
Escritura, 2.º grado de secundaria	492,931	141,747

Referencias

Referencias

- American Educational Research Association, American Psychological Association y National Council of Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Bond, T. G. y Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human Sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- Charter, R. A. (2003). A breakdown of reliability coefficients by test type and reliability method, and the clinical implications of low reliability. *The Journal of General Psychology*, 130(3), 290-304. <https://doi.org/10.1080/00221300309601160>
- Cizek, G. J. (1993). Reconsidering standards and criteria. *Journal of Educational Measurement*, 30(2), 93-106. <https://doi.org/10.1111/j.1745-3984.1993.tb01068.x>
- Cizek, G. J. y Bunch, M. B. (2007). *Standard settings*. California, CA: SAGE.
- Embretson, S. y Reise, S. (2000). *Item Response Theory for psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Fisher, W. (1992). Reliability statistics. En J. M. Linacre (Ed.), *Rasch Measurement Transactions Part 2, 1996* (p. 238). Chicago, IL: MESA Press.
- Fox, C. (1999). An introduction to the Partial Credit model for developing nursing assessments. *Journal of Nursing Education*, 38(8), 340-346.
- Hambleton, R., Swaminathan, H. y Rogers, J. (1991). *Fundamentals of Item Response Theory*. Thousand Oaks, CA: SAGE.
- Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9(2), 139-164. <https://doi.org/10.1177/014662168500900204>
- Ho, C. y Osborn, S. E. (2005). Test equating by common items and common subjects: Concepts and applications. *Practical Assessment, Research & Evaluation*, 10(4), 1-19.
- Ingebo, G. (1997). *Probability in the measure of achievement*. Chicago, IL: MESA.
- Karantonis, A. y Sireci, S. G. (2006). The bookmark standard-setting method: A literature review. *Educational Measurement: Issues and Practice*, 25(1), (4-12). <https://doi.org/10.1111/j.1745-3992.2006.00047.x>
- Kim, D., Choi, S., Um, K. y Kim, J. (2006, Abril). *A comparison of methods for estimating classification consistency*. Trabajo presentado en el Annual Meeting of the National Council on Education in Measurement, San Francisco, CA.
- Kolen, M. y Brennan, R. (2004). *Test equating, scaling and linking* (2nd ed.). New York, NY: Springer.
- Lee, W. (2010). Classification consistency and accuracy for complex assessments using Item Response Theory. *Journal of Educational Measurement*, 47(1), 1-17. <https://doi.org/10.1111/j.1745-3984.2009.00096.x>

- Lehtonen, R. y Pahkinen, E. (2004). *Practical methods for design and analysis of complex surveys* (2nd ed.). Chichester: John Wiley and Sons.
- Lewis, D. M., Mitzel, H. C. y Green, D. R. (1996, Junio). Standard setting: A bookmark approach. En D. R. Green (Ed.), *Irt-based standard-setting procedures utilizing behavioral anchoring*. Simposio presentado en el Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.
- Linacre, J. M. (1998). Detecting multidimensionality: Which residual data-type works best? *Journal of Outcome Measurement*, 2(3), 266-283.
- Linacre, J. M. (2001). Category, step and threshold: Definitions & disordering. *Rasch Measurement Transactions*, 15(1), 794.
- Linacre, J. M. (2019). *Winsteps manual*. Recuperado de <https://www.winsteps.com/a/Winsteps-Manual.pdf>
- Livingston, S. A. (2004). *Equating test scores (without IRT)*. New Jersey, NJ: Educational Testing Service.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, 47(2), 149-174. <https://doi.org/10.1007/BF02296272>
- McGartland, D., Berg-Weger, M., Tebb, S. S., Lee, E. S. y Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research*, 27(2), 94-104. <https://doi.org/10.1093/swr/27.2.94>
- Ministerio de Educación. (2005). *Evaluación nacional del rendimiento estudiantil 2004. Informe descriptivo de resultados*. Lima: Oficina de Medición de la Calidad de los Aprendizajes.
- Ministerio de Educación. (2009). *Evaluación Censal de Estudiantes (ECE). Segundo grado de primaria y cuarto grado de primaria de IE EIB. Marco de trabajo*. Lima: Oficina de Medición de la Calidad de los Aprendizajes.
- Ministerio de Educación. (2015). *Marco de fundamentación de las pruebas de rendimiento de la Evaluación Censal de Estudiantes de 2.º grado de secundaria 2015*. Lima: Oficina de Medición de la Calidad de los Aprendizajes.
- Ministerio de Educación. (2019). *Reporte técnico de la Evaluación Muestral de Estudiantes (EM 2018) de 2.º grado de primaria*. Lima: Oficina de Medición de la Calidad de los Aprendizajes.
- Moreno, R., Martínez, R. J. y Muñoz, J. (2004). Directrices para la construcción de ítems de elección múltiple. *Psicothema*, 16(3), 490-497.
- Muñoz, J. (1998). *Teoría clásica de los tests* (2.ª ed.). Madrid: Ediciones Pirámide.
- Navas, M. (1996). Equiparación de puntuaciones. En J. Muñoz (Ed.), *Psicometría* (pp. 293-370). Madrid: Universitas.
- Nunnally, J. y Bernstein, I. (n.d.). *Teoría psicométrica* (3.ª ed.).
- Park, H. S., Dailey, R. y Lemus, D. (2002). The use of exploratory factor analysis and principal components analysis in communication research. *Human*

Communication Research, 28(4), 562-577.
<https://doi.org/10.1111/j.1468-2958.2002.tb00824.x>

- Paz, M. (1996). Validez. En J. Muñiz (Ed.), *Psicometría* (pp. 49-103). Madrid: Universitas.
- Preacher, K. J. y MacCallum, R. C. (2003). Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics*, 2(1), 13-43.
https://doi.org/10.1207/S15328031US0201_02
- Schulz, E. M. (1990). Functional assessment of fit. En J. M. Linacre (Ed.), *Rasch Measurement Transactions*, 1990 (p. 82). Chicago, IL: MESA Press.
- Schumacker, R. E. y Smith, E. V. (2007). Reliability: A Rasch perspective. *Educational and Psychological Measurement*, 67(3), 394-409.
<https://doi.org/10.1177/0013164406294776>
- Shepard, L. (1980). Standard setting issues and methods. *Applied Psychological Measurement*, 4(4), 447-467.
<https://doi.org/10.1177/014662168000400403>
- Smith, R. y Kramer, G. (1989). Response pattern analysis with supplemental store reports. En J. M. Linacre (Ed.), *Rasch Measurement Transactions Part 1, 1995* (pp. 33-35). Chicago, IL: MESA Press.
- Suen, H. (1990). *The psychometric process*. Pennsylvania, PA: Pennsylvania State University.
- Wilson, M. (2005). *Constructing measures. An item response modeling approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wright, B. D. (1999). Model selection: Rating Scale Model (RSM) or Partial Credit Model (PCM)? *Rasch Measurement Transactions*, 12(3), 641-642.
- Wright, B. D. y Linacre, J. M. (1989). Differences between scores and measures. En J. M. Linacre (Ed.), *Rasch Measurement Transactions Part 1, 1995* (pp. 63-65). Chicago, IL: MESA Press.
- Wright, B. D. y Linacre, J. M. (1994). Reasonable mean-square fit values. En J. M. Linacre (Ed.), *Rasch Measurement Transactions*, 1994 (p. 370). Chicago, IL: MESA Press.
- Wright, B. D. y Stone, M. (1998). *Diseño de mejores pruebas*. México: CENEVAL.
- Wright, B. D. y Masters, G. (1982). *Rating scale analysis*. Chicago, IL: MESA.
- Zhu, W. (1998). Test equating: What, why, how? *Research Quarterly for Exercise and Sport*, 69(1), 11-23. <https://doi.org/10.1080/02701367.1998.10607662>

Ministerio de Educación

**Calle Del Comercio 193,
San Borja - Lima, Perú
Tel.: (511) 615-5800**

<http://www.minedu.gob.pe/>



PERÚ

Ministerio
de Educación