

COMPETENCIA
LECTURA
MATEMÁTICA
CIENCIA
SOCIOEMOCIONAL
MEDICIÓN
DESEMPEÑO
PSICOMETRÍA
APRENDIZAJE
EDUCACIÓN
CALIDAD

Reporte técnico de la Evaluación Muestral de Estudiantes 2022

COMPETENCIA
LECTURA
MATEMÁTICA
CIENCIA
SOCIOEMOCIONAL
MEDICIÓN
DESEMPEÑO
PSICOMETRÍA
APRENDIZAJE
EDUCACIÓN
CALIDAD

Reporte técnico de la Evaluación Muestral de Estudiantes 2022



PERÚ

Ministerio
de Educación

Oscar Manuel Becerra Tresierra

Ministro de Educación del Perú

Magnet Carmen Márquez Ramírez

Viceministro de Gestión Institucional

Miriam Janette Ponce Vertiz

Viceministra de Gestión Pedagógica

Luis Humberto Náñez Aldaz

Secretario de Planificación Estratégica

Gloria María Zambrano Rozas

Jefa de la Oficina de Medición de la Calidad de los Aprendizajes

Reporte técnico de la Evaluación Muestral de estudiantes 2022

Responsables del documento

Andrés Alberto Burga León (Coordinador)

Tania Pacheco Valenzuela

Yoni Arámbulo Mogollón

Rosina Vanessa Sánchez Jiménez

Luis Mejía Campos

Gabriela Santibáñez Rojas

Wilmer Hernández Cabrera

Marco Carranza Rodríguez

Alejandro Romani Valdivieso

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, 2023

Calle Del Comercio 193, San Borja

Lima, Perú

Teléfono: (511) 615-5800

www.minedu.gob.pe

Todos los derechos son 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	7
Capítulo 1: Diseño y construcción de las pruebas aplicadas	9
1.1 Construcción de los ítems	12
1.2 Procesos de validación aplicados.....	16
1.2.1 Validación de expertos de la UMC.....	16
1.2.2 Revisión de ítems y consulta con la población objetivo.....	17
1.2.3 Validación de campo: aplicación piloto.....	17
1.2.4 Análisis de las propiedades psicométricas y revisión de aspectos pedagógicos	18
1.2.5 Validación de expertos de otras oficinas del Minedu.....	19
1.2.6 Adaptación de instrumentos para estudiantes con discapacidad	19
1.2.7 Aplicación definitiva.....	22
1.3 Ensamblaje de las pruebas mediante el diseño bloques	23
Capítulo 2: Población y muestra	46
2.1 Población objetivo	47
2.2 Marco muestral.....	47
2.3 Diseño muestral.....	48
2.3.1 Conglomerados	48
2.3.2 Estratificación	49
2.4 Tamaño muestral y cobertura.....	49
2.5 Pesos muestrales	53
2.6 Validación de la muestra seleccionada.....	55
Capítulo 3: Operativo de campo	64
3.1 Organización territorial del operativo	66
3.2 Conformación de la red administrativa de aplicación.....	67
3.3 Proceso de selección y capacitación del personal de la RA.....	70
3.4 Características de los instrumentos y documentos utilizados	73
3.5 Proceso de aplicación.....	75

Capítulo 4: Procesamiento de datos y análisis psicométrico	80
4.1 Gestión y depuración de datos	81
4.1.1 Gestión física	81
4.1.2 Digitalización y captura de datos	82
4.1.3 Depuración de datos	83
4.2 Estrategias de análisis psicométrico	84
4.2.1 Modelo Rasch para ítems dicotómicos	84
4.2.2 Modelo Rasch de crédito parcial	87
4.2.3 Análisis de ítems	90
4.2.4 Evidencias de validez vinculadas a la estructura interna de las pruebas	92
4.2.5 Confiabilidad y consistencia de la clasificación	93
4.3 Resultados del análisis psicométrico en Lectura, Matemática, y Ciencia y Tecnología	94
4.3.1 Calibración de los ítems	94
4.3.2 Evidencias de validez vinculadas a la estructura interna	133
4.3.3 Confiabilidad y consistencia de la clasificación	134
4.4 Equiparación de medidas	137
4.5 Niveles de logro y preparación de resultados	151
4.6 Resultados del análisis psicométrico de habilidades socioemocionales	154
4.6.1 Calibración de los ítems	155
4.6.2 Estructura de umbrales	162
4.6.3 Evidencias de validez vinculadas a la estructura interna	164
4.6.4 Confiabilidad	165
 Referencias	 168
 Anexos	 172

Introducción

El Ministerio de Educación (Minedu), a través de la Oficina de Medición de la Calidad de los Aprendizajes (UMC) diseña e implementa las Evaluaciones Nacionales de Logros de Aprendizaje. Mediante ellas se obtiene información confiable que puede ser utilizada para mejorar los aprendizajes en las áreas y competencias evaluadas tanto por el sistema educativo como por todos sus actores (familias, directores, docentes, especialistas UGEL, DRE/GRE y Minedu). Estas evaluaciones contribuyen a identificar cuáles son los aprendizajes que nuestros estudiantes han logrado en relación a lo que el Currículo Nacional de Educación Básica (CNEB) espera para el grado.

En marzo de 2020, se declaró la emergencia sanitaria por la COVID-19 a nivel nacional, la cual continuó hasta el 2022. Esta emergencia sanitaria originó condiciones atípicas en el sistema educativo. Luego de esa coyuntura atípica en que se desarrolló el servicio educativo, el Minedu a través de la Resolución Ministerial N° 531-2021-MINEDU aprueba el Documento Normativo denominado “Disposiciones para el retorno a la presencialidad y/o semipresencialidad, así como la prestación del servicio educativo para el año escolar 2022 en instituciones y programas educativos de la Educación Básica, ubicadas en los ámbitos urbano y rural, en el marco de la emergencia sanitaria por la COVID-19”, el cual dispuso que el servicio educativo para el año 2022 iniciaría el 14 de marzo y el tipo de prestación de este servicio se desarrollaría por modalidad presencial o semipresencial.

En ese marco, y conforme a la RM 074-2022-Minedu, la UMC llevó a cabo en el 2022 la Evaluación Muestral de Estudiantes (EM 2022), la misma que se aplicó de manera presencial en una muestra representativa de escuelas a nivel nacional.

La EM 2022 evaluó las áreas de Comunicación (Lee diversos tipos de textos escritos en su lengua materna) y Matemática (Resuelve problemas de cantidad) en 2. ° grado de primaria; Comunicación (Lee diversos tipos de textos escritos en su lengua materna) y Matemática (todas las competencias) en 4. ° grado de primaria y 6. ° grado de primaria. Asimismo, evaluó las áreas de Comunicación (Lee diversos tipos de textos escritos en su lengua materna), Matemática (todas las competencias) y Ciencia y Tecnología (todas las competencias) en 2. ° grado de secundaria. Además, se aplicó un cuestionario de Habilidades socioemocionales a los estudiantes de 6. ° grado de primaria y 2. ° grado de secundaria. Esta información fue complementada con cuestionarios sobre factores contextuales aplicados a directores, docentes, familias y estudiantes.

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 las aplicaciones 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 cómo se organizó la red administrativa, cómo fueron seleccionados y capacitados los aplicadores, cuáles fueron los instrumentos utilizados y cómo se recolectaron los datos en el contexto de las evaluaciones 2022.

Finalmente, el cuarto capítulo describe el proceso de gestión y depuración de datos, así como el modelo Rasch utilizado para los análisis psicométricos. Se pone énfasis en 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 y la forma en que se establecieron los niveles de logro a partir de los puntos de corte.

Diseño y construcción de las pruebas aplicadas

Capítulo 1

La Oficina de Medición de la Calidad de los Aprendizajes (UMC) cuenta con un equipo de evaluación conformado por distintos subequipos organizados por especialidades: Comunicación, Matemática, Lenguas originarias, Ciudadanía y Ciencias Sociales, Ciencia y Tecnología. Además, cuenta con un equipo de Habilidades Socioemocionales. 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 técnica del Ministerio de Educación del Perú (Minedu). Además, la oficina cuenta con un subequipo de Atención a la Diversidad que orienta y contribuye en la aplicación del diseño universal de evaluación en la construcción de todas las pruebas, así como en la adaptación de estas y de los procedimientos de aplicación 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 de forma autónoma.

Cada año, las pruebas que se aplican en el marco de las evaluaciones nacionales de logros de aprendizaje con alcance censal o muestral son especificadas mediante una resolución ministerial publicada en los primeros meses del año en el diario oficial El Peruano. Es así que, en 2022, mediante la Resolución Ministerial 074-2022-Minedu, se aprobó la implementación de la Evaluación Muestral de Estudiantes tal como se detalla en la tabla 1.1.

Tabla 1.1 Evaluaciones muestrales implementadas por la UMC en el 2022

Nivel	Grado	Área	Competencia(s) evaluada(s)	Denominación de la prueba
Primaria	2.º	Comunicación	Lee diversos tipos de textos escritos en su lengua materna.	Lectura
		Matemática	Resuelve problemas de cantidad.	Matemática

Nivel	Grado	Área	Competencia(s) evaluada(s)	Denominación de la prueba	
Primaria	4.º	Comunicación	Lee diversos tipos de textos escritos en su lengua materna.	Lectura	
		Matemática	Resuelve problemas de cantidad.	Matemática	
			Resuelve problemas de regularidad, equivalencia y cambio.		
			Resuelve problemas de gestión de datos e incertidumbre.		
	6.º	Matemática	Resuelve problemas de forma, movimiento y localización.	Matemática	
			Resuelve problemas de cantidad.		
			Resuelve problemas de regularidad, equivalencia y cambio.		
		Habilidades Socioemocionales	Autonomía y relación, Autoeficacia, Autorregulación, Empatía, Responsabilidad y Toma de decisiones	Habilidades Socioemocionales	
	Secundaria	2.º	Comunicación	Lee diversos tipos de textos escritos en su lengua materna.	Lectura
			Matemática	Resuelve problemas de cantidad.	Matemática
Resuelve problemas de regularidad, equivalencia y cambio.					
Resuelve problemas de gestión de datos e incertidumbre.					
	Resuelve problemas de forma, movimiento y localización.				

Nivel	Grado	Área	Competencia(s) evaluada(s)	Denominación de la prueba
Secundaria	2.º	Ciencia y Tecnología	Indaga mediante métodos científicos para construir sus conocimientos.	Ciencia y Tecnología
			Explica el mundo físico basándose en conocimientos sobre los seres vivos, materia y energía, biodiversidad, Tierra y universo.	
		Habilidades Socioemocionales	Diseña y construye soluciones tecnológicas para resolver problemas de su entorno Autonomía y relación, Autoeficacia, Autorregulación, Empatía, Responsabilidad y Toma de decisiones.	Habilidades Socioemocionales

1.1 Construcción de los ítems

Los procesos de diseño y la construcción tanto de los ítems como de las pruebas tienen como referente fundamental el Currículo Nacional de Educación Básica (CNEB, aprobado mediante Resolución Ministerial N.º 281-2016-MINEDU), y siguen los principios de validez, confiabilidad e imparcialidad. Por un lado, tomar como referente el currículo involucra la evaluación de las competencias establecidas en este documento, desde la perspectiva del enfoque propio de cada área: el enfoque comunicativo (y sociocultural) del área de Comunicación, el enfoque de resolución de problemas del área de Matemática y el enfoque de indagación científica y alfabetización científica y tecnológica del área de Ciencia y Tecnología. Por otro lado, seguir los principios de validez, confiabilidad e imparcialidad implica 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 en 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, los instrumentos deben reflejar una concepción inclusiva de la educación, conforme a los lineamientos de la política educativa nacional (Ministerio de Educación, 2015).

En las evaluaciones nacionales de logros de aprendizaje, se emplean dos formatos de ítems: de opción múltiple y de respuesta construida. En seguida, se describe cada uno.

Ítem de opción múltiple. El formato de opción múltiple (el más usual) presenta en su estructura un estímulo, un enunciado o pregunta y tres o cuatro alternativas de respuesta, entre las cuales el estudiante debe identificar la única que es correcta.

Esto implica que dos o tres alternativas funcionan como distractores (respuestas incorrectas). En algunos ítems, alguno de los distractores puede constituir una respuesta intermedia o de logro parcial y, por lo tanto, esta se modela como un ítem de crédito parcial (puntaje parcial). En las evaluaciones implementadas por la UMC, las pruebas de 2.º grado de primaria en su mayoría están conformadas por ítems que presentan tres alternativas para evitar que el estudiante, debido a su edad o al nivel de desarrollo de su competencia lectora, haga un mayor esfuerzo cognitivo en la lectura del ítem que en su solución. Para los demás grados, las pruebas están conformadas por ítems que presentan cuatro alternativas.

La construcción del enunciado y de las alternativas de un ítem de opción múltiple 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 en que se encuentra.
- El ítem debe estar libre de todo sesgo sociocultural, socioeconómico, de género y 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.
- El ítem debe ser fácil de entender; es decir, la persona evaluada debe comprender con claridad qué tarea debe realizar, al margen de que pueda realizarla.
- El ítem debe tener elementos visuales y gráficos solo cuando sean necesarios para responderlo.
- El ítem debe tener una alta legibilidad material; es decir, el tamaño y el tipo de fuente, los espacios entre caracteres y entre líneas, el contraste, entre otras características, deben ser los adecuados.

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 alternativa correcta debe ser solo una y debe estar acompañada por distractores plausibles.
- A lo largo de la prueba, la alternativa correcta debe estar distribuida entre las distintas ubicaciones.
- Las alternativas deben ser preferiblemente tres o cuatro.
- Las alternativas deben disponerse, de preferencia, de forma vertical.
- Las alternativas de cada ítem deben organizarse de acuerdo con un criterio; por ejemplo, del número menor al número mayor.
- Las alternativas 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 alternativa 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 demandan al estudiante proporcionar, desarrollar o crear una respuesta particular. El estímulo consiste en una situación contextualizada o un texto (en el caso de la prueba de Lectura) que le da un marco tanto al ítem como a la respuesta que se espera. Este tipo de ítem también puede incluir instrucciones o especificaciones para responder.

La inclusión de un ítem de respuesta construida responde a la necesidad de explorar tanto los procedimientos como las habilidades complejas de argumentación, evaluación, juicio crítico y toma de decisiones razonada que están involucrados en la resolución de ciertas tareas. Si bien se pueden formular ítems cerrados de opción múltiple para evaluar estas habilidades complejas, los ítems de respuesta construida permiten medirlas en toda su dimensión y con mayor profundidad.

Las respuestas de los estudiantes a un ítem de respuesta construida son codificadas por expertos en el área especialmente capacitados para este fin. Los codificadores pasan por un riguroso proceso de capacitación que les permite apropiarse de la lógica de los ítems, los criterios de codificación y los posibles tipos de respuestas. Tanto para el proceso de capacitación como para la codificación en sí, se utiliza un manual de codificación que especifica el indicador y el descriptor que el ítem evalúa, y establece los tipos de respuesta (de crédito total, de crédito parcial y sin crédito), así como los criterios de codificación para cada uno ellos. Además, este manual brinda, al menos, cuatro ejemplos de respuestas reales de los estudiantes para cada uno de los criterios establecidos.

Las respuestas de crédito total son aquellas que muestran el logro completo de los procedimientos y de las habilidades implicadas en la resolución de la tarea. En este proceso, un estudiante puede utilizar diferentes procedimientos. Por su parte, las respuestas de crédito parcial son aquellas que reflejan un logro parcial del estudiante en la resolución de la tarea; mientras que las respuestas sin crédito evidencian una comprensión errada y, por tanto, no responden a los criterios de codificación establecidos.

En las pruebas de las evaluaciones nacionales de logros de aprendizaje, los ítems de respuesta construida se encuentran en menor cantidad que los ítems de opción múltiple. Esto se debe a que el procesamiento de las respuestas de este tipo de ítems es complejo y, en aplicaciones masivas como la EM 2022, demanda muchos recursos y tiempo.

Ítem tipo Likert. Este formato de ítem usualmente se compone de un enunciado y diferentes alternativas de respuesta que conforman una escala ordinal (p. ej. nunca, casi nunca, a veces, casi siempre, siempre). Este tipo de ítem surge con la publicación de Rensis Likert en 1932, en la que se presenta un tipo de instrumento para la medición de las actitudes (Likert, 1932). Estos ítems son preguntas individuales que pueden ser colocadas en un instrumento sin intención de combinarlas en un puntaje compuesto (composite score) y que se analizan a nivel ordinal. También permiten diseñar escalas, en las que se combinan las respuestas de los distintos ítems para representar la variable o constructo a medir. El tipo de escala usado en un ítem depende del constructo que se está tratando de medir, siendo las anclas las “palabras” que se utilizan para identificar cada una de las categorías de respuesta. Es común el uso de escalas de acuerdo, en que las palabras ancla suelen ser Totalmente de acuerdo, De acuerdo, Ni de acuerdo ni en desacuerdo, En desacuerdo, Totalmente en desacuerdo. En la EM 2022, se propone a los estudiantes una serie de afirmaciones o ítems que buscan medir las habilidades socioemocionales seleccionadas, por ende, estos ítems hacen referencia a aspectos característicos de las mismas.

1.2 Procesos de validación aplicados

Estos procesos se realizan para obtener evidencias de validez vinculadas al contenido de las pruebas. Este tipo de evidencias se vincula con la necesidad de garantizar que la prueba sea una muestra adecuada y representativa de los contenidos que pretende evaluar (McGartland et al. 2003; Muñiz, 1998). En esta misma línea, Suen (1990) indica que los ítems en un test o prueba, 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. Esta se definiría, entonces, como la representatividad de los ítems empleados en un test para medir un constructo particular. A su vez, la representatividad de un test puede definirse como la precisión con que se puede 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 (o dimensiones) 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 construir y seguir matrices (o tablas) de especificaciones.

La elaboración de estas matrices de especificaciones tiene como base el CNEB. Con ello, se garantiza el alineamiento de las pruebas con el currículo vigente, lo cual es una condición sine qua non para que estas brinden evidencias suficientes del logro de la competencia a evaluar en diferentes niveles de desarrollo por parte de los estudiantes. Asimismo, las matrices de un año actual guardan una estrecha relación con las del año anterior en cuanto a su contenido pedagógico, así como respecto de las proporciones o pesos de cada dimensión de la competencia evaluada en la prueba para, en conjunción con los procesos de equiparación, asegurar su comparabilidad en el tiempo. En tal sentido, los porcentajes de cada dimensión (competencias, capacidades, conocimientos y contextos) son muy similares año tras año. Por ejemplo, las matrices de especificaciones de la Evaluación Muestral 2022 son muy similares a las de la Evaluación Censal de Estudiantes y de la Evaluación Muestral 2019. Esta correspondencia de las matrices es validada por los especialistas de las direcciones de línea del Ministerio de Educación.

En los siguientes acápitales, se describen los procesos de validación de los instrumentos de evaluación elaborados sobre la base de las matrices de especificaciones.

1.2.1 Validación de expertos de la UMC

Después de que los equipos de especialistas de las áreas (Matemática, Comunicación, Ciudadanía y Ciencias Sociales, Ciencia y Tecnología y EIB) han

construido los ítems, estos son revisados por expertos de la UMC. Para esto, se cuenta con guías que orientan el proceso y con fichas que permiten evaluar cada ítem en aspectos como calidad, vigencia y veracidad de la información según cada disciplina científica, la correspondencia con las matrices de especificaciones, la adecuación de la complejidad del ítem a la población evaluada, y la construcción del enunciado y de 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 (2.º, 4.º y 6.º grado de primaria, 2.º grado de secundaria, y 4.º grado de primaria EIB), correspondientes al área de Comunicación, se revisan también diversos aspectos referidos a los textos a partir de los cuales se construyen los ítems, tales como el tipo y el género textuales, la estructura textual, la densidad de la información, el vocabulario, la sintaxis, la familiaridad del tema, la extensión y el formato.

1.2.2 Revisión de ítems y consulta con la población objetivo

En el caso de HSE, tanto en 6.º grado de primaria como en 2.º grado de secundaria se realizó, en diferentes momentos, un proceso de revisión de los ítems que conformaban las escalas seleccionadas para su medición. Las revisiones consistieron en reuniones virtuales con estudiantes, las cuales permitieron socializar los ítems y asegurar su comprensión y reformulación. En 6.º grado de primaria, durante el 2022, participaron estudiantes de diferentes zonas urbanas y rurales del país. En 2.º grado de secundaria, durante el 2020, participaron estudiantes de Lima. Como resultado de estas consultas con estudiantes, se realizaron ajustes en los términos que no eran bien comprendidos. Finalmente, se debe tener en cuenta que en la EM 2022 se aplicaron las mismas escalas de HSE en ambos grados.

1.2.3 Validación de campo: aplicación piloto

Después de la revisión de expertos al interior de la UMC, los subequipos de evaluación responsables de las pruebas subsanan las observaciones e incorporan las sugerencias hechas por los revisores. Luego de esto, los ítems pasan a ser empleados en la construcción de las pruebas destinadas a la aplicación piloto, la cual consiste en la aplicación de pruebas a una muestra representativa de estudiantes a nivel nacional.

Un diseño estadístico permite determinar las instituciones educativas (IE) que formarán 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 mediante cada una de las pruebas. Asimismo, se recogen evidencias acerca del funcionamiento de los ítems y su ajuste al modelo de medición empleado en la UMC.

Cabe señalar que las aplicaciones piloto, además de alimentar el banco de ítems de la UMC, permiten renovar un porcentaje de los ítems de las pruebas cada año y reemplazar aquellos ítems liberados; es decir, que son publicados en los informes de resultados que produce y distribuye la UMC.

1.2.4 Análisis de las propiedades psicométricas y revisión de aspectos pedagógicos

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 de dificultad. Entre los aspectos psicométricos, se analizan el ajuste de los ítems, la correlación ítem-medida, la confiabilidad de las medidas, la unidimensionalidad, la correspondencia entre la media de habilidad de los estudiantes y la media de dificultad de los ítems, y el funcionamiento diferencial de los ítems según los distintos estratos (urbano y rural, estatal y no estatal, hombre y mujer) y según su posición en la prueba. Igualmente, se ensayan hipótesis sobre las posibles causas por las que un ítem no ajustó al modelo psicométrico, tuvo un funcionamiento diferencial o su dificultad no concordó con la complejidad propuesta. En el capítulo 4 de este documento, se detallan dichas propiedades psicométricas con la finalidad de ampliar más este aspecto.

En la revisión de aspectos pedagógicos, se analiza la correspondencia entre la complejidad pedagógica estimada de un ítem y la dificultad dada por el modelo psicométrico, el funcionamiento de los distractores, 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 del área de Comunicación (Lectura), 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 funcionamiento y ajuste al modelo psicométrico para la conformación de las pruebas de la aplicación definitiva.

1.2.5 Validación de expertos de otras oficinas del Minedu

De acuerdo con las recomendaciones técnicas, el proceso de validación de los ítems incluye que un conjunto de expertos en cada una de las áreas evaluadas emita su juicio sobre tres aspectos: la construcción, la correspondencia con el indicador señalado y la dificultad de la tarea solicitada de 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 para que sean validados por expertos de las distintas direcciones del Minedu involucradas, tales como la Dirección de Educación Primaria (DEP) y la Dirección de Educación Secundaria (DES). En este proceso de validación, también participan especialistas de la Dirección de Evaluación Docente (DIED), por su experiencia en el diseño y la construcción de ítems, así como por su nivel de aporte relacionado con esa experiencia.

En esta validación, se solicita a los jueces que evalúen los aspectos mencionados en párrafo anterior, usando 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 con las características de la competencia evaluada. 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 particular, estime convenientes. Luego, los subequipos de cada área se encargan de sistematizar las fichas, absolver los cuestionamientos e incorporar las sugerencias. Ello implica, en algunos casos, realizar algún tipo de ajuste al ítem y/o al texto.

1.2.6 Adaptación de instrumentos para estudiantes con discapacidad

Desde 2015, la UMC adapta (es decir, realiza cambios pertinentes) 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 y autónoma de los estudiantes con discapacidad sensorial (baja visión, ceguera, hipoacusia y sordera), discapacidad física y trastorno del espectro autista de las evaluaciones nacionales de logros de aprendizaje.

Dependiendo de su tipo y su profundidad, Browning (2003) clasifica las adaptaciones en dos categorías: acomodaciones y modificaciones. Las acomodaciones son cambios introducidos a los instrumentos de evaluación o a los procedimientos de aplicación que posibilitan a los estudiantes con discapacidad acceder autónomamente a las evaluaciones, sin alterar el constructo medido. En contraste,

las modificaciones son cambios que alteran el estándar o las expectativas de logro y, por tanto, el constructo medido, ya que no se evalúan las mismas habilidades que forman parte de la competencia como a los demás estudiantes.

En la EM 2022, se implementaron las siguientes acomodaciones en las pruebas y procesos de aplicación:

Tabla 1.2 Acomodaciones a las pruebas de la EM 2022

Tipo de discapacidad	Tipo de acomodación	Descripción
Baja visión	Presentación	Prueba en macrotipo
	Equipos y materiales	Plumón delgado, lupa, lápiz jumbo (extragrueso)
	Extensión de la prueba	Menor cantidad de preguntas a resolver en el mismo tiempo establecido para todos los estudiantes
	Ubicación	Carpeta junto a acceso de luz natural
	Apoyo	Aplicador personalizado
Visual	Presentación	Prueba en sistema braille e imágenes en relieve Prueba asistida (impresa en tinta e imágenes en relieve, para estudiantes que no dominan el sistema braille) con el apoyo en la lectura de un aplicador
	Equipos y materiales	Regleta, punzón, ábaco, cubarritmo
	Extensión de la prueba	Menor cantidad de preguntas a resolver en el mismo tiempo establecido para todos los estudiantes
	Ubicación	Carpeta cerca a puerta de acceso al aula, para el caso de los estudiantes que rinden la prueba en braille Aula diferente para los estudiantes que recibieron la prueba asistida
	Apoyo	Aplicador personalizado

Tipo de discapacidad		Tipo de acomodación	Descripción
Auditiva	Hipoacusia/ Sordera	Presentación	Prueba con léxico y sintaxis simplificados
		Equipos/materiales	Lápiz, borrador
		Extensión de la prueba	Menor cantidad de preguntas a resolver en el mismo tiempo establecido para todos los estudiantes
		Ubicación	Carpeta al fondo del aula
		Apoyo	Intérprete de lengua de señas peruana (si el estudiante lo requería)
Sordoceguera	Según la severidad de esta discapacidad, se aplicaron las mismas acomodaciones que en los casos de baja visión o ceguera.		
Física		Presentación	Prueba sin adaptaciones en el formato o prueba en macrotipo (si el estudiante lo requería debido a la distancia con el objeto de lectura)
		Equipos/materiales	Lápiz jumbo, soporte de lápiz (si el estudiante lo requería)
		Extensión de la prueba	Igual cantidad de preguntas a resolver que todos los estudiantes y en el mismo tiempo establecido para todos los estudiantes
		Ubicación	Carpeta cerca a puerta de acceso al aula
		Apoyo	Aplicador personalizado para asistirlo en el marcado de las respuestas (si el estudiante lo requería)
Trastorno del espectro autista	Asperger	Presentación	Prueba sin adaptaciones en el formato
		Equipos/materiales	Tarjetas de comunicación alternativa (si el estudiante lo requería)

Tipo de discapacidad	Tipo de acomodación	Descripción	
Trastorno del espectro autista	Asperger	Extensión de la prueba	Igual cantidad de preguntas a resolver que todos los estudiantes y en el mismo tiempo establecido para todos los estudiantes
		Ubicación	Lugar y carpeta que usa habitualmente
		Apoyo	Aplicador personalizado
Trastorno del espectro autista	Autismo	Presentación	Prueba con léxico y sintaxis simplificados
		Equipos/materiales	Tarjetas de comunicación alternativa
		Extensión de la prueba	Menor cantidad de preguntas a resolver en el mismo tiempo establecido para todos los estudiantes
		Ubicación	Lugar y carpeta que usa habitualmente
		Apoyo	Aplicador personalizado

Tabla 1.3 Modificaciones a las pruebas de la EM 2022

Discapacidad	Tipo de modificación	Modificación
Intelectual	Presentación	Prueba diferenciada para estudiantes lectores y lectores iniciales
		Prueba de complejidad escalonada
	Equipos/materiales	Lápiz, borrador
	Extensión de la prueba	Menor cantidad de preguntas a resolver en el mismo tiempo establecido para todos los estudiantes
	Ubicación	Carpeta al fondo del aula
Apoyo	Aplicador personalizado para asistirlo en la lectura y/o marcado de las respuestas (si el estudiante lo requiriera)	

1.2.7 Aplicación definitiva

La aplicación definitiva de la EM 2022 se realizó en el último trimestre del año (octubre, noviembre y diciembre). Esta se llevó a cabo durante dos días en 2.º, 4.º y 6.º grado de primaria, y tres días en 2.º grado de secundaria. En el primer día,

tanto en 2.º y 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, este orden de aplicación de las pruebas fue inverso. Este cambio obedece a la necesidad de neutralizar los posibles efectos del cansancio de los estudiantes, pues rinden dos pruebas en un día. En el caso particular de 6.º grado de primaria, en el primer día solo se aplicó la prueba de Lectura y en el segundo día, solo la de Matemática.

Asimismo, en el caso de 2.º grado de secundaria, el tercer día se destinó a la aplicación de la prueba de Ciencia y Tecnología.

Cabe resaltar que, a partir de 2015, los instrumentos utilizados en la aplicación definitiva de las evaluaciones nacionales de logros de aprendizaje, incluidos los cuadernillos de las pruebas, retornan en su totalidad al Minedu y se mantienen bajo estricta confidencialidad. Este procedimiento, aunque complicado, se realiza porque permite alimentar un banco de ítems creciente y asegurar el proceso de equiparación de las pruebas por ítems en común y, con ello, la comparabilidad de los resultados en el tiempo.

En resumen, mediante los procesos de construcción y validación, se logró que los instrumentos de evaluación de la EM 2022 cumplan con los requerimientos técnicos y consideraciones de contenido que permitan generar información válida y confiable sobre los niveles de logro de los estudiantes evaluados. Asimismo, se garantizó que los instrumentos guardaran correspondencia con los diversos documentos curriculares del sistema educativo nacional y que los resultados derivados de su aplicación sean comparables a lo largo del tiempo.

1.3 Ensamblaje de las pruebas mediante el diseño bloques

Conforme avanzan en su escolaridad, los estudiantes deben lograr aprendizajes cada vez más complejos. Así, por ejemplo, en Lectura, el lenguaje de los textos se complejiza y se especializa progresivamente, los temas tratados en los textos son más diversos, y las capacidades de inferencia y reflexión se diversifican y manifiestan en desempeños cada vez más complejos. 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 forma precisa lo que realmente saben y pueden hacer los estudiantes.

Para lograr este objetivo, se ha utilizado el diseño de bloques no balanceados incompletos que permitió generar diversas formas de la misma prueba. Las evidencias generadas mediante las aplicaciones piloto, así como por las simulaciones realizadas (las cuales consideraron distintos escenarios) han permitido tomar esta decisión con la certeza de contar con buenos o aceptables indicadores psicométricos de

confiabilidad de los resultados, así como de precisión y consistencia en la clasificación de los estudiantes en múltiples iteraciones. De la misma manera, estos diseños de pruebas han posibilitado que las formas se adaptaran al nivel de habilidad de grupos de estudiantes diferentes. En este sentido, se han elaborado formas con un mayor nivel de dificultad que han sido aplicadas por estudiantes de escuelas urbanas y formas con menor nivel de dificultad que han sido aplicadas por estudiantes de escuelas rurales y por todos los estudiantes de las regiones Loreto y Ucayali.

En concordancia con lo anterior, la organización y ensamblaje de estas formas contempló la presencia de bloques de ítems comunes (generalmente, en una posición fija), que formaran parte de las formas de mayor y menor dificultad, y de bloques no comunes (generalmente en diferentes posiciones), que formaran parte solo de las formas de mayor dificultad o solo de aquellas de menor dificultad. La combinación de bloques comunes y no comunes permitió el ensamblaje de formas que se distribuyeron aleatoriamente entre los estudiantes de cada grupo antes mencionado. De este modo y gracias al uso de técnicas de equiparación del modelo Rasch, las distintas pruebas estuvieron conformadas por una mayor cantidad de ítems en la misma métrica sin que ello implicara incrementar la cantidad de ítems a los que cada estudiante se enfrentó. Además, estas formas tenían un nivel de dificultad ajustado al nivel de habilidad del grupo destinatario.

En relación con la conformación de los bloques, es preciso señalar que el bloque común (distribuido en pequeños bloques comunes) tenía también la finalidad de optimizar la equiparación de las medidas y fue construido de tal forma que los ítems reflejaran la tabla de especificaciones, así como la dispersión de las medidas a lo largo de la escala de dificultad.

Por último, en el caso de las pruebas con adaptaciones destinadas a estudiantes con discapacidad, se ha considerado el diseño de prueba única. En las siguientes tablas, se presenta el diseño de bloques de las pruebas aplicadas. Las tablas de especificaciones de estos instrumentos pueden ser consultadas en los anexos.

Diseño de bloques de las pruebas de Lectura de 2.º grado de primaria

La prueba de Lectura de 2.º grado de primaria elaborada para la EM 2022 estuvo conformada por 90 ítems, de los cuales 82 se construyeron a partir de un conjunto de 18 textos de diferente tipo, género, formato y complejidad; mientras que 8 ítems fueron de lectura inicial, por lo que no están vinculados a ningún texto. Los 18 textos y 90 ítems se distribuyeron en 26 bloques. De ellos, 6 bloques son comunes a todas las formas (3 por día) y 20 bloques, no comunes (10 por día), es decir, solo se encuentran en un tipo de formas, pero no en el otro tipo. Las diferentes combinaciones de los bloques permitieron ensamblar 8 formas, 4 de mayor dificultad y otras 4 de

menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (capacidades, tipos de textos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

Tabla 1.3 Matriz de bloques de la prueba de Lectura de 2.º grado de primaria. Día 1

Forma	Bloque	Ítems por forma
1	B03	25
	B05	
	B06	
	B07	
	B08	
	B10	
2	B04	25
	B05	
	B06	
	B07	
	B09	
	B11	
3	B01	25
	B03	
	B05	
	B06	
	B07	
	B08	
4	B12	25
	B02	
	B04	
	B05	
	B06	
	B07	
	B09	
	B13	

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

Forma	Bloque	Ítems por forma
5	B16	25
	B18	
	B19	
	B20	
	B21	
	B23	
6	B17	25
	B18	
	B19	
	B20	
	B22	
7	B24	25
	B14	
	B16	
	B18	
	B19	
	B20	
8	B21	25
	B25	
	B15	
	B17	
	B18	
	B19	
	B20	
B22		
	B26	

De las 8 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 esta prueba, cada estudiante respondió 50.

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

La prueba de Lectura de 4.º grado de primaria elaborada para la EM 2022 estuvo conformada por 17 textos de diferente tipo, género, formato y complejidad, a partir de los cuales se elaboraron 96 ítems. Dichos textos e ítems se distribuyeron en un total de

16 bloques. De ellos, 4 bloques son comunes (dos por día) y 12 bloques, no comunes (6 por día). Las diferentes combinaciones de los bloques permitieron ensamblar 12 formas, 6 de mayor dificultad y 6 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (capacidades, tipos de textos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

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

Forma	Bloque	Ítems por forma
1	B05	24
	B01	
	B02	
	B06	
2	B06	24
	B01	
	B02	
3	B07	24
	B10	
	B01	
4	B02	24
	B05	
	B09	
5	B09	24
	B01	
	B02	
6	B10	24
	B01	
	B02	
	B08	

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

Forma	Bloque	Ítems por forma
7	B11	24
	B03	
	B04	
	B12	
8	B12	24
	B03	
	B04	
	B16	
9	B13	24
	B03	
	B04	
	B14	
10	B11	24
	B03	
	B04	
	B15	
11	B15	24
	B03	
	B04	
	B16	
12	B16	24
	B03	
	B04	
	B14	

De las 12 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 96 ítems que tiene esta prueba, cada estudiante respondió 48.

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

La prueba de Lectura de 6.º grado de primaria elaborada para la EM 2022 estuvo conformada por 18 textos de diferente tipo, género, formato y complejidad, a partir de los cuales se elaboraron 108 ítems. Dichos textos e ítems se distribuyeron en 18 bloques. De ellos, 11 bloques son comunes y 7, no comunes. Las diferentes combinaciones de los bloques permitieron ensamblar 10 formas, 6 de mayor

dificultad y 4 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (capacidades, tipos de textos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

Tabla 1.7 Matriz de bloques de la prueba de Lectura de 6.º grado de primaria. Día único

Forma	Bloque	Ítems por forma
1	B09	24
	B01	
	B05	
	B11	
2	B11	24
	B02	
	B06	
	B12	
3	B12	24
	B03	
	B07	
	B17	
4	B13	24
	B04	
	B08	
	B14	
5	B14	24
	B01	
	B05	
	B15	
6	B15	24
	B02	
	B06	
	B10	
7	B09	24
	B01	
	B05	
	B16	

Forma	Bloque	Ítems por forma
8	B16	24
	B02	
	B06	
	B17	
9	B17	24
	B03	
	B07	
10	B18	24
	B18	
	B04	
	B08	
	B10	

De las 10 formas ensambladas, cada estudiante evaluado se enfrentó a una de ellas. En consecuencia, de los 108 ítems que tiene esta prueba, cada estudiante respondió 24.

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

La prueba de Lectura de 2.º grado de secundaria elaborada para la EM 2022 estuvo conformada por 20 textos de diferente tipo, género, formato y complejidad, a partir de los cuales se elaboraron 122 ítems. Dichos textos e ítems se distribuyeron en 20 bloques. De ellos, 4 bloques son comunes (dos por día) y 16 bloques no comunes (8 por día). Las diferentes combinaciones de los bloques permitieron ensamblar 18 formas, 10 de mayor dificultad y 8 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (capacidades, tipos de textos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

Tabla 1.8 Matriz de bloques de la prueba de Lectura de 2.º grado de secundaria. Día 1

Forma	Bloque	Ítems por forma
1	B05	25
	B01	
	B02	
	B06	

Forma	Bloque	Ítems por forma
2	B06	25
	B01	
	B02	
	B07	
3	B07	25
	B01	
	B02	
4	B08	25
	B08	
	B01	
	B02	
5	B09	25
	B11	
	B01	
	B02	
6	B05	25
	B05	
	B01	
	B02	
7	B10	25
	B10	
	B01	
	B02	
8	B06	25
	B11	
	B01	
	B02	
9	B12	25
	B12	
	B01	
	B02	
	B09	

**Tabla 1.9 Matriz de bloques de la prueba de Lectura de 2.º grado de secundaria.
Día 2**

Forma	Bloque	Ítems por forma
10	B13	25
	B03	
	B04	
	B14	
11	B14	25
	B03	
	B04	
	B15	
12	B15	25
	B03	
	B04	
	B16	
13	B16	25
	B03	
	B04	
	B17	
14	B12	25
	B03	
	B04	
	B17	
15	B13	25
	B03	
	B04	
	B18	
16	B18	25
	B03	
	B04	
	B19	
17	B19	25
	B03	
	B04	
	B20	
18	B20	25
	B03	
	B04	
	B17	

De las 18 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 122 ítems que tiene esta prueba, cada estudiante respondió 50.

Diseño de bloques de las pruebas de Matemática de 2.º grado de primaria

La prueba de Matemática de 2.º grado de primaria elaborada para la EM 2022 estuvo conformada por un total de 92 ítems distribuidos en 8 formas. Cada forma contiene 23 ítems, 11 ítems comunes organizados en 2 bloques comunes en posición fija y 12 ítems no comunes organizados en 2 bloques rotados que varían de una forma a otra. Las diferentes combinaciones de los bloques permitieron ensamblar 4 formas de mayor dificultad y 4 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (competencias, capacidades, conocimientos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

Tabla 1.10 Matriz de bloques de la prueba de Matemática de 2.º grado de primaria. Día 1

Forma	Bloque	Ítems por forma
1	B01	23
	B09	
	B14	
	B02	
2	B02	23
	B10	
	B13	
3	B01	23
	B05	
	B09	
	B13	
4	B06	23
	B10	
	B14	
	B05	

Tabla 1.11 Matriz de bloques de la prueba de Matemática de 2.º grado de primaria. Día 2

Forma	Bloque	Ítems por forma
5	B03	23
	B11	
	B16	
	B04	
6	B04	23
	B12	
	B15	
7	B03	23
	B07	
	B11	
	B15	
8	B08	23
	B08	
	B12	
	B16	
	B07	

De las 8 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 92 ítems que tiene esta prueba, cada estudiante respondió 46.

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

La prueba de Matemática de 4.º grado de primaria elaborada para la EM 2022 estuvo conformada por un total de 102 ítems distribuidos en 12 formas. Cada forma contiene 25 ítems, 11 ítems comunes organizados en 4 bloques comunes y 14 ítems no comunes que corresponden a 2 bloques rotados que varían de una forma a otra. Las diferentes combinaciones de los bloques permitieron ensamblar 6 formas de mayor dificultad y 6 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (competencias, capacidades, conocimientos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

**Tabla 1.12 Matriz de bloques de la prueba de Matemática de 4.º grado de primaria.
Día 1**

Forma	Bloque	Ítems por forma
1	B01	25
	B09	
	B13	
	B19	
	B10	
	B02	
2	B02	25
	B10	
	B14	
	B19	
	B09	
	B01	
3	B06	25
	B10	
	B15	
	B19	
	B09	
	B02	
4	B05	25
	B09	
	B13	
	B19	
	B10	
	B06	
5	B06	25
	B10	
	B14	
	B19	
	B09	
	B05	
6	B05	25
	B10	
	B15	
	B19	
	B09	
	B06	

**Tabla 1.13 Matriz de bloques de la prueba de Matemática de 4.º grado de primaria.
Día 2**

Forma	Bloque	Ítems por forma
7	B03	25
	B11	
	B16	
	B20	
	B12	
8	B04	25
	B04	
	B12	
	B17	
	B20	
9	B11	25
	B03	
	B08	
	B12	
	B18	
10	B20	25
	B11	
	B04	
	B07	
	B11	
11	B16	25
	B20	
	B12	
	B08	
	B08	
12	B12	25
	B17	
	B20	
	B11	
	B07	
12	B07	25
	B12	
	B18	
	B20	
	B11	
	B08	

De las 12 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 102 ítems que tiene esta prueba, cada estudiante respondió 50.

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

La prueba de Matemática de 6.º grado de primaria elaborada para la EM 2022 estuvo conformada por un total de 94 ítems distribuidos en 10 formas. Cada forma contiene 25 ítems, 12 ítems comunes organizados en 3 bloques comunes y 13 ítems no comunes que corresponden a bloques que varían de una forma a otra. Las diferentes combinaciones de los bloques permitieron ensamblar 6 formas de mayor dificultad y 4 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (competencias, capacidades, conocimientos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

Tabla 1.14 Matriz de bloques de la prueba de Matemática de 6.º grado de primaria. Día 2

Forma	Bloque	Ítems por forma
1	B01	25
	B09	
	B17	
	B21	
	B10	
	B02	
2	B03	25
	B11	
	B18	
	B21	
	B12	
	B04	
3	B05	25
	B13	
	B19	
	B21	
	B14	
	B06	

Forma	Bloque	Ítems por forma
4	B02	25
	B10	
	B20	
	B21	
	B11	
	B03	
5	B04	25
	B14	
	B17	
	B21	
	B09	
	B05	
6	B06	25
	B12	
	B18	
	B21	
	B13	
	B01	
7	B02	25
	B15	
	B17	
	B22	
	B12	
	B07	
8	B04	25
	B16	
	B18	
	B22	
	B14	
	B08	
9	B08	25
	B14	
	B19	
	B22	
	B15	
	B02	

Forma	Bloque	Ítems por forma
10	B07	25
	B12	
	B20	
	B22	
	B16	
	B04	

De las 10 formas ensambladas, cada estudiante evaluado se enfrentó a una de ellas. En consecuencia, de los 94 ítems que tiene esta prueba, cada estudiante respondió 25.

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

La prueba de Matemática de 2.º grado de secundaria elaborada para la EM 2022 estuvo conformada por un total de 122 ítems distribuidos en 18 formas. Cada forma contiene 25 ítems, 13 ítems comunes organizados en 4 bloques comunes y 12 ítems no comunes que corresponden a un bloque que varía de una forma a otra. Las diferentes combinaciones de los bloques permitieron ensamblar 10 formas de mayor dificultad y 8 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (competencias, capacidades, conocimientos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

Tabla 1.15 Matriz de bloques de la prueba de Matemática de 2.º grado de secundaria. Día 1

Forma	Bloque	Ítems por forma
1	B06	25
	B09	
	B13	
	B21	
	B10	
	B02	
2	B02	25
	B10	
	B14	
	B21	
	B09	
	B06	

Forma	Bloque	Ítems por forma
3	B06	25
	B10	
	B15	
	B21	
	B09	
	B02	
4	B02	25
	B09	
	B16	
	B21	
	B10	
	B01	
5	B01	25
	B10	
	B16	
	B21	
	B09	
	B06	
6	B05	25
	B09	
	B13	
	B21	
	B10	
	B06	
7	B06	25
	B10	
	B14	
	B21	
	B09	
	B05	
8	B05	25
	B10	
	B15	
	B21	
	B09	
	B06	

Forma	Bloque	Ítems por forma
9	B06	25
	B09	
	B16	
	B21	
	B10	
	B05	

Tabla 1.16 Matriz de bloques de la prueba de Matemática de 2.º grado de secundaria. Día 2

Forma	Bloque	Ítems por forma
10	B08	25
	B11	
	B17	
	B22	
	B12	
	B04	
11	B04	25
	B12	
	B18	
	B22	
	B11	
	B08	
12	B08	25
	B12	
	B19	
	B22	
	B11	
	B04	
13	B04	25
	B11	
	B20	
	B22	
	B12	
	B03	

Forma	Bloque	Ítems por forma
14	B03	25
	B12	
	B20	
	B22	
	B11	
	B08	
15	B07	25
	B11	
	B17	
	B22	
	B12	
	B08	
16	B08	25
	B12	
	B18	
	B22	
	B11	
	B07	
17	B07	25
	B12	
	B19	
	B22	
	B11	
	B08	
18	B08	25
	B11	
	B20	
	B22	
	B12	
	B07	

De las 18 formas ensambladas, cada estudiante evaluado se enfrentó a dos de ellas (una por cada día de aplicación). En consecuencia, de los 122 ítems que tiene esta prueba, cada estudiante respondió 50.

Diseño de bloques de las pruebas de Ciencia y Tecnología de 2.º grado de secundaria

La prueba de Ciencia y Tecnología de 2.º grado de secundaria elaborada para la EM 2022 estuvo conformada por un total de 118 ítems distribuidos en 8 bloques comunes (4 bloques de 7 ítems cada uno y 4 bloques de 8 ítems cada uno) y 11 bloques no comunes (10 bloques de 5 ítems cada uno y 1 bloque de 8 ítems). Las diferentes combinaciones de los bloques permitieron ensamblar 6 formas de mayor dificultad y 4 de menor dificultad. De acuerdo con esta distribución, cada forma y el conjunto de estas reflejan los contenidos (competencias, capacidades, conocimientos, indicadores, entre otros) de las tablas de especificaciones de la prueba. Dentro de su tipo (de mayor o menor dificultad), estas formas presentan una dificultad similar entre sí, pero una dificultad diferenciada, según el grupo de estudiantes al que están dirigidas.

Tabla 1.17 Matriz de bloques de la prueba de Ciencia y Tecnología de 2.º grado de secundaria. Día 3

Forma	Bloque	Ítems por forma
1	B10	32
	B06	
	B01	
	B07	
	B11	
2	B11	32
	B07	
	B02	
	B08	
3	B12	32
	B08	
	B03	
	B09	
4	B13	32
	B13	
	B09	
	B04	
	B06	
	B14	

Forma	Bloque	Ítems por forma
5	B14	32
	B06	
	B05	
	B07	
	B15	
6	B15	32
	B07	
	B01	
	B08	
7	B10	32
	B16	
	B08	
	B02	
	B09	
8	B17	32
	B17	
	B09	
	B03	
	B06	
9	B18	32
	B06	
	B04	
	B07	
10	B19	32
	B19	
	B07	
	B05	
	B08	
	B16	

De las 10 formas ensambladas, cada estudiante evaluado se enfrentó a una de ellas. En consecuencia, de los 118 ítems que tiene esta prueba, cada estudiante respondió 32.

Diseño de bloques de los cuestionarios de Habilidades Socioemocionales de 6.º grado de primaria y de 2.º grado de secundaria

El cuestionario de Habilidades Socioemocionales de 6.º grado de primaria y 2.º grado de secundaria elaborado para la EM 2022 cuenta con un total de 94 ítems distribuidos en 7 bloques. Estos se distribuyen de la siguiente manera: B01 Responsabilidad, 7 ítems; B02 Toma de decisiones, 17 ítems; B03 Autoeficacia, 18 ítems; B04 Deseabilidad social, 9 ítems; B05 Empatía, 16 ítems; B06 Autorregulación conductual, 15 ítems; y, B07 Autonomía y relación, 12 ítems.

Tabla 1.18 Matriz de bloques del cuestionario de Habilidades Socioemocionales de 6.º grado de primaria (día 1) y 2.º grado de secundaria (día 3).

Forma	Bloque	Ítems por forma
1	B01 - Responsabilidad	94
	B02 - Toma de decisiones	
	B03 - Autoeficacia	
	B04 - Deseabilidad social	
	B05 - Empatía	
	B06 - Autorregulación conductual	
	B07 - Autonomía y relación	

Cada estudiante evaluado se enfrentó a los 7 bloques de preguntas. En consecuencia, cada estudiante respondió 94 ítems que tiene el cuestionario.

Población y muestra

Capítulo 2

Capítulo 2

Población y muestra

2.1 Población objetivo

La EM 2022 tuvo como población objetivo a los estudiantes que asistieron a clases en 2.º grado de primaria, 4.º grado de primaria, 6.º grado de primaria y 2.º grado de secundaria en instituciones educativas (IE) de Educación Básica Regular (EBR).

En la evaluación, se consideraron los grados mencionados porque en ellos se concluyen los ciclos III, IV, V y VI de la EBR respectivamente. Evaluar estos grados brinda la oportunidad de planificar intervenciones en el sistema educativo para mejorar los grados y ciclos posteriores, pues, conforme se avanza en la escolaridad, los conocimientos se especializan y se hacen más complejos.

Se excluyó de la población de estudiantes a aquellos que asisten a una IE que registra menos de cinco estudiantes matriculados¹ en el grado evaluado. En secundaria, la exclusión fue del 0,7 % de los estudiantes a nivel nacional. Por su parte, en primaria, fue de 5,7 % en 2.º grado, 5,5% en 4.º grado y 5,3% en 6.º grado. Las razones de esta exclusión se encuentran detalladas en el marco de trabajo de la ECE (Ministerio de Educación, 2009).

En el 2022, la EM tenía programado evaluar 11 824 IE y aproximadamente a 422 969 estudiantes entre todos los grados programados.

2.2 Marco muestral

El marco muestral ha sido elaborado con información proveniente de diferentes fuentes: el Padrón de Instituciones Educativas de la Unidad de Estadística Educativa (UEE), la cantidad de estudiantes y secciones del Censo Educativo 2021 y la información del Sistema de Información de Apoyo a la Gestión de la Institución Educativa (Siagie) del 2021. Anualmente, la Unidad de Estadística Educativa (UEE) del Minedu elabora y difunde el Padrón de Instituciones Educativas con fines estadísticos, el cual está siendo actualizado constantemente. 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

¹ La 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.

la información de la cantidad de estudiantes y secciones que proviene del Censo Educativo es declarada por el director de la IE. En algunos casos, el director no reporta el dato, por ello la Unidad de Estadística Educativa le imputa un valor. El porcentaje de escuelas con datos de matrícula imputados en el Censo Escolar 2021 es de alrededor del 1,6%.

Para elaborar el listado de IE a evaluar en la EM 2022, se aplicó el siguiente procedimiento:

1. Se utilizó el Padrón de Instituciones Educativas para determinar el listado de IE activas². Luego, se seleccionaron las IE de primaria y secundaria de Educación Básica Regular (EBR). Tras ello, en primaria, se excluyeron las IE que pertenecieran al servicio con el modelo educativo en Educación Intercultural Bilingüe (EIB).
2. Se calculó la cantidad de secciones y estudiantes programados. Esta información corresponde al valor registrado en Siagie 2021. En los casos en que la IE no contaba con información en el Siagie, se tomó el valor del Censo Educativo 2021.
3. Se seleccionaron las IE que tuvieron cinco o más estudiantes asignados en el punto 2.

2.3 Diseño muestral

El diseño de muestreo empleado fue de tipo probabilístico, estratificado, por conglomerados y bietápico, donde la unidad primaria de muestreo corresponde a las IE y la unidad secundaria corresponde a las secciones completas dentro de cada IE. En la primera etapa, la selección de escuelas fue proporcional al tamaño, por lo que las escuelas con mayor cantidad de estudiantes en el grado a evaluar tuvieron una mayor probabilidad de ser seleccionadas. En la segunda etapa, se realizó una selección aleatoria simple de hasta tres secciones.

2.3.1 Conglomerados

Cada IE es considerada un conglomerado, pues agrupa a un conjunto de secciones 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 exige contar con un marco muestral de cada elemento por analizar, pues solo es necesario tener el marco muestral a nivel de conglomerados (IE y secciones) para realizar la selección de estos (Lohr, 2022). Luego, en cada conglomerado seleccionado, se puede obtener el listado de todas las secciones y seleccionar una muestra de ellas.

² El 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/>

2.3.2 Estratificación

Para la EM 2022, se realizó una estratificación principalmente porque, en un contexto de tanta diversidad como el peruano, la variable de interés (rendimiento) toma diferentes valores en diferentes subpoblaciones (área³ y gestión para todos los grados, o característica en los grados de primaria). Por tal motivo, es importante reportar resultados para cada una de ellas. La combinación de estas tres subpoblaciones permite generar estratos mutuamente excluyentes (cada elemento pertenece únicamente a un estrato) y colectivamente exhaustivos.

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).

La EM 2022, además, cuenta con 26⁴ dominios conformados por cada Dirección Regional de Educación (DRE) bajo la consideración de que las IE son administradas por estas instancias. En cada DRE, se conforman los estratos mencionados, salvo en los casos de Lima Metropolitana y Callao, donde solo se consideran los estratos 1 y 2. Por ello, el total de estratos considerados es 124.

2.4 Tamaño muestral y cobertura

El tamaño de muestra es determinado por cada dominio. Por eso, una vez estratificado, se empleó en la muestra la asignación de Neyman (Lohr, 2022) para determinar el tamaño en cada uno de los estratos, el cual es directamente proporcional a su variabilidad y a su tamaño.

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 población evaluada y la cobertura alcanzada.

El porcentaje de estudiantes evaluados respecto de los estudiantes programados a nivel nacional varía entre el 93 % y el 95 % dependiendo del grado. Esta cobertura se

³ El marco muestral de la EM 2022 se elaboró empleando el padrón de IE, el cual adoptaba una nueva definición de ámbito urbano y rural (en el 2021, el Minedu adoptó la definición urbano/rural del INEI). Esta adopción implicó que alrededor de 3063 IE en primaria y 2405 IE en secundaria cambiara de clasificación de urbana a rural.

⁴ La EM de 6.º grado de primaria, a diferencia de las de los otros grados, no tiene desagregación por DRE.

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 2022.

Tabla 2.1 Cobertura de la EM 2022 a nivel de IE y estudiantes, 2.º grado de primaria

2.º grado de primaria	Programados		Evaluados		Cobertura (%)	
	IE	Estudiantes	IE	Estudiantes	IE	Estudiantes
Nacional	4 877	151 553	4 874	142 602	99,9	94,1
Amazonas	210	3 464	210	3 315	100,0	95,7
Áncash	154	5 260	154	4 949	100,0	94,1
Apurímac	116	3 838	116	3 747	100,0	97,6
Arequipa	218	8 208	218	7 866	100,0	95,8
Ayacucho	165	5 029	165	4 878	100,0	97,0
Cajamarca	541	9 010	541	8 610	100,0	95,6
Callao	101	5 268	101	4 792	100,0	91,0
Cusco	190	7 502	190	7 341	100,0	97,9
Huancavelica	210	3 353	210	3 209	100,0	95,7
Huánuco	212	4 951	212	4 679	100,0	94,5
Ica	178	7 746	178	7 257	100,0	93,7
Junín	258	7 420	258	7 037	100,0	94,8
La Libertad	230	7 517	229	7 089	99,6	94,3
Lambayeque	182	7 475	182	6 899	100,0	92,3
Lima Metropolitana	201	10 822	201	9 805	100,0	90,6
Lima Provincias	146	5 167	146	4 861	100,0	94,1
Loreto	205	6 538	205	5 798	100,0	88,7
Madre de Dios	51	2 270	51	2 141	100,0	94,3
Moquegua	42	1 727	42	1 662	100,0	96,2
Pasco	155	3 150	155	2 958	100,0	93,9
Piura	316	10 113	316	9 515	100,0	94,1
Puno	192	6 201	192	6 039	100,0	97,4
San Martín	310	7 300	310	6 884	100,0	94,3
Tacna	80	3 596	80	3 475	100,0	96,6
Tumbes	59	2 390	59	2 206	100,0	92,3
Ucayali	155	6 238	153	5 590	98,7	89,6

Tabla 2.2 Cobertura de la EM 2022 a nivel de IE y estudiantes, 4.° grado de primaria

4.° grado de primaria	Programados		Evaluados		Cobertura (%)	
	IE	Estudiantes	IE	Estudiantes	IE	Estudiantes
Nacional	3 946	128 856	3 942	121 635	99,9	94,4
Amazonas	146	2 657	146	2 547	100,0	95,9
Áncash	178	6 090	178	5 734	100,0	94,2
Apurímac	91	3 261	91	3 216	100,0	98,6
Arequipa	142	5 718	142	5 554	100,0	97,1
Ayacucho	129	4 260	129	4 145	100,0	97,3
Cajamarca	407	7 097	406	6 743	99,8	95,0
Callao	90	5 263	90	4 847	100,0	92,1
Cusco	148	5 946	148	5 808	100,0	97,7
Huancavelica	167	2 896	167	2 793	100,0	96,4
Huánuco	220	5 121	220	4 839	100,0	94,5
Ica	118	5 179	118	4 938	100,0	95,3
Junín	230	6 387	230	6 116	100,0	95,8
La Libertad	237	8 204	235	7 530	99,2	91,8
Lambayeque	130	5 742	130	5 312	100,0	92,5
Lima Metropolitana	120	6 727	120	6 105	100,0	90,8
Lima Provincias	105	3 704	105	3 503	100,0	94,6
Loreto	137	4 541	136	3 997	99,3	88,0
Madre de Dios	61	2 516	61	2 374	100,0	94,4
Moquegua	41	1 483	41	1 445	100,0	97,4
Pasco	137	3 077	137	2 880	100,0	93,6
Piura	250	8 947	250	8 487	100,0	94,9
Puno	161	5 714	161	5 592	100,0	97,9
San Martín	238	6 665	238	6 338	100,0	95,1
Tacna	80	3 648	80	3 545	100,0	97,2
Tumbes	68	2 994	68	2 749	100,0	91,8
Ucayali	115	5 019	115	4 498	100,0	89,6

Tabla 2.3 Cobertura de la EM 2022 a nivel de IE y estudiantes, 6.º grado de primaria

6.º grado de primaria	Programados		Evaluados		Cobertura (%)	
	IE	Estudiantes	IE	Estudiantes	IE	Estudiantes
Nacional	268	10 179	267	9 694	99,6	95,2
Estatad	209	7 780	209	7 433	100,0	95,5
No estatal	59	2 399	58	2 261	98,3	94,2
Urbano	140	8 206	139	7 836	99,3	95,5
Rural	128	1 973	128	1 858	100,0	94,2
Polidocente completo	171	9 227	170	8 808	99,4	95,5
Unidocente / multigrado	97	952	97	886	100,0	93,1

Tabla 2.4 Cobertura de la EM 2022 a nivel de IE y estudiantes, 2.º de grado secundaria

2.º de grado secundaria	Programados		Evaluados		Cobertura (%)	
	IE	Estudiantes	IE	Estudiantes	IE	Estudiantes
Nacional	2 733	132 381	2 733	123 948	100,0	93,6
Amazonas	162	5 859	162	5 261	100,0	89,8
Áncash	117	5 673	117	5 380	100,0	94,8
Apurímac	98	3 974	98	3 866	100,0	97,3
Arequipa	116	6 085	116	5 962	100,0	98,0
Ayacucho	129	4 510	129	4 320	100,0	95,8
Cajamarca	173	6 816	173	6 452	100,0	94,7
Callao	63	4 297	63	4 072	100,0	94,8
Cusco	140	6 490	140	6 284	100,0	96,8
Huancavelica	115	2 906	115	2 757	100,0	94,9
Huánuco	112	4 763	112	4 317	100,0	90,6
Ica	61	4 248	61	4 078	100,0	96,0
Junín	164	7 580	164	7 079	100,0	93,4
La Libertad	106	5 867	106	5 421	100,0	92,4
Lambayeque	84	5 164	84	4 883	100,0	94,6
Lima Metropolitana	144	9 271	144	8 812	100,0	95,0
Lima Provincias	79	4 379	79	4 174	100,0	95,3
Loreto	148	7 282	148	6 244	100,0	85,7

2.° de grado secundaria	Programados		Evaluados		Cobertura (%)	
	IE	Estudiantes	IE	Estudiantes	IE	Estudiantes
Madre de Dios	48	2 763	48	2 588	100,0	93,7
Moquegua	32	1 631	32	1 591	100,0	97,5
Pasco	87	2 912	87	2 654	100,0	91,1
Piura	132	8 176	132	7 658	100,0	93,7
Puno	114	5 277	114	5 185	100,0	98,3
San Martín	94	4 728	94	4 338	100,0	91,8
Tacna	50	2 814	50	2 764	100,0	98,2
Tumbes	40	2 479	40	2 307	100,0	93,1
Ucayali	125	6 437	125	5 501	100,0	85,5

Respecto de los niveles de inferencia de la EM 2022 en 2.° y 4.° grado de primaria, 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), e) por característica (polidocente completo / unidocente-multigrado), f) a nivel de DRE y g) a nivel de gestión, área o característica dentro de la DRE. En 6.° grado de primaria, los niveles de inferencia no incluyen la DRE.

2.5 Pesos muestrales

Los pesos son ponderadores que ayudan a corregir la distribución de la muestra en los estratos y a expandir la información muestral a la población.

Peso por IE. El peso, en la primera etapa, 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, tal como se observa en la siguiente fórmula:

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

donde:

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

pIE_{ij} = 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, la selección de secciones fue aleatoria simple, tal como se observa en la siguiente fórmula:

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

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, tal como se observa en la siguiente fórmula:

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

donde:

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

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

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

Peso final. Se obtiene un peso separado para cada área evaluada. Ambos pesos se consiguen mediante 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 de la i-ésima IE.

2.6 Validación de la muestra seleccionada

Las características del diseño muestral de la EM 2022 fueron replicadas en los operativos censales de 2.º grado de primaria del 2016, 4.º grado de primaria del 2018 y 2.º grado de secundaria del 2019, considerando las mismas Instituciones Educativas. Al realizar los análisis correspondientes, se obtuvieron resultados muy cercanos a los resultados censales de esos mismos años a nivel nacional y en las distintas DRE. Por ejemplo, las mismas IE seleccionadas en la EM 2022 de 2.º grado de primaria fueron utilizadas para sacar una muestra de la ECE 2016 del mismo grado. Los resultados que se obtuvieron a partir de esta muestra fueron muy similares a los resultados censales reportados en ese mismo año. Esto significa que la muestra seleccionada para la EM 2022 está representando adecuadamente a los operativos censales anteriores.

Tabla 2.5 Resultados de 2.º grado de primaria, Lectura

		ECE 2016		Muestra	
		Medida	Medida	LI	LS
Nacional	Medida promedio	584	585,3	583,7	586,9
	En inicio (%)	5,7	5,39	5,13	5,65
	En proceso (%)	46,2	45,71	44,91	46,50
	Satisfactorio (%)	48,1	48,90	48,02	49,78
Hombre	Medida promedio	579	580,2	578,5	581,9
	En inicio (%)	6,2	5,89	5,56	6,23
	En proceso (%)	48,4	47,68	46,84	48,51
	Satisfactorio (%)	45,4	46,43	45,50	47,36
Mujer	Medida promedio	589	590,5	588,8	592,3
	En inicio (%)	5,1	4,88	4,59	5,17
	En proceso (%)	44,0	43,69	42,77	44,61
	Satisfactorio (%)	50,9	51,43	50,43	52,43
Público	Medida promedio	579	580,5	578,9	582,1
	En inicio (%)	6,7	6,21	5,93	6,50
	En proceso (%)	47,1	47,00	46,22	47,79
	Satisfactorio (%)	46,2	46,79	45,92	47,65
Privado	Medida promedio	595	597,7	593,7	601,7
	En inicio (%)	3,0	3,28	2,69	3,86
	En proceso (%)	44,1	42,38	40,37	44,38
	Satisfactorio (%)	52,9	54,34	52,11	56,57

		ECE 2016		Muestra	
		Medida	Medida	LI	LS
Urbano	Medida promedio	597	597,3	595,4	599,3
	En inicio (%)	3,0	3,31	3,04	3,58
	En proceso (%)	41,9	41,61	40,64	42,58
	Satisfactorio (%)	55,0	55,09	54,00	56,17
Rural	Medida promedio	543	540,9	538,9	542,8
	En inicio (%)	12,9	13,10	12,37	13,84
	En proceso (%)	60,0	60,88	59,85	61,91
	Satisfactorio (%)	27,2	26,02	24,97	27,07

Tabla 2.6 Resultados de 2.º grado de primaria, Matemática

		ECE 2016		Muestra	
		Medida	Medida	LI	LS
Nacional	Medida promedio	599	600,8	598,2	603,4
	En inicio (%)	26,6	26,01	25,21	26,80
	En proceso (%)	37,4	37,25	36,67	37,84
	Satisfactorio (%)	36,1	36,74	35,85	37,63
Hombre	Medida promedio	601	603,7	600,9	606,4
	En inicio (%)	26,3	25,52	24,69	26,35
	En proceso (%)	36,5	36,67	35,99	37,35
	Satisfactorio (%)	37,1	37,81	36,83	38,79
Mujer	Medida promedio	597	597,9	595,1	600,7
	En inicio (%)	26,8	26,51	25,58	27,43
	En proceso (%)	38,2	37,85	37,17	38,53
	Satisfactorio (%)	35,0	35,64	34,69	36,59
Público	Medida promedio	609	610,2	607,6	613,1
	En inicio (%)	23,5	22,94	22,29	23,59
	En proceso (%)	37,1	37,30	36,64	37,97
	Satisfactorio (%)	39,4	39,76	38,79	40,72
Privado	Medida promedio	573	576,5	570,2	582,9
	En inicio (%)	34,4	33,92	31,64	36,21
	En proceso (%)	38,1	37,12	35,92	38,33
	Satisfactorio (%)	27,4	28,95	26,96	30,94

		ECE 2016		Muestra	
		Medida	Medida	LI	LS
Urbano	Medida promedio	612	610,4	607,3	613,6
	En inicio (%)	22,2	23,02	22,07	23,98
	En proceso (%)	37,8	37,29	36,59	37,99
	Satisfactorio (%)	40,0	39,69	38,60	40,78
Rural	Medida promedio	567	565,3	562,0	568,5
	En inicio (%)	36,2	37,04	35,87	38,22
	En proceso (%)	37,0	37,11	36,21	38,02
	Satisfactorio (%)	26,8	25,84	24,79	26,89

Tabla 2.7 Resultados de 4.º grado de primaria, Lectura

		ECE 2018		Muestra	
		Medida	Medida	LI	LS
Nacional	Medida promedio	482	483,9	481,8	485,9
	Previo al inicio (%)	10,1	9,7	9,3	10,1
	En inicio (%)	24,2	23,94	23,35	24,53
	En proceso (%)	30,9	30,90	30,36	31,44
	Satisfactorio (%)	34,8	35,42	34,55	36,29
Hombre	Medida promedio	478	479,5	477,0	482,0
	Previo al inicio (%)	11,0	11,0	10,4	11,5
	En inicio (%)	25,2	24,92	24,12	25,71
	En proceso (%)	30,4	30,11	29,44	30,77
	Satisfactorio (%)	33,4	34,01	32,97	35,04
Mujer	Medida promedio	486	488,5	486,4	490,5
	Previo al inicio (%)	9,2	8,4	8,0	8,9
	En inicio (%)	23,2	22,92	22,27	23,57
	En proceso (%)	31,4	31,73	31,08	32,39
	Satisfactorio (%)	36,3	36,90	35,92	37,88
Público	Medida promedio	474	475,6	473,5	477,6
	Previo al inicio (%)	11,7	11,2	10,8	11,7
	En inicio (%)	25,5	25,58	25,03	26,13
	En proceso (%)	30,9	31,08	30,56	31,60
	Satisfactorio (%)	31,9	32,10	31,26	32,93

Privado	Medida promedio	504	506,9	501,7	512,2
	Previo al inicio (%)	5,7	5,6	4,6	6,5
	En inicio (%)	20,5	19,40	17,79	21,01
	En proceso (%)	30,9	30,41	28,96	31,86
	Satisfactorio (%)	42,9	44,63	42,29	46,97
Urbano	Medida promedio	497	497,6	495,1	500,1
	Previo al inicio (%)	6,1	6,3	5,9	6,8
	En inicio (%)	21,8	21,55	20,84	22,26
	En proceso (%)	32,3	32,12	31,46	32,77
	Satisfactorio (%)	39,8	40,01	38,93	41,08
Rural	Medida promedio	433	429,9	427,4	432,4
	Previo al inicio (%)	22,5	23,1	22,2	24,1
	En inicio (%)	32,7	33,36	32,56	34,17
	En proceso (%)	26,6	26,13	25,41	26,85
	Satisfactorio (%)	18,2	17,36	16,55	18,17

Tabla 2.8 Resultados de 4.º grado de primaria, Matemática

		ECE 2018		Muestra	
		Medida	Medida	LI	LS
Nacional	Medida promedio	480	480,5	478,4	482,6
	Previo al inicio (%)	9,3	9,3	8,8	9,7
	En inicio (%)	19,3	18,85	18,33	19,36
	En proceso (%)	40,7	40,64	40,07	41,22
	Satisfactorio (%)	30,7	31,25	30,35	32,14
Hombre	Medida promedio	484	483,4	480,9	485,9
	Previo al inicio (%)	9,1	9,6	9,0	10,2
	En inicio (%)	18,5	18,07	17,43	18,71
	En proceso (%)	39,9	39,44	38,69	40,19
	Satisfactorio (%)	32,5	32,87	31,81	33,93
Mujer	Medida promedio	476	477,5	475,3	479,6
	Previo al inicio (%)	9,5	8,9	8,4	9,3
	En inicio (%)	20,1	19,66	18,97	20,35
	En proceso (%)	41,5	41,91	41,23	42,59
	Satisfactorio (%)	28,9	29,54	28,58	30,50

		ECE 2018		Muestra	
		Medida	Medida	LI	LS
Público	Medida promedio	475	476,9	474,7	479,1
	Previo al inicio (%)	10,5	10,3	9,9	10,7
	En inicio (%)	19,9	19,38	18,88	19,88
	En proceso (%)	40,3	40,21	39,59	40,83
	Satisfactorio (%)	29,4	30,12	29,19	31,04
Privado	Medida promedio	491	490,5	485,4	495,6
	Previo al inicio (%)	6,0	6,4	5,4	7,5
	En inicio (%)	17,7	17,37	16,00	18,74
	En proceso (%)	41,9	41,85	40,53	43,16
	Satisfactorio (%)	34,4	34,37	32,19	36,55
Urbano	Medida promedio	493	492,2	489,6	494,7
	Previo al inicio (%)	6,0	6,5	6,0	7,0
	En inicio (%)	17,2	16,81	16,20	17,43
	En proceso (%)	42,1	41,76	41,07	42,44
	Satisfactorio (%)	34,7	34,92	33,82	36,02
Rural	Medida promedio	438	434,6	432,0	437,2
	Previo al inicio (%)	19,4	20,1	19,2	21,0
	En inicio (%)	26,2	26,85	26,06	27,65
	En proceso (%)	36,7	36,26	35,37	37,15
	Satisfactorio (%)	17,7	16,78	15,94	17,62

Tabla 2.9 Resultados de 2.º grado de secundaria, Lectura

		ECE 2019		Muestra	
		Medida	Medida	LI	LS
Nacional	Medida promedio	567	568,7	567,4	569,9
	Previo al inicio (%)	17,7	17,1	16,7	17,5
	En inicio (%)	42,0	41,42	40,79	42,06
	En proceso (%)	25,8	26,42	26,00	26,84
	Satisfactorio (%)	14,5	15,04	14,42	15,67
Hombre	Medida promedio	563	565,5	564,1	566,9
	Previo al inicio (%)	18,9	18,1	17,6	18,6
	En inicio (%)	43,0	42,17	41,45	42,89

		ECE 2019		Muestra	
		Medida	Medida	LI	LS
	En proceso (%)	25,0	25,93	25,40	26,46
	Satisfactorio (%)	13,1	13,78	13,10	14,45
Mujer	Medida promedio	571	572,0	570,4	573,6
	Previo al inicio (%)	16,4	16,0	15,5	16,5
	En inicio (%)	40,9	40,63	39,85	41,41
	En proceso (%)	26,7	26,95	26,38	27,52
	Satisfactorio (%)	16,0	16,39	15,61	17,17
	Público	Medida promedio	556	557,8	556,5
Previo al inicio (%)		21,3	20,7	20,3	21,2
En inicio (%)		45,1	44,57	43,99	45,15
En proceso (%)		23,1	23,65	23,22	24,07
Satisfactorio (%)		10,6	11,04	10,53	11,55
Privado		Medida promedio	600	603,0	599,3
	Previo al inicio (%)	6,7	5,7	4,9	6,5
	En inicio (%)	32,4	31,52	29,61	33,43
	En proceso (%)	34,2	35,17	34,05	36,28
	Satisfactorio (%)	26,7	27,64	25,61	29,67
	Urbano	Medida promedio	580	580,9	579,3
Previo al inicio (%)		11,6	11,3	10,8	11,7
En inicio (%)		41,1	40,31	39,52	41,09
En proceso (%)		29,7	30,22	29,71	30,73
Satisfactorio (%)		17,7	18,21	17,41	19,00
Rural		Medida promedio	523	524,4	522,9
	Previo al inicio (%)	39,0	38,2	37,2	39,3
	En inicio (%)	45,2	45,46	44,68	46,24
	En proceso (%)	12,3	12,71	12,12	13,30
	Satisfactorio (%)	3,5	3,63	3,27	3,98

Tabla 2.10 Resultados de 2.º grado de secundaria, Matemática

		ECE 2019		Muestra	
		Medida	Medida	LI	LS
Nacional	Medida promedio	567	569,7	567,8	571,6
	Previo al inicio (%)	33,0	32,1	31,5	32,8
	En inicio (%)	32,1	31,78	31,35	32,20
	En proceso (%)	17,3	17,46	17,12	17,80
	Satisfactorio (%)	17,7	18,64	17,96	19,33
Hombre	Medida promedio	573	576,8	574,7	578,9
	Previo al inicio (%)	30,9	29,8	29,1	30,5
	En inicio (%)	31,6	31,41	30,86	31,95
	En proceso (%)	17,8	17,88	17,45	18,31
	Satisfactorio (%)	19,7	20,92	20,15	21,69
Mujer	Medida promedio	560	562,2	560,0	564,4
	Previo al inicio (%)	35,2	34,6	33,8	35,4
	En inicio (%)	32,5	32,17	31,62	32,73
	En proceso (%)	16,7	17,01	16,55	17,48
	Satisfactorio (%)	15,6	16,22	15,43	17,00
Público	Medida promedio	554	556,3	554,5	558,0
	Previo al inicio (%)	37,6	36,8	36,1	37,5
	En inicio (%)	32,9	32,69	32,30	33,08
	En proceso (%)	15,8	15,98	15,65	16,31
	Satisfactorio (%)	13,7	14,52	13,97	15,08
Privado	Medida promedio	607	612,0	606,5	617,4
	Previo al inicio (%)	18,7	17,4	15,8	19,0
	En inicio (%)	29,4	28,90	27,65	30,16
	En proceso (%)	21,8	22,11	21,17	23,06
	Satisfactorio (%)	30,0	31,61	29,41	33,81
Urbano	Medida promedio	581	583,8	581,5	586,1
	Previo al inicio (%)	26,9	26,2	25,5	27,0
	En inicio (%)	32,9	32,37	31,87	32,88
	En proceso (%)	19,3	19,43	19,02	19,84
	Satisfactorio (%)	20,9	21,97	21,11	22,83

		ECE 2019		Muestra	
		Medida	Medida	LI	LS
Rural	Medida promedio	517	518,9	516,7	521,1
	Previo al inicio (%)	54,3	53,4	52,2	54,5
	En inicio (%)	29,0	29,62	28,87	30,36
	En proceso (%)	10,2	10,37	9,87	10,88
	Satisfactorio (%)	6,5	6,64	6,13	7,15

Tabla 2.11 Resultados de 2.º grado de secundaria, Ciencia y Tecnología

		ECE 2019		Muestra	
		Medida	Medida	LI	LS
Nacional	Medida promedio	501	503,1	501,5	504,6
	Previo al inicio (%)	10,1	9,8	9,5	10,0
	En inicio (%)	43,8	43,57	43,00	44,14
	En proceso (%)	36,3	36,65	36,19	37,11
	Satisfactorio (%)	9,7	10,02	9,60	10,43
Hombre	Medida promedio	506	508,2	506,5	510,0
	Previo al inicio (%)	9,6	9,2	8,8	9,5
	En inicio (%)	42,6	42,26	41,58	42,94
	En proceso (%)	37,1	37,39	36,82	37,96
	Satisfactorio (%)	10,7	11,19	10,69	11,68
Mujer	Medida promedio	497	497,5	495,6	499,5
	Previo al inicio (%)	10,7	10,4	10,0	10,8
	En inicio (%)	45,1	44,96	44,26	45,66
	En proceso (%)	35,5	35,86	35,26	36,46
	Satisfactorio (%)	8,7	8,78	8,29	9,27
Público	Medida promedio	488	490,4	488,9	491,9
	Previo al inicio (%)	12,0	11,6	11,3	11,9
	En inicio (%)	47,5	47,09	46,56	47,63
	En proceso (%)	33,2	33,62	33,13	34,12
	Satisfactorio (%)	7,3	7,65	7,34	7,97

		ECE 2019		Muestra	
		Medida	Medida	LI	LS
Privado	Medida promedio	541	542,7	538,2	547,1
	Previo al inicio (%)	4,5	3,9	3,3	4,5
	En inicio (%)	32,3	32,51	30,87	34,16
	En proceso (%)	46,1	46,14	45,00	47,28
	Satisfactorio (%)	17,2	17,45	16,04	18,86
Urbano	Medida promedio	516	516,8	514,9	518,7
	Previo al inicio (%)	7,1	6,9	6,7	7,2
	En inicio (%)	41,2	40,78	40,08	41,47
	En proceso (%)	40,3	40,49	39,95	41,04
	Satisfactorio (%)	11,4	11,79	11,26	12,31
Rural	Medida promedio	451	454,2	451,8	456,7
	Previo al inicio (%)	21,0	19,8	19,0	20,5
	En inicio (%)	53,0	53,47	52,71	54,23
	En proceso (%)	22,3	23,02	22,20	23,84
	Satisfactorio (%)	3,7	3,75	3,38	4,12

Operativo de campo

Capítulo 3

Capítulo 3

Operativo de campo

Al diseñar el operativo de campo, se tuvo en cuenta que los procedimientos y las condiciones de aplicación se alineen con las evaluaciones de los años anteriores de tal manera que los resultados puedan ser comparables. En este marco, la organización, la implementación y los procedimientos del operativo de campo de la EM 2022 se rigieron por los siguientes tres principios:

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

Las estrategias propuestas para alcanzar estos principios se plasmaron en las especificaciones técnicas de los Términos de Referencia del servicio de aplicación de la EM 2022. El Instituto Nacional de Estadística e Informática (INEI) fue el operador logístico responsable de su ejecución en todo el país. Dicha institución tuvo bajo su responsabilidad garantizar los siguientes procesos:

- 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 IE para actualizar información de las IE y de los estudiantes;
- la seguridad y la confidencialidad de los instrumentos desde su llegada a cada local de jurisdicción hasta su recojo de los mismos y
- la correcta aplicación de los instrumentos de evaluación en las IE.

Por su parte, el Minedu asumió por primera vez los siguientes procesos:

- realizar la impresión, el inventario y el modulado de los instrumentos de aplicación (cuadernillos de pruebas, cuestionarios y documentos auxiliares);
- realizar la impresión y el modulado de los materiales de capacitación;
- realizar la distribución de materiales de capacitación;
- realizar la distribución y el repliegue de materiales de aplicación y
- conseguir los locales de las sedes jurisdiccionales a nivel nacional.

Además, el Minedu asignó 193 monitores, uno por sede. Sus principales funciones fueron las siguientes: liderar procesos que eran responsabilidad del Minedu en sus respectivas sedes, asegurar que los procesos de capacitación cumplieran con el objetivo de transmitir de manera estandarizada los procedimientos de aplicación, garantizar que el operador logístico llevara a cabo las acciones necesarias para asegurar la confidencialidad de los instrumentos de evaluación y apoyar en la solución de incidentes..

Teniendo en cuenta el contexto mencionado, se realizaron coordinaciones previas con las IE que eran parte de la muestra para recoger información de la modalidad en la cual se encontraban. Con dicha información, se consideró una misma fecha de inicio de aplicación para todas las IE y grados programados. Para cumplir con este propósito, algunos aplicadores tuvieron que desplazarse días previos al primer día de aplicación con el fin de llegar en las fechas establecidas.

3.1 Organización territorial del operativo

Para gestionar territorialmente el operativo de campo, el país se dividió en 29 sedes de jurisdicciones regionales. La conformación de estas sedes de jurisdicciones regionales no se ciñe estrictamente a los límites del departamento donde se ubican. Para conformarlas, se tuvo en cuenta el criterio de accesibilidad y grados programados. En cada jurisdicción regional de aplicación, se establecieron sedes provinciales y se implementaron, en total, 193 locales de jurisdicción provinciales.

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

Tabla 3.1 Sedes de jurisdicciones regionales de aplicación y distribución de IE según grado a evaluar

Regiones	2.º grado de primaria Regional	4.º grado de primaria Regional	6.º grado de primaria Nacional	2.º grado de secundaria Regional	Total general
Amazonas	349	252	9	215	825
Ancash-Chimbote	90	91	6	64	251
Ancash-Huaraz	63	87	2	91	243
Apurímac	116	91	3	114	324
Arequipa	219	148	8	133	508
Ayacucho	151	123	1	117	392
Cajamarca	423	330	19	171	943
Callao	102	96	10	63	271
Cusco	191	151	8	167	517
Huancavelica	199	158	8	115	480
Huánuco	192	198	7	132	529

Regiones	2.° grado de primaria Regional	4.° grado de primaria Regional	6.° grado de primaria Nacional	2.° grado de secundaria Regional	Total general
Ica	224	166	8	114	512
Junín	259	232	10	198	699
La Libertad	227	249	15	157	648
Lambayeque	191	143	13	101	448
Lima Metropolitana	242	193	55	157	647
Lima-Cañete	56	43	2	36	137
Lima-Huacho	79	64	2	70	215
Loreto	161	110	17	134	422
Madre de Dios	52	62		51	165
Moquegua	43	42	2	32	119
Pasco	153	135	6	100	394
Piura	316	268	23	174	781
Puno	209	178	8	135	530
San Martín-Moyobamba	151	112	6	78	347
San Martín-Tarapoto	211	170	12	127	520
Tacna	81	80	3	56	220
Tumbes	59	64	1	41	165
Ucayali	195	157	5	165	522
Total general	5004	4193	269	3308	12774

3.2 Conformación de la Red Administrativa de aplicación

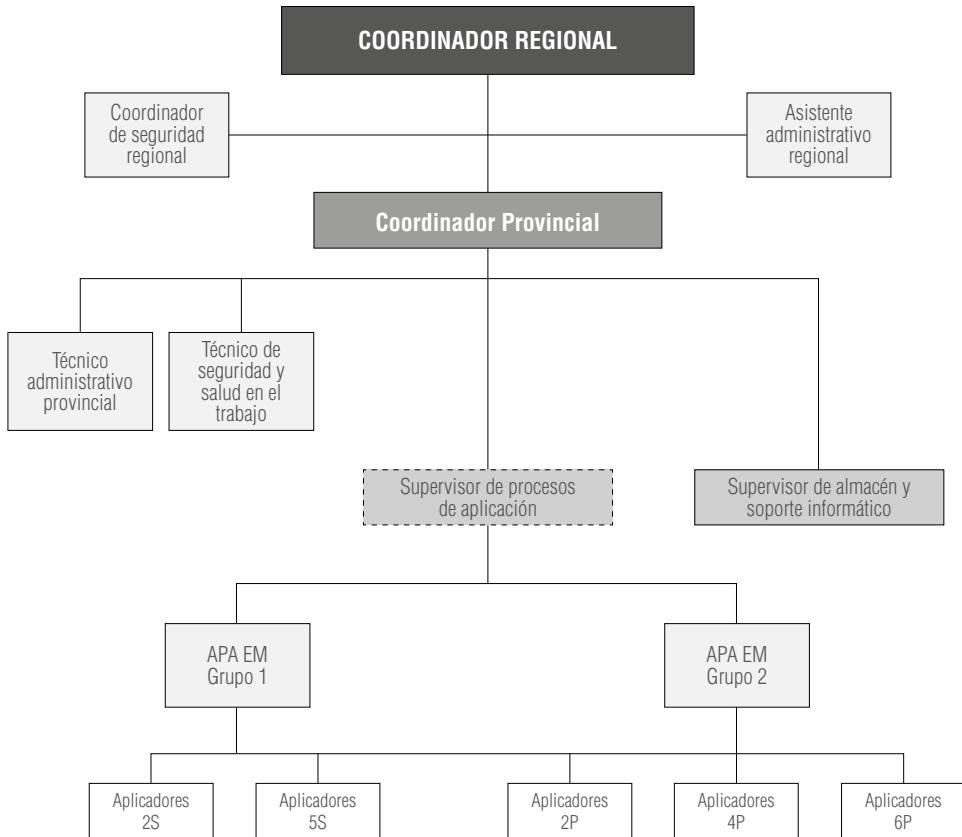
El personal necesario para ejecutar el operativo de campo estuvo organizado de la siguiente manera:

Tabla 3.2 Personal del operativo

Nivel	Cargo en la RA de campo	Total RA 2022
Nivel IA	Supervisores regionales	29
Nivel IB	Coordinadores regionales	29
	Asistente administrativo regional	29
	Coordinador de seguridad regional	29
Nivel IC	Coordinadores provinciales	193
	Técnicos administrativos (incluye todas las sedes provinciales)	193
	Técnico de seguridad y salud en el trabajo	31

Nivel	Cargo en la RA de campo	Total RA 2022
Nivel IIA	Supervisor de procesos de aplicación	26
Nivel IIB	Supervisores de almacén y soporte informático	193
Nivel III	Asistentes de procesos de aplicación	1549
	Aplicadores de 2.º de primaria	8803
	Aplicadores de 4.º de primaria	7899
Nivel IV	Aplicadores de 6.º de primaria	525
	Aplicadores de 2.º de secundaria	8641
	Aplicadores para atención de estudiantes con discapacidad	1948

Figura 3.1 Organigrama de la Red Administrativa EM - 2023



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

- Los supervisores regionales velan por el cumplimiento de los procesos y procedimientos del operativo en la región asignada. Además, garantizan las condiciones materiales necesarias para la evaluación y aseguran la consecución de personal en los diferentes niveles de la RA.
- Los coordinadores regionales y provinciales son responsables de 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 o provincia). 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 técnicos de seguridad y salud en el trabajo apoyan en la implementación de las condiciones materiales y logísticas de las medidas de seguridad y de prevención del COVID-19 en su sede jurisdiccional provincial y/o regional.
- Los supervisores de procesos de aplicación se encargan de la supervisión e implementación de los procesos de convocatoria, selección y capacitación de asistentes de procesos de aplicación. También, lideran y asisten técnicamente a los aplicadores (APA).
- Los supervisores de almacén y soporte informático son responsables de gestionar la recepción, el almacenamiento, el inventario y el despacho del material de aplicación, así como del uso de los sistemas de información implementados para estas actividades.
- Los asistentes de procesos de aplicación son responsables de llevar a cabo 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.

Para cada uno de los niveles de la RA, el Minedu entregó manuales de aplicación y de funciones. En ese sentido, se elaboraron 16 manuales para el personal de la RA que atienden las particularidades de los grados evaluados.

- Manual del coordinador regional y provincial
- Manual del supervisor de almacén y soporte informático
- Manual del usuario del sistema
- Manual del asistente de procesos de aplicación secundaria

- Manual del asistente de procesos de aplicación primaria
- Manual del aplicador 2.º grado de primaria
- Manual del aplicador 4.º grado de primaria
- Manual del aplicador 6.º grado de primaria
- Manual del aplicador 2.º grado de secundaria
- Protocolo del aplicador de apoyo a estudiantes con discapacidad visual - macrotipo
- Protocolo del aplicador de apoyo a estudiantes con discapacidad visual – prueba en braille
- Protocolo del aplicador de apoyo a estudiantes con discapacidad auditiva
- Protocolo del aplicador de apoyo a estudiantes con discapacidad motora
- Protocolo del aplicador de apoyo a estudiantes con discapacidad del Trastorno del Espectro Autista
- Protocolo del aplicador de apoyo a estudiantes con discapacidad intelectual - lector
- Protocolo del aplicador de apoyo a estudiantes con discapacidad intelectual - lector inicial

3.3 Proceso de selección y capacitación del personal de la RA

El proceso de selección del personal estuvo a cargo del INEI y constó de tres etapas cancelatorias:

1. convocatoria y revisión de hojas de vida;
2. preselección; y
3. capacitación.

El Minedu estableció, en los términos de referencia, el perfil profesional para cada uno de los cargos. Para los aplicadores, en la mayoría de los casos, los perfiles fueron los siguientes:

- Se consideraron tres grupos para los aplicadores:
 - I. Con experiencia en la ECE - EM: (no necesitaba cumplir los otros requisitos del perfil)

Experiencia demostrada como aplicador en dos (2) de los operativos ECE 2015, 2016, 2018 o 2019, y que su desempeño no hubiera sido observado.

II. Sin experiencia en la ECE - EM y con el siguiente perfil:

- docente con experiencia en aula, sin carga docente o administrativa en el sector estatal
- egresado universitario o egresado técnico, excepto carreras vinculadas al cuidado personal u oficios
- estudiantes de Institutos Pedagógicos de los últimos tres años o
- estudiantes universitarios de carreras de humanidades de los últimos tres años.

III. Sin experiencia en la ECE - EM y con práctica sustentada en

- la aplicación de instrumentos estandarizados de evaluación educativa;
- la aplicación de instrumentos de recojo de información o
- participación en actividades con niñas, niños y adolescentes o en proyectos sociales.

Con el siguiente perfil:

- o estudiantes universitarios de otras carreras que no sean humanidades, de los últimos tres años;
- o estudiantes de carreras técnicas del último año o
- o estudiantes con estudios de CETPRO (excepto carreras vinculadas al cuidado personal u oficios).

Además, para los grupos II y III se consideran otras características:

- edad entre 19 y 60 años, salvo excepciones que deberán ser aprobadas por la UMC.
- deseable: habilidades para el buen manejo de grupos de niños y adolescentes.
- deseable: experiencia de trabajo con niños y niñas, o adolescentes en aula o animación sociocultural o aplicación de pruebas estandarizadas de rendimiento escolar.
- buen nivel de lectura comprensiva.
- adecuada dicción.
- obligatoriedad de asistir a las sesiones de capacitación.
- disposición a tiempo completo durante las fechas de capacitación y aplicación, y para viajar o trasladarse al lugar que se le designe.
- deseable: que disponga de teléfono móvil.

- Para el caso de aplicadores para la atención de estudiantes con discapacidad, se solicitó como mínimo (en orden de prioridad):
 - egresado de carrera de educación especial;
 - estudiante universitario de carreras de educación especial de los dos últimos años;
 - egresado de educación superior universitaria, titulado técnico con especialidad, diplomado en educación inclusiva o con experiencia en la atención de estudiantes con discapacidad;
 - egresados técnicos o estudiantes del último año de la carrera de educación, enfermería, fisioterapia y rehabilitación o carreras afines o
 - estudiantes universitarios de los dos últimos años o egresados de las carreras de psicología, educación, educación especial, trabajo social, sociología, antropología, comunicación social, enfermería u obstetricia o tecnología médica con mención en terapia física, terapia de lenguaje o terapia ocupacional o carreras afines.
 - Contar con:
 - o Una (1) experiencia de trabajo con niños y niñas, o adolescentes o con población vulnerable (personas con discapacidad, poblaciones indígenas, adultos mayores o afines).
 - o Para atención de estudiantes que presentan discapacidad auditiva (según registro RNEE): la persona deberá manejar la Lengua de Señas Peruana (LSP)
 - o Para atención de estudiantes con autismo, deseable: manejo del sistema de comunicación por intercambio de tarjetas.

Teniendo en cuenta el contexto de emergencia sanitaria, las capacitaciones se programaron de manera descentralizada. Para este operativo, los asistentes de procesos de aplicación y los aplicadores se capacitaron en sus sedes provinciales.

Con el objetivo de desarrollar capacidades en los equipos de trabajo y de asegurar que recibieran 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 dicha capacitación, la metodología incluyó actividades de lectura individual y grupal, ejercicios prácticos para el uso de las fichas de recojo de información y actividades sobre manejo de procedimientos y documentos de aplicación, así como actividades de simulación de las indicaciones para estudiantes en el aula 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 tuvo una duración de dos días para todos los aplicadores. Las aulas estuvieron conformadas por un máximo de 25 candidatos. Además, los aplicadores que atendieron a estudiantes con discapacidad recibieron medio día adicional de capacitación, contemplada en sus respectivos manuales, según el tipo de discapacidad que iban a atender.

Durante la capacitación, se instruyó a los candidatos en la aplicación estandarizada de instrumentos. Para ello, se utilizó el Manual del aplicador y los siguientes 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:

2.º grado de primaria

- Cuadernillos integrados de pruebas (un cuadernillo por día de evaluación)
- Ficha Óptica de Asistencia y Respuestas (FOAR)
- Cuestionario para la familia
- Carta al director
- Cuestionario al director
- Cuestionario al docente tutor
- Compromiso de confidencialidad para el docente
- Ficha del aplicador líder
- Registro de préstamo
- Ficha de verificación

4.º grado de primaria

- Cuadernillos integrados de pruebas (un cuadernillo por día de evaluación)
- Ficha Óptica de Asistencia y Respuestas (FOAR)
- Cuestionario para la familia
- Carta al director
- Cuestionario al director
- Compromiso de confidencialidad para el docente
- Cuestionario al docente de matemática
- Cuestionario al docente de lectura

- Ficha del aplicador líder
- Registro de préstamo
- Ficha de verificación

6.º grado de primaria

- Cuadernillos de pruebas (un cuadernillo por día de evaluación)
- Ficha Óptica de Asistencia y Respuestas (FOAR)
- Cuestionario al estudiante HSE
- Cuestionario al estudiante, factores asociados
- Cuestionario para la familia
- Carta al director
- Cuestionario al director
- Compromiso de confidencialidad para el docente
- Cuestionario al docente tutor
- Ficha del aplicador líder
- Registro de préstamo
- Ficha de verificación

2.º grado de secundaria

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

Las cajas de aplicación que llegaron a cada sede provincial fueron inventariadas y dispuestas en el almacén del local de jurisdicción, el cual permaneció bajo llave y fue custodiado por la RA y contó con resguardo policial o privado hasta el final del operativo.

3.5 Proceso de aplicación

La duración y las fechas de aplicación de cada operativo de la EM 2022 se muestran en la siguiente tabla:

Tabla 3.3 Duración y fechas de la evaluación

Operativo	Duración (días)	Fechas de evaluación*	Criterio
2.º grado de primaria	2	Del 28 de noviembre al 02 de diciembre	Un aplicador líder por IE que requiera pernocte de los aplicadores
			Un aplicador por sección programada
			Un aplicador de apoyo solo para secciones que estén organizadas en dos grupos y que reciben clases de manera simultánea
			Un aplicador para cada estudiante con discapacidad según registro de información UMC_ R-NEE
4.º grado de primaria	2	Del 28 de noviembre al 02 de diciembre	Un aplicador líder por IE que requiera pernocte de los aplicadores
			Un aplicador por sección programada
			Un aplicador de apoyo solo para secciones que estén organizadas en dos grupos y que reciben clases de manera simultánea
			Un aplicador para cada estudiante con discapacidad según registro de información UMC_ R-NEE
6.º grado de primaria	2	Del 28 de noviembre al 02 de diciembre	Un aplicador líder por IE que requiera pernocte de los aplicadores
			Un aplicador por sección programada
			Un aplicador de apoyo solo para secciones que estén organizadas en dos grupos y que reciben clases de manera simultánea
			Un aplicador para cada estudiante con discapacidad según registro de información UMC_ R-NEE

Operativo	Duración (días)	Fechas de evaluación*	Criterio
2.º grado de secundaria	3	Del 7 al 11 de noviembre*	Un aplicador líder por IE que requiera pernocte de los aplicadores
			Un aplicador por sección programada
			Un aplicador de apoyo solo para secciones que estén organizadas en dos grupos y que reciben clases de manera simultánea
			Un aplicador para cada estudiante con discapacidad según registro de información UMC_R-NEE

(*) La duración en días de aplicación se planteó a lo largo de la semana para considerar la posibilidad de que una misma sección requiera de más días para su participación.

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

Aplicación primaria EM 2022 (2.º, 4.º y 6.º grado)

- I. Recepción de materiales. Los aplicadores de las IE que requieren pernocte se desplazaron días antes del primer día de aplicación para llegar a tiempo a las fechas programadas. En las IE que no requieren pernocte, el aplicador acudió cada día de aplicación a la sede para recoger los instrumentos de evaluación.
- II. Presentación ante el director de la IE y organización de la aplicación. El aplicador llegó temprano a la IE, se presentó ante el director, le entregó los documentos enviados por el Minedu y le solicitó la nómina de matrícula del grado a evaluar y la lista auxiliar de asistencia del docente. El docente tutor firmó el compromiso de confidencialidad.
- III. Preparación de los materiales antes de la evaluación. El aplicador solicitó al director un espacio privado para preparar los cuadernillos y demás instrumentos. Solo y sin ayuda, el aplicador verificó los instrumentos que iba aplicar. Adicionalmente, registró los nombres de los estudiantes en los cuadernillos de pruebas con la información de la nómina de matrícula.
- IV. Preparación del aula e indicaciones. El aplicador organizó el mobiliario del aula, ubicó a los estudiantes según el orden de la nómina, entregó los cuadernillos de las pruebas y dio lectura a las indicaciones de acuerdo con el guion de aplicación.
- V. Orden y duración de las sesiones de aplicación. El aplicador siguió el guion de aplicación y administró los cuadernillos siguiendo la secuencia y los tiempos indicados en la siguiente tabla:

Tabla 3.4 Orden y duración de las sesiones de aplicación de los grados de primaria EM 2022

2.° grado de primaria		4.° grado de primaria		6.° grado de primaria	
Primer día	Segundo día	Primer día	Segundo día	Primer día	Segundo día
Lectura (45 m)	Matemática (45 m)	Lectura (60 m)	Matemática (60 m)	Lectura (60 m)	Matemática (60 m)
Descanso (30 m)	Descanso (30 m)	Descanso (30 m)	Descanso (30 m)	Descanso (30 m)	Descanso (30 m)
Matemática (45 m)	Lectura (45 m)	Matemática (60 m)	Lectura (60 m)	Cuestionario HSE (30 m)	Cuestionario FFAA (30 m)

VI. Después de la aplicación, cada día, el aplicador se dirigió a un ambiente privado dentro de la IE para realizar el traslado de las respuestas de los cuadernillos a la FOAR. Luego, el director verificó que el traslado haya sido correcto, respetando los procedimientos. Al culminar la revisión, el aplicador contó la totalidad de cuadernillos y procedió a lacrarlos utilizando una bolsa de seguridad. El director presenció el conteo y firmó un acta de conformidad, dando fe de la cantidad de cuadernillos que retornan al Minedu.

Al terminar las actividades previstas en la IE, el aplicador retornó inmediatamente a la sede para entregar todos los instrumentos de aplicación de la EM.

VII. Al finalizar el último día de la aplicación, el director o personal responsable de la IE dio conformidad a las condiciones de la aplicación en una declaración jurada de aplicación, la cual fue firmada para cada sección evaluada. Para los casos en los que no se concretó la aplicación, se contó con una constancia de no aplicación, la cual detallaba el motivo de la no aplicación, el mismo que fue corroborado con una constancia de la UGEL.

Al culminar el trabajo en la IE, el aplicador se dirigió a la sede donde el asistente de procesos verificó 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 soporte informático hicieron un inventario de los instrumentos. Las fichas ópticas fueron retiradas de sus cajas, clasificadas y ordenadas según correlativo e instrumento. Los paquetes de cuadernillos retornaron en sus mismas cajas.

Aplicación secundaria EM 2022 (2.º grado de secundaria)

- I. Recepción de materiales. En las IE que no requerían pernocte, el aplicador acudió cada día de aplicación a la sede para recoger los instrumentos de evaluación. Los aplicadores de las IE que requerían pernocte se desplazaron con el tiempo necesario para iniciar la aplicación en las fechas programadas.
- II. Presentación ante el director de la IE y organización de la aplicación. El aplicador llegó temprano a la IE, se presentó ante el director, le entregó los documentos enviados por el Minedu y le solicitó la nómina de matrícula del grado a evaluar y la lista auxiliar de asistencia del docente. Se le solicitó que designara a un docente para el apoyo en la organización de los estudiantes, y se le solicitó que firmara el compromiso de confidencialidad.
- III. Preparación del aula e indicaciones. El aplicador organizó el mobiliario del aula, ubicó a los estudiantes según el orden del registro de estudiantes, entregó los cuadernillos de las pruebas y las fichas de respuestas, y luego dio lectura a las indicaciones de acuerdo con el guion de aplicación.
- IV. Orden y duración de las sesiones de aplicación. El aplicador administró los cuadernillos siguiendo la secuencia y los tiempos indicados en la siguiente tabla:

Tabla 3.5 Orden y duración de las sesiones de aplicación de 2.º grado de secundaria

Grado y operativo	Primer día	Segundo día	Tercer día
2.º grado de secundaria	Lectura (60 m)	Matemática (70 m)	Ciencia y Tecnología (80 m)
	Descanso (30 m)	Descanso (30 m)	Descanso (30 m)
	Matemática (70 m)	Lectura (60 m)	Cuestionario HSE (30 m)
	Cuestionario día 1 (25 m)	Cuestionario día 2 (25 m)	

- V. Después de la aplicación, cada día, en un ambiente privado, el aplicador revisó las fichas de respuestas y organizó todos los instrumentos que debían ser devueltos al Minedu.

- VI. Al finalizar cada día de aplicación, el aplicador, en la dirección, contó la totalidad de cuadernillos y procedió a lacrarlos utilizando una bolsa de seguridad. El director, que presenció el conteo, firmó un acta de conformidad respecto de las condiciones de aplicación, dando fe de la cantidad de cuadernillos que retornaban al Minedu. Al terminar las actividades previstas en la IE, el aplicador retornó inmediatamente a la sede para entregar todos los instrumentos de aplicación.
- VII. Al finalizar el último día de la aplicación, el director o personal responsable de la IE dio conformidad a las condiciones de la aplicación en una declaración jurada de aplicación, la cual fue firmada para cada sección evaluada. Para los casos en los que no se concretó la aplicación, se contó con una constancia de no aplicación, la cual detallaba el motivo de la no aplicación, el mismo que fue corroborado con una constancia de la UGEL.
- VIII. Al culminar el trabajo en la IE, el aplicador se dirigió a la sede donde el asistente de procesos verificó 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 soporte informático hicieron un inventario de los instrumentos. Las fichas ópticas fueron retiradas de sus cajas, clasificadas y ordenadas según correlativo e instrumento. Los paquetes de cuadernillos retornaron en sus mismas cajas.

Procesamiento de datos y análisis psicométrico

Capítulo 4

Capítulo 4

Procesamiento de datos y análisis psicométrico

Para el procesamiento de datos y el análisis psicométrico, 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). De esta forma, garantiza la calidad y comparabilidad de la información obtenida de sus evaluaciones.

4.1 Gestión y depuración de datos

La gestión y depuración de datos tiene como finalidad convertir la información contenida en los documentos físicos (fichas ópticas) en información digital (base de datos) consistente y confiable. Este proceso es realizado por una empresa especializada en digitalización y captura de datos atendiendo a una serie de especificaciones técnicas delineadas por la UMC. Es de suma importancia que esta tarea sea desarrollada cumpliendo los plazos establecidos, por lo que la empresa responsable también debe tener la capacidad logística necesaria para trabajar con una gran cantidad de datos.

La UMC se encarga de supervisar todos los procesos, subprocesos y actividades que realice la empresa de captura de datos. Esta última es la responsable de contratar a todo el personal necesario para el cumplimiento de todos los procesos y subprocesos que se lleven a cabo. La UMC, por su parte, es responsable de la capacitación de este personal.

De manera general, este proceso está dividido en tres etapas: i) gestión física, ii) digitalización y captura de datos, y iii) depuración de bases de datos.

4.1.1 Gestión física

Este proceso tiene como objetivo asegurar que la empresa contratada cuente con toda la documentación física a procesar, la cual debe estar completa y lista para la captura de datos. Por ello, todos los documentos y cajas utilizados en la evaluación tienen impreso un código de barras que los identifica de forma única en todo el proyecto. Además, todos los documentos vienen clasificados en cajas según su tipo. Por ejemplo, hay cajas distintas para fichas ópticas de respuestas, fichas ópticas de información complementaria de la aplicación, cuestionarios de factores asociados, cuadernillos o pruebas.

En esta etapa, se realizan las siguientes actividades:

- Recepción e inventario. Se reciben todas las cajas con los documentos físicos a procesar y se realiza el inventario en dos momentos: primero se revisa la cantidad total de cajas y luego se revisa el contenido de cada caja. Este inventario se realiza escaneando el código de barras impreso en cada caja y en cada documento.
- Clasificación (en caso haya documentos sin clasificar). Se clasifican todos los documentos físicos según sus características.
- Preparación y loteo. Se quitan las grapas de los documentos y se limpian los residuos dejados por el uso de borrador en las fichas ópticas para que esto no altere la imagen a digitalizar y, por ende, la captura de datos. Luego, se divide el total de documentos en paquetes pequeños (lotes) que puedan gestionarse fácilmente.
- Almacenamiento. Se resguardan, en un almacén, todos los documentos físicos de forma ordenada para poder acceder a estos en caso sean requeridos en etapas siguientes.

Para asegurar una correcta gestión física de los documentos, la empresa contratada proporciona un local con el espacio suficiente para todas las etapas, incluyendo un espacio exclusivo para el almacenamiento de todos los documentos, dado su carácter confidencial.

4.1.2 Digitalización y captura de datos

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

- Digitalización. Consiste en pasar por un escáner todos los documentos físicos inventariados. Los escáner funcionan con un software de control de calidad que valida que la cantidad de documentos digitalizados sea la misma que la cantidad de documentos inventariados.
- Captura de datos. Utilizando las imágenes digitalizadas y un software desarrollado por la empresa contratada, se capturan los datos 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. Cuando el software de reconocimiento de datos presenta un bajo nivel de confiabilidad, un grupo de personas se dedica a revisar y asegurar que los datos capturados se correspondan con los datos consignados en la imagen del documento.
- Revisiones previas a la depuración. Finalizada la captura de datos, se debe

asegurar que los archivos que se utilizarán en la depuración se han generado siguiendo las especificaciones brindadas a la empresa contratada por la UMC. Las principales revisiones son las siguientes:

- Cantidad de campos igual a la del diccionario de datos
- Longitud y tipo de dato de todos los campos según diccionario de datos
- Valores fuera de rango permitido según diccionario de datos
- Campos de identificación completos
- Correspondencia entre la data capturada y la imagen del documento de una muestra aleatoria del 10 % de todo el archivo.

4.1.3 Depuración de datos

En esta etapa, 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 ello, la UMC elabora un manual del depurador, el cual contiene el detalle de todas las revisiones a realizar en cada campo de todas las bases de datos. Con este manual, se capacita a todo el personal que participa de este proceso.

Durante esta etapa, algunas de las revisiones realizadas son las siguientes:

- Revisiones iniciales:
 - 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 el código modular y el 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 los correlativos de los estudiantes

- Duplicados en los correlativos de los estudiantes
- Caracteres no válidos en el sexo de los estudiantes
- Caracteres no válidos en la lengua materna de los estudiantes
- Caracteres no válidos en la discapacidad de los estudiantes
- Caracteres no válidos en los números de forma de los cuadernillos
- Caracteres no válidos en las respuestas de los estudiantes
- Inconsistencias en las respuestas y la asistencia de los estudiantes
- Revisiones finales:
 - Repetir todas las revisiones iniciales
 - Contrastar el listado de IE aplicadas, el cual es proporcionado por el equipo encargado del trabajo de campo de la UMC
 - Identificar valores fuera del rango permitido

4.2 Estrategias de análisis psicométrico

4.2.1 Modelo Rasch para ítems dicotómicos

El análisis psicométrico de las pruebas aplicadas en el 2022 se basó en el modelo Rasch para ítems dicotómicos (Bond et al., 2021; Wright y Stone, 1998), el cual se centra en el análisis de cada ítem, concretamente en la interacción entre un ítem y una persona. Este modelo 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 de dificultad del ítem utilizado (β)¹. Por este motivo, se le denomina usualmente modelo de un parámetro (Hambleton et al., 1991).

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 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 que el 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

¹ En 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.

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:

$$\theta > \beta; p(X = 1|\theta, \beta) \in]0,5; 1,0[$$

$$\theta < \beta; p(X = 1|\theta, \beta) \in [0,0; 0,5[$$

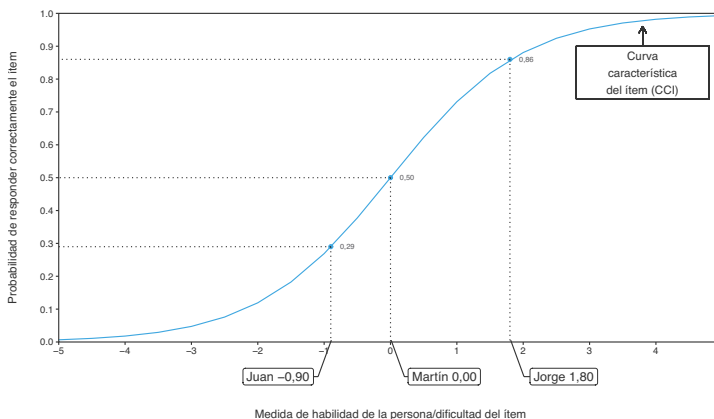
$$\theta = \beta; p(X = 1|\theta, \beta) = 0,5$$

En el primer caso, si la habilidad de la persona es mayor que la dificultad del ítem, la probabilidad de responder correctamente el 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 que 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:

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

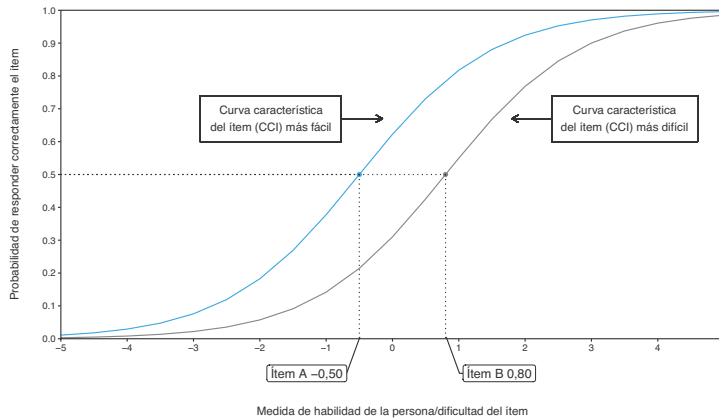
Donde θ representa el rasgo latente de cada persona n , y β representa la medida de dificultad de cada ítem i . Esta relación entre la habilidad de una persona y la dificultad de un ítem se presenta en la figura 4.1.

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



Como se observa, Juan tiene una habilidad de $-0,90$, por lo que su probabilidad de acertar este ítem es de $0,29$; es decir, lo más probable es que falle y obtenga 0 puntos. En cambio, Jorge, 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 la misma probabilidad de acertar o de 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, la dificultad de un ítem se expresa numéricamente como el valor en la escala de habilidad que posee una probabilidad del 50% de acertar dicho ítem. En la figura 4.2, 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 la pudiese evaluar en condiciones óptimas y con todos los ítems que potencialmente pueden usarse para medir dicho rasgo latente. No obstante, como esto 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 lo 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:

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

4.2.2 Modelo Rasch de crédito parcial

Masters (1982) introdujo el modelo de crédito parcial para trabajar ítems politómicos de categorías ordenadas. Se deriva de las pruebas 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 aplicable en pruebas 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:

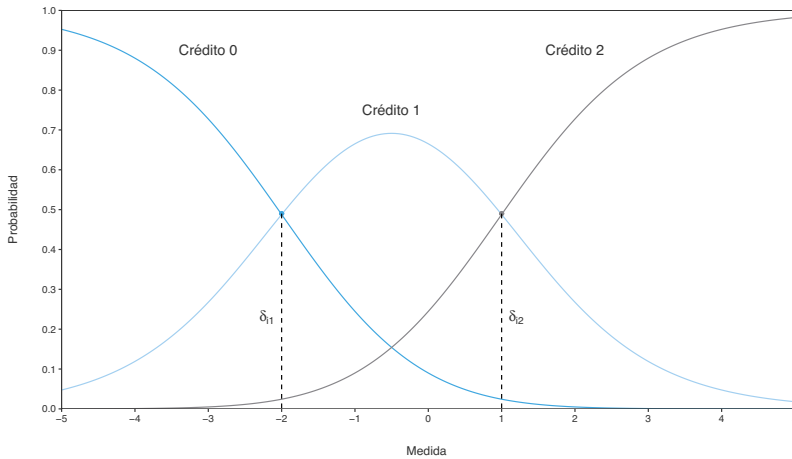
$$\sqrt{\frac{196}{4}} = x$$

Existe una secuencia de 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 el 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_i y F_{ij} , de tal manera que

$$\delta_{ij} = D_i + F_{i1}$$

y

$$\delta_{ij} = D_i + F_{i2}$$

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, además, se pueden señalar tres situaciones para una persona: si esta tiene una medida del rasgo latente menor a $-1,74$, lo más probable es que reciba 0 puntos (crédito 0) en este ítem; si su medida se encuentra entre $-1,74$ y $0,52$, lo más probable es que reciba 1 punto (crédito 1); y, 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 personas 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:

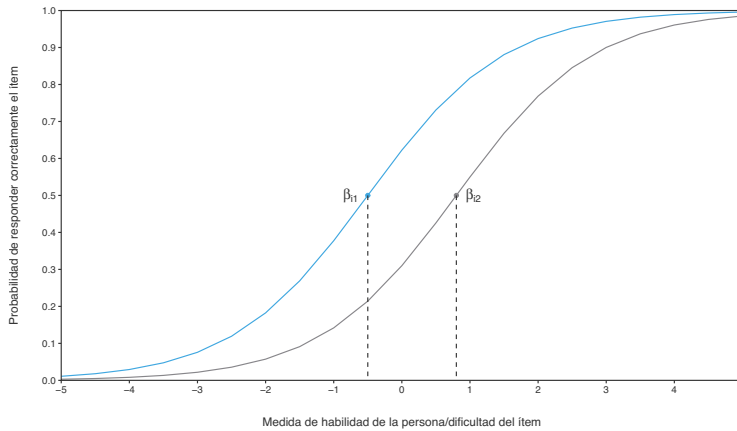
$$P(X_{ni} = 0) = \frac{\exp(0)}{\exp(0) + \exp(0 + (\theta_n - \delta_{i1})) + \exp(0 + (\theta_n - \delta_{i1}) + (\theta_n - \delta_{i2}))}$$

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

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

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 en la figura 4.4.

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).

4.2.3 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 respecto del modelo suele dar lugar a la desestimación del modelo estadístico. 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. 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 y no cómo el modelo se ajusta a los datos (Bond et al., 2021). Como

señala Wilson (2005), el ajuste de los ítems al modelo Rasch se fundamenta en dos datos estadísticos: el outfit y el infit, cuyo cálculo se basa en los residuos cuadráticos estandarizados del modelo (Z^2_{ni}). A partir de las propuestas de diversos autores (Bond et al., 2021; 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^2_{ni}), de tal manera que

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

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:

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

Como $\sum_{n=1}^N R^2_{ni} = VAR(P_{ni}) Z^2_{ni}$, se puede redefinir el infit como una media cuadrática ponderada de residuales (Z^2_{ni}) que es sensible a patrones de respuesta irregulares.

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

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

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

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 los datos presentan un 25 % más de “ruido” 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, 2023):

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

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

Siguiendo las recomendaciones de Linacre (2023), 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, 2023).

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 varían 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:

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

4.2.4 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 de la prueba. 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 es 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. Por ello, 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 al., 1991). No obstante, 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 un 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 busca reducir las variables a un conjunto

menor de componentes enfocándose en la varianza total (Park et al., 2002; Preacher y MacCallum, 2003).

En el caso de los modelos Rasch, lo usual es efectuar un análisis de componentes principales de los residuos estandarizados luego de ajustar el modelo Rasch unidimensional (Linacre, 2023). Este análisis de los residuos sirve para detectar cualquier varianza que quede en los datos luego de extraer la dimensión principal (Bond et al., 2021), lo cual se realiza tratando de modelar la varianza que no es explicada por el constructo que se pretende medir. La idea básica detrás de este análisis es que, si las respuestas cumplen adecuadamente con el requerimiento de unidimensionalidad, no debería haber patrones significativos en los residuos. En otras palabras, no debería haber una variable latente adicional, distinta de la variable latente principal, que haya tenido un impacto en las respuestas a los ítems (Wind y Hua, 2022).

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.

De igual modo, Linacre (2023) 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, entonces no existen serias evidencias que atenten contra el supuesto de unidimensionalidad y, por lo tanto, los datos se pueden analizar adecuadamente utilizando modelos Rasch. Asimismo, se está frente a datos unidimensionales cuando la cantidad de varianza empírica es similar a la cantidad de varianza predicha por el modelo.

4.2.5 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 por la UMC. Este índice es análogo al de consistencia interna alpha (α) (Cronbach, 1951), pero produce mejores estimaciones, pues los valores numéricos son lineales si los datos se ajustan al modelo Rasch aplicado (Wind y Hua, 2022). 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 del 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 et al., 2021).

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 mucha variabilidad; asimismo, observó valores propuestos con los diversos métodos para obtener la confiabilidad, los cuales 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 (1995).

Estos coeficientes también pueden ser expresados como índices de separación de personas. Estos índices se refieren a la dispersión de los datos medidos como el número de errores estándar que separan a las personas (Schumacker y Smith, 2007).

Otros indicadores importantes para el análisis son los de consistencia y precisión de la clasificación. Por un lado, la consistencia de la clasificación es el grado de acuerdo en dos aplicaciones independientes o paralelas de un instrumento de medición. Se espera que el alumno sea clasificado en la misma categoría al repetirse la evaluación. Por su parte, la precisión de la clasificación implica el grado en el cual la clasificación observada coincide con la clasificación verdadera (Kim et al., 2006). En las evaluaciones realizadas, 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.

4.3 Resultados del análisis psicométrico en Lectura, Matemática, y Ciencia y Tecnología

4.3.1 Calibración de los ítems

En las tablas 4.1, 4.2 y 4.3, se muestra el detalle de las características psicométricas de los ítems de la EM 2022 de 2.º grado de secundaria. Asimismo, en las tablas 4.4, 4.5, 4.6, 4.7, 4.8 y 4.9, se muestran las características psicométricas de los ítems de las EM de 2.º, 4.º y 6.º grado de primaria. Tenga en cuenta que Bloque A corresponde a las formas con un mayor nivel de dificultad y Bloque B a las formas con menor nivel de dificultad.

Tabla 4.1 Dificultad y ajuste de los ítems al modelo Rasch de la EM 2022 de Lectura de 2.º grado de secundaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT001	B06	B06	2,230	0,011	0,93	0,92	0,47	0,32	4
CT002	B06	B06	1,813	0,010	1,00	1,02	0,42	0,42	4
CT003	B06	B06	1,339	0,010	1,03	1,04	0,40	0,53	3
CT004_1	B06	B06	0,430	0,007	1,11	1,13	0,50	1,18	2
CT004_2	B06	B06	1,787	0,007	1,11	1,13	0,50	1,18	4
CT005_1	B06	B06	0,681	0,007	1,19	1,25	0,43	0,96	2
CT005_2	B06	B06	2,422	0,007	1,19	1,25	0,43	0,96	4
CT006	-	B19	0,268	0,016	0,94	0,89	0,49	0,65	2
CT007	B16	-	0,563	0,013	0,91	0,86	0,45	0,74	2
CT008	B16	-	0,975	0,012	1,06	1,12	0,33	0,66	3
CT009	B16	-	1,007	0,012	1,12	1,15	0,29	0,65	3
CT010	B16	-	1,138	0,012	0,96	0,94	0,44	0,62	3
CT011	B06	B06	0,545	0,011	1,01	0,97	0,40	0,71	2
CT012	B13	B13	-0,312	0,017	0,93	0,79	0,44	0,83	1
CT013	B07	-	2,375	0,012	0,96	0,98	0,43	0,32	4
CT014	B09	B09	0,749	0,014	0,89	0,81	0,52	0,66	2
CT015	B09	B09	1,998	0,014	1,17	1,32	0,29	0,37	4
CT016	B09	B09	0,594	0,014	1,01	0,96	0,42	0,69	2
CT017	B09	B09	0,936	0,013	0,92	0,86	0,50	0,61	2
CT018	B09	B09	1,965	0,013	1,12	1,23	0,33	0,37	4
CT019	B01	B01	-1,392	0,012	0,85	0,58	0,39	0,94	1
CT020	B01	B01	0,557	0,007	0,93	0,88	0,48	0,69	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT021	B01	B01	-0,837	0,010	0,92	0,76	0,39	0,9	1
CT022	B01	B01	-0,517	0,009	0,85	0,69	0,46	0,87	1
CT023	B01	B01	0,955	0,006	1,06	1,11	0,38	0,61	3
CT024	B01	B01	1,980	0,006	0,87	0,91	0,52	0,37	4
CT025	B01	B01	1,902	0,078	0,78	0,70	0,60	0,39	4
CT026	B16	-	1,277	0,012	0,85	0,80	0,54	0,59	3
CT027	B17	B17	2,468	0,011	1,11	1,34	0,30	0,28	4
CT028	B17	B17	2,818	0,012	1,01	1,43	0,34	0,22	4
CT029	B17	B17	1,887	0,010	1,01	1,12	0,41	0,40	4
CT030	B17	B17	1,984	0,010	0,88	0,98	0,51	0,38	4
CT031	B17	B17	2,331	0,011	1,08	1,27	0,33	0,31	4
CT032	B17	B17	2,078	0,011	0,96	1,03	0,45	0,36	4
CT033	B09	B09	1,751	0,013	0,95	0,97	0,47	0,42	3
CT034	B02	B02	1,503	0,056	0,98	0,97	0,44	0,47	3
CT035	B08	-	2,478	0,012	1,21	1,45	0,19	0,31	4
CT036	B08	-	1,651	0,011	1,06	1,08	0,36	0,50	3
CT037	B08	-	1,777	0,012	0,98	0,98	0,43	0,47	4
CT038	B03	B03	0,643	0,006	0,96	0,92	0,46	0,67	2
CT039	B03	B03	1,585	0,006	1,26	1,44	0,22	0,46	3
CT040	B03	B03	0,853	0,006	1,10	1,20	0,35	0,63	2
CT041	B03	B03	1,385	0,006	1,04	1,09	0,40	0,50	3
CT042	B03	B03	0,921	0,057	1,05	1,14	0,37	0,61	2
CT043	-	B20	-0,126	0,017	0,90	0,89	0,50	0,73	1
CT044	-	B20	-1,016	0,022	0,83	0,61	0,49	0,87	1
CT045	-	B20	0,279	0,016	0,86	0,79	0,55	0,65	2
CT046	-	B20	0,705	0,015	0,99	1,02	0,45	0,56	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT047	-	B20	0,105	0,016	0,94	0,86	0,49	0,69	2
CT048	B05	B05	-0,112	0,013	1,03	1,2	0,33	0,82	1
CT049	B05	B05	-0,136	0,013	0,96	0,88	0,40	0,82	1
CT050	B05	B05	1,535	0,010	0,94	0,95	0,48	0,48	3
CT051	B15	-	-0,638	0,019	1,02	1,00	0,24	0,91	1
CT052	B15	-	2,291	0,012	1,00	1,02	0,40	0,34	4
CT053	B15	-	1,178	0,012	1,00	0,98	0,41	0,60	3
CT054	B15	-	2,376	0,012	1,10	1,27	0,29	0,32	4
CT055	B15	-	1,328	0,011	1,04	1,06	0,37	0,57	3
CT056	B05	B05	0,366	0,011	0,86	0,77	0,52	0,74	2
CT057	B03	B03	0,638	0,006	1,13	1,21	0,32	0,67	2
CT058	B12	B12	0,718	0,012	1,13	1,20	0,34	0,62	2
CT059	B12	B12	0,02	0,013	0,83	0,72	0,54	0,76	2
CT060	B12	B12	1,374	0,011	1,00	1,02	0,45	0,47	3
CT061	B11	B11	0,569	0,014	1,03	1,05	0,41	0,67	2
CT062	B11	B11	-	-	-	-	-	-	-
CT063	B11	B11	2,016	0,014	1,11	1,32	0,33	0,34	4
CT064	B11	B11	1,13	0,013	1,10	1,17	0,36	0,55	3
CT065	B11	B11	-	-	-	-	-	-	-
CT066	-	B10	-0,412	0,018	0,88	0,76	0,50	0,78	1
CT067	-	B10	1,374	0,015	1,17	1,25	0,30	0,39	3
CT068	-	B10	0,111	0,016	0,98	0,98	0,44	0,68	2
CT069	-	B18	0,094	0,016	0,99	0,96	0,44	0,69	2
CT070	-	B18	0,137	0,016	1,09	1,11	0,36	0,68	2
CT071	-	B18	-	-	-	-	-	-	-
CT072	-	B18	0,376	0,015	1,00	0,97	0,44	0,62	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT073	-	B18	0,462	0,015	1,00	0,98	0,44	0,60	2
CT074	-	B18	0,574	0,015	1,11	1,15	0,35	0,58	2
CT075	B03	B03	0,715	0,006	0,96	0,93	0,46	0,66	2
CT076	B05	B05	0,691	0,011	0,92	0,89	0,48	0,68	2
CT077	B05	B05	1,436	0,010	1,12	1,23	0,33	0,50	3
CT078	B08	-	0,859	0,012	1,05	1,13	0,33	0,68	2
CT079	B08	-	1,049	0,012	1,04	1,07	0,36	0,64	3
CT080	-	B10	-0,426	0,018	0,97	1,10	0,41	0,78	1
CT081	-	B10	-0,474	0,018	0,91	0,80	0,47	0,79	1
CT082	-	B10	-	-	-	-	-	-	-
CT083	B11	B11	0,900	0,013	1,09	1,12	0,38	0,60	2
CT084	B12	B12	0,809	0,011	0,94	0,90	0,49	0,60	2
CT085	B12	B12	2,042	0,012	1,17	1,35	0,29	0,32	4
CT086	B12	B12	2,402	0,013	1,06	1,43	0,32	0,25	4
CT087	B15	-	1,847	0,011	1,09	1,15	0,32	0,44	4
CT088	B16	-	0,551	0,013	0,92	0,86	0,44	0,75	2
CT089	-	B19	1,845	0,016	1,10	1,30	0,31	0,29	4
CT090	-	B20	1,166	0,015	0,95	0,96	0,48	0,45	3
CT091	B08	-	1,394	0,012	0,93	0,91	0,47	0,56	3
CT092	B04	B04	-0,769	0,009	0,87	0,72	0,42	0,89	1
CT093	-	B19	-1,335	0,024	0,91	0,72	0,40	0,90	1
CT094	-	B19	0,031	0,016	1,04	1,01	0,40	0,70	2
CT095	-	B19	0,945	0,015	0,98	1,01	0,46	0,49	3
CT096	-	B19	1,119	0,015	1,01	1,05	0,43	0,45	3
CT097	B07	-	-0,356	0,017	0,95	0,81	0,34	0,88	1
CT098	B07	-	0,625	0,013	1,00	1,03	0,36	0,73	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT099	B07	-	2,210	0,012	1,03	1,07	0,37	0,36	4
CT100	B07	-	1,187	0,012	1,06	1,08	0,34	0,60	3
CT101	B07	-	0,806	0,012	0,90	0,82	0,48	0,69	2
CT102	B14	-	1,396	0,011	1,05	1,08	0,36	0,56	3
CT103	B14	-	1,177	0,012	0,95	0,92	0,45	0,61	3
CT104	B14	-	1,527	0,011	1,02	1,04	0,39	0,52	3
CT105	B14	-	1,583	0,011	1,14	1,21	0,28	0,51	3
CT106	B14	-	1,665	0,011	0,97	0,98	0,43	0,49	3
CT107	B14	-	1,942	0,012	0,93	0,94	0,47	0,42	4
CT108	B02	B02	1,359	0,006	1,06	1,08	0,39	0,51	3
CT109	B02	B02	0,270	0,007	0,96	0,91	0,44	0,75	2
CT110	B02	B02	1,421	0,006	0,98	0,99	0,46	0,50	3
CT111	B02	B02	-0,361	0,008	0,92	0,81	0,43	0,85	1
CT112	B02	B02	1,507	0,006	0,93	0,94	0,49	0,48	3
CT113	B04	B04	-0,628	0,009	1,01	1,32	0,30	0,88	1
CT114	B04	B04	-0,564	0,009	0,79	0,54	0,51	0,87	1
CT115_1	B04	B04	-0,171	0,043	0,91	0,98	0,56	1,42	1
CT115_2	B04	B04	1,187	0,043	0,91	0,98	0,56	1,42	3
CT116	B13	B13	0,691	0,014	0,92	0,90	0,49	0,65	2
CT117	B13	B13	0,140	0,015	0,99	0,96	0,42	0,76	2
CT118	B13	B13	0,573	0,014	1,20	1,33	0,27	0,68	2
CT119	B13	B13	0,151	0,015	1,01	0,94	0,41	0,76	2
CT120	B13	B13	1,116	0,013	0,99	1,01	0,45	0,56	3
CT121	B04	B04	-0,549	0,009	0,88	0,74	0,44	0,87	1
CT122	B04	B04	-0,429	0,008	0,88	0,75	0,45	0,86	1

Tabla 4.2 Dificultad y ajuste de los ítems al modelo Rasch de la EM 2022 de 2.º grado de secundaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA001	B18	B18	1,133	0,013	0,99	1,01	0,46	0,52	2
MA002	B20	B20	2,199	0,011	0,92	1,06	0,47	0,33	4
MA003	B16	B16	0,126	0,011	0,88	0,8	0,50	0,73	1
MA004	B16	B16	0,735	0,01	1,02	1,04	0,41	0,62	2
MA005	B17	B17	0,792	0,013	1,01	1,00	0,43	0,59	2
MA006	B06	B06	0,639	0,007	1,23	1,40	0,25	0,62	2
MA007	B06	B06	0,336	0,007	1,17	1,29	0,28	0,68	2
MA008	B18	B18	2,013	0,014	0,94	1,01	0,47	0,35	3
MA009	B20	B20	1,015	0,010	1,11	1,14	0,34	0,56	2
MA010	B02	-	1,426	0,008	1,22	1,28	0,23	0,52	3
MA011	B02	-	1,586	0,008	1,14	1,21	0,29	0,49	3
MA012	B02	-	0,971	0,008	0,86	0,81	0,53	0,61	2
MA013	B14	B14	-0,908	0,018	1,01	1,14	0,33	0,86	1
MA014	B18	B18	1,438	0,013	0,99	1,00	0,46	0,46	3
MA015	B14	B14	1,064	0,013	0,87	0,83	0,55	0,54	2
MA016	B15	B15	2,072	0,014	1,01	1,12	0,42	0,35	4
MA017	B02	-	1,928	0,008	0,94	0,95	0,47	0,42	3
MA018	B19	B19	1,902	0,013	1,02	1,18	0,41	0,38	3
MA019	-	B05	-0,617	0,012	0,98	0,97	0,42	0,76	1
MA020	B08	B08	0,718	0,007	0,92	0,88	0,51	0,61	2
MA021	-	B07	0,156	0,011	1,04	1,08	0,41	0,62	1
MA022	B16	B16	2,412	0,011	0,91	1,00	0,47	0,29	4
MA023	B08	B08	-	-	-	-	-	-	-
MA024	B11	B11	3,196	0,073	0,86	0,71	0,48	0,17	4

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA025	B13	B13	2,664	0,015	0,94	1,31	0,41	0,24	4
MA026	B16	B16	0,872	0,010	0,96	0,96	0,46	0,59	2
MA027	B04	-	2,464	0,009	0,92	0,95	0,47	0,31	4
MA028	B20	B20	2,009	0,011	1,00	1,09	0,42	0,36	3
MA029	B19	B19	-0,287	0,015	1,01	1,00	0,39	0,78	1
MA030	B14	B14	2,371	0,014	1,05	1,27	0,36	0,29	4
MA031	B04	-	0,327	0,009	0,99	0,97	0,39	0,73	2
MA032	B13	B13	0,837	0,013	0,91	0,87	0,51	0,59	2
MA033	B19	B19	-0,266	0,015	0,92	0,83	0,46	0,78	1
MA034	-	B07	-0,420	0,012	0,96	0,96	0,45	0,73	1
MA035	B03	-	0,282	0,013	0,93	0,88	0,44	0,73	2
MA036	B17	B17	0,775	0,013	0,97	0,99	0,46	0,60	2
MA037	B03	-	0,303	0,013	0,93	0,87	0,44	0,73	2
MA038	B03	-	0,907	0,012	1,09	1,22	0,32	0,62	2
MA039	B02	-	1,160	0,008	0,90	0,87	0,50	0,57	2
MA040	B03	-	1,486	0,011	0,90	0,89	0,50	0,50	3
MA041	B20	B20	1,655	0,010	1,09	1,16	0,36	0,43	3
MA042	B04	-	2,149	0,008	1,02	1,03	0,40	0,37	4
MA043	B13	B13	1,467	0,013	1,01	1,03	0,43	0,46	3
MA044	B02	-	3,686	0,012	1,01	1,43	0,29	0,14	4
MA045	B03	-	-	-	-	-	-	-	-
MA046	B01	-	1,766	0,012	0,99	1,01	0,42	0,44	3
MA047	B13	B13	-	-	-	-	-	-	-
MA048	B01	-	-	-	-	-	-	-	-
MA049	B16	B16	1,533	0,010	0,98	0,99	0,45	0,46	3
MA050	B04	-	0,490	0,009	0,92	0,90	0,45	0,70	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA051	B12	B12	3,359	0,081	1,04	1,03	0,31	0,14	4
MA052	B15	B15	-	-	-	-	-	-	-
MA053	B04	-	0,994	0,008	1,16	1,24	0,27	0,61	2
MA054	B21	B21	1,874	0,006	1,22	1,43	0,24	0,38	3
MA055	B21	B21	1,832	0,006	1,15	1,27	0,31	0,39	3
MA056	B21	B21	-	-	-	-	-	-	-
MA057	B22	B22	-	-	-	-	-	-	-
MA058	B22	B22	-	-	-	-	-	-	-
MA059	B01	-	1,117	0,012	0,99	0,99	0,42	0,58	2
MA060	B20	B20	1,916	0,010	1,06	1,28	0,35	0,38	3
MA061	B01	-	0,819	0,012	0,96	0,96	0,43	0,64	2
MA062	B22	B22	0,309	0,006	0,92	0,88	0,49	0,69	2
MA063_1	B10	B10	2,768	0,058	0,91	0,78	0,48	0,25	4
MA063_2	B10	B10	4,344	0,058	0,91	0,78	0,48	0,25	4
MA064	B14	B14	1,605	0,013	1,07	1,13	0,38	0,43	3
MA065	-	B05	-0,990	0,013	0,87	0,78	0,49	0,81	1
MA066_1	B09	B09	2,104	0,053	0,92	0,83	0,52	0,37	4
MA066_2	B09	B09	4,543	0,053	0,92	0,83	0,52	0,37	4
MA067	-	B07	-1,275	0,014	0,87	0,69	0,48	0,84	1
MA068	B13	B13	0,803	0,013	0,96	0,92	0,47	0,59	2
MA069	-	B05	-1,379	0,015	0,84	0,66	0,48	0,86	1
MA070	B06	B06	-0,185	0,008	0,88	0,79	0,50	0,77	1
MA071	B18	B18	0,360	0,014	0,96	0,92	0,47	0,67	2
MA072	B13	B13	-0,313	0,016	0,82	0,67	0,54	0,79	1
MA073	B16	B16	-0,101	0,012	0,99	1,00	0,39	0,77	1
MA074	B14	B14	1,149	0,013	1,05	1,06	0,41	0,52	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA075	-	B05	0,456	0,011	0,94	0,92	0,49	0,57	2
MA076	B18	B18	0,286	0,014	0,89	0,84	0,52	0,69	2
MA077	B08	B08	-1,044	0,009	0,91	0,86	0,41	0,87	1
MA078	B01	-	1,313	0,012	1,00	1,00	0,42	0,54	3
MA079	B17	B17	1,209	0,013	1,07	1,11	0,39	0,51	2
MA080	B15	B15	0,677	0,013	0,99	0,98	0,45	0,62	2
MA081	-	B05	0,984	0,011	0,97	0,97	0,46	0,46	2
MA082	B13	B13	-0,186	0,015	0,89	0,77	0,49	0,77	1
MA083	B20	B20	-1,665	0,019	0,86	0,65	0,40	0,93	1
MA084	B19	B19	0,486	0,014	0,98	0,92	0,46	0,66	2
MA085	B17	B17	1,748	0,013	0,92	0,90	0,50	0,40	3
MA086	B08	B08	-1,093	0,010	0,90	0,88	0,41	0,88	1
MA087	B06	B06	1,571	0,007	1,11	1,18	0,35	0,44	3
MA088	-	B05	1,957	0,012	0,97	1,15	0,42	0,28	3
MA089	B08	B08	0,347	0,007	0,87	0,79	0,53	0,68	2
MA090	B08	B08	-0,831	0,009	1,11	1,41	0,25	0,85	1
MA091	B15	B15	1,327	0,013	0,91	0,90	0,52	0,49	3
MA092	B15	B15	0,845	0,013	0,97	0,95	0,47	0,59	2
MA093	B18	B18	0,806	0,013	1,11	1,15	0,36	0,59	2
MA094	B06	B06	-0,475	0,008	0,87	0,74	0,49	0,81	1
MA095	B20	B20	-0,345	0,012	0,83	0,69	0,52	0,80	1
MA096	B16	B16	0,903	0,010	1,00	1,00	0,43	0,59	2
MA097	-	B07	-1,262	0,014	0,84	0,61	0,51	0,84	1
MA098	B15	B15	0,547	0,014	0,87	0,82	0,54	0,65	2
MA099	B03	-	0,046	0,013	0,90	0,87	0,44	0,77	1
MA100	B18	B18	1,071	0,013	0,91	0,89	0,52	0,54	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA101	B15	B15	-0,238	0,015	1,04	0,97	0,37	0,78	1
MA102	B17	B17	-0,715	0,017	0,90	0,83	0,45	0,84	1
MA103	B14	B14	0,525	0,014	0,91	0,89	0,50	0,64	2
MA104	B17	B17	0,884	0,013	0,94	0,90	0,49	0,57	2
MA105	B01	-	-0,549	0,016	0,92	0,86	0,38	0,85	1
MA106	B06	B06	0,442	0,007	0,91	0,85	0,50	0,66	2
MA107	B19	B19	-	-	-	-	-	-	-
MA108	B17	B17	1,449	0,013	0,93	0,93	0,49	0,46	3
MA109	-	B07	-1,036	0,013	0,95	0,90	0,43	0,82	1
MA110	B15	B15	-	-	-	-	-	-	-
MA111	B14	B14	1,466	0,013	1,02	1,04	0,43	0,46	3
MA112	B19	B19	1,919	0,013	0,98	1,10	0,44	0,37	3
MA113	B19	B19	0,644	0,013	0,96	0,95	0,47	0,63	2
MA114	B17	B17	1,161	0,013	1,13	1,18	0,34	0,52	2
MA115	B16	B16	2,093	0,011	1,14	1,30	0,30	0,35	4
MA116	B14	B14	0,583	0,014	1,05	1,06	0,40	0,63	2
MA117	-	B07	-0,518	0,012	0,91	0,80	0,50	0,74	1
MA118	B20	B20	0,580	0,011	1,03	1,03	0,40	0,65	2
MA119	B13	B13	1,690	0,013	1,22	1,47	0,24	0,41	3
MA120	B18	B18	0,940	0,013	0,91	0,88	0,52	0,56	2
MA121	B04	-	2,448	0,009	1,12	1,30	0,29	0,31	4
MA122	B19	B19	0,652	0,013	0,93	0,90	0,50	0,62	2

Tabla 4.3 Dificultad y ajuste de los ítems al modelo Rasch de la EM 2022 de Ciencia y Tecnología de 2.º grado de secundaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CN001	B03	B03	-0,078	0,013	0,96	0,94	0,44	0,51	3
CN002	B10	-	-0,142	0,012	0,96	0,95	0,40	0,56	3
CN003	B10	-	0,887	0,013	0,96	0,98	0,39	0,33	4
CN004	B02	B02	-0,051	0,013	1,18	1,25	0,15	0,50	3
CN005	B04	B04	0,737	0,014	1,00	1,04	0,37	0,33	4

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CN006	B01	-	0,163	0,012	1,04	1,04	0,32	0,49	3
CN007	B01	-	0,219	0,012	1,04	1,04	0,32	0,48	3
CN008	B04	B04	0,497	0,014	1,08	1,13	0,30	0,38	4
CN009	B14	-	1,136	0,013	1,06	1,13	0,27	0,29	4
CN010	B01	-	-1,461	0,015	0,94	0,84	0,37	0,81	1
CN011	B01	-	-2,048	0,018	0,91	0,75	0,36	0,88	1
CN012	B12	-	0,427	0,012	0,94	0,94	0,42	0,44	3
CN013	B02	B02	0,415	0,013	1,04	1,06	0,30	0,39	3
CN014	B02	B02	-2,421	0,021	0,95	0,84	0,33	0,90	1
CN015	B03	B03	-1,166	0,015	1,00	1,03	0,34	0,73	2
CN016	B03	B03	1,186	0,015	1,13	1,37	0,20	0,25	4
CN017_1	B09	B09	0,506	0,051	1,10	1,41	0,48	0,61	4
CN017_2	B09	B09	0,873	0,051	1,10	1,41	0,48	0,61	4
CN018	B03	B03	-0,205	0,013	0,99	1,00	0,40	0,54	3
CN019	B03	B03	-0,147	0,013	0,89	0,86	0,50	0,52	3
CN020	B03	B03	-0,374	0,014	1,00	1,00	0,39	0,57	3
CN021	B04	B04	0,365	0,014	0,92	0,93	0,46	0,41	3
CN022	B04	B04	-0,949	0,014	0,92	0,87	0,45	0,69	2
CN023	B02	B02	0,156	0,013	1,09	1,11	0,26	0,45	3
CN024	B09	B09	-0,077	0,009	1,04	1,04	0,36	0,51	3
CN025	B05	B05	-0,688	0,014	1,04	1,13	0,29	0,64	2
CN026	B05	B05	0,259	0,013	1,04	1,07	0,32	0,42	3
CN027	B13	-	0,738	0,013	1,07	1,10	0,31	0,37	4
CN028	-	B17	-1,895	0,018	0,94	0,86	0,39	0,80	1
CN029	B04	B04	0,633	0,014	1,16	1,24	0,22	0,35	4
CN030	B04	B04	-0,639	0,014	0,89	0,84	0,49	0,62	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CN031	B05	B05	-1,877	0,018	0,90	0,74	0,42	0,84	1
CN032	B05	B05	0,241	0,013	1,02	1,04	0,34	0,43	3
CN033	B05	B05	-1,221	0,015	0,98	0,94	0,36	0,74	2
CN034	B03	B03	-1,500	0,016	1,01	1,08	0,31	0,79	1
CN035	B05	B05	-0,409	0,013	1,11	1,15	0,24	0,57	2
CN036	B07	B07	0,597	0,008	1,10	1,17	0,23	0,37	4
CN037	B11	-	-0,730	0,013	0,93	0,89	0,42	0,69	2
CN038	B14	-	-0,493	0,013	1,06	1,11	0,28	0,64	2
CN039	B13	-	-0,660	0,013	0,96	0,97	0,39	0,68	2
CN040	B04	B04	-0,542	0,014	0,97	0,96	0,41	0,60	2
CN041	B05	B05	-1,657	0,017	0,90	0,78	0,43	0,81	1
CN042	B02	B02	-2,444	0,021	0,92	0,76	0,36	0,90	1
CN043	B02	B02	-1,348	0,015	0,99	1,01	0,33	0,76	2
CN044	B07	B07	0,181	0,007	1,08	1,12	0,26	0,46	3
CN045_1	B06	B06	1,150	0,063	0,84	0,68	0,56	0,33	4
CN045_2	B06	B06	1,441	0,063	0,84	0,68	0,56	0,33	4
CN046	-	B16	-2,157	0,019	0,89	0,75	0,43	0,84	1
CN047	-	B16	-0,012	0,015	1,19	1,25	0,16	0,43	3
CN048	B04	B04	-0,638	0,014	0,93	0,89	0,45	0,62	2
CN049	B09	B09	-0,862	0,010	0,87	0,80	0,50	0,67	2
CN050	B05	B05	-0,14	0,013	0,95	0,94	0,42	0,51	3
CN051	B02	B02	1,016	0,015	1,02	1,07	0,29	0,27	4
CN052	B03	B03	-1,136	0,015	0,84	0,74	0,52	0,73	2
CN053	B12	-	1,210	0,014	1,02	1,05	0,31	0,27	4
CN054	B12	-	0,175	0,012	0,96	0,96	0,40	0,49	3
CN055	B02	B02	-0,689	0,014	0,93	0,90	0,44	0,64	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CN056	B12	-	0,091	0,012	1,09	1,11	0,25	0,51	3
CN057	B06	B06	0,392	0,008	1,18	1,24	0,19	0,41	3
CN058	B06	B06	-0,028	0,008	1,03	1,04	0,35	0,50	3
CN059	B08	B08	0,973	0,009	1,00	1,05	0,33	0,29	4
CN060	B08	B08	0,663	0,009	0,96	0,98	0,40	0,35	4
CN061	B09	B09	0,600	0,010	1,09	1,15	0,29	0,36	4
CN062	B09	B09	-0,143	0,009	0,94	0,93	0,45	0,52	3
CN063	B11	-	1,190	0,014	1,01	1,06	0,30	0,27	4
CN064	B08	B08	-1,156	0,009	0,97	0,93	0,37	0,74	2
CN065	B08	B08	0,001	0,008	0,88	0,86	0,50	0,50	3
CN066_1	B08	B08	2,270	0,120	1,00	0,98	0,25	0,08	4
CN066_2	B08	B08	2,864	0,120	1,00	0,98	0,25	0,08	4
CN067	B06	B06	0,237	0,008	0,90	0,90	0,48	0,44	3
CN068	B06	B06	-0,784	0,009	0,94	0,91	0,43	0,66	2
CN069	B06	B06	0,067	0,008	0,94	0,94	0,44	0,48	3
CN070	B14	-	-0,201	0,012	0,94	0,92	0,42	0,57	3
CN071	B14	-	0,314	0,012	0,99	0,99	0,38	0,46	3
CN072	-	B17	-0,136	0,015	1,09	1,11	0,29	0,45	3
CN073	-	B17	-0,876	0,015	0,87	0,82	0,50	0,62	2
CN074	B11	-	1,630	0,015	1,16	1,41	0,07	0,20	4
CN075	B11	-	0,650	0,012	1,00	1,03	0,33	0,38	4
CN076	B07	B07	-0,538	0,008	1,02	1,01	0,33	0,62	2
CN077_1	B07	B07	0,403	0,060	0,89	0,86	0,52	0,46	3
CN077_2	B07	B07	2,487	0,060	0,89	0,86	0,52	0,46	4
CN078	-	B17	-1,244	0,016	1,06	1,12	0,29	0,69	2
CN079	B12	-	-0,395	0,013	1,09	1,13	0,23	0,62	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CN080	-	B19	-0,635	0,015	1,05	1,04	0,32	0,56	2
CN081	-	B19	-2,041	0,019	0,91	0,81	0,41	0,82	1
CN082	-	B19	0,238	0,015	0,94	0,95	0,42	0,37	3
CN083	-	B19	-0,998	0,015	0,91	0,9	0,45	0,64	2
CN084	-	B19	-1,416	0,016	0,93	0,86	0,42	0,72	1
CN085	B15	-	-0,144	0,012	1,01	1,01	0,33	0,56	3
CN086	B15	-	0,200	0,012	1,11	1,14	0,22	0,48	3
CN087	B15	-	0,93	0,013	0,92	0,92	0,43	0,32	4
CN088	B15	-	1,257	0,014	1,15	1,27	0,14	0,26	4
CN089	B15	-	1,229	0,014	0,96	0,99	0,36	0,26	4
CN090	B08	B08	1,298	0,010	1,10	1,28	0,19	0,23	4
CN091	B08	B08	-0,157	0,008	1,13	1,17	0,23	0,53	3
CN092	B13	-	1,685	0,015	1,04	1,15	0,27	0,2	4
CN093	B13	-	-0,127	0,012	0,92	0,89	0,47	0,56	3
CN094	B10	-	0,082	0,012	0,97	0,96	0,40	0,51	3
CN095	B10	-	0,234	0,012	1,20	1,26	0,13	0,47	3
CN096	B10	-	-0,097	0,012	0,93	0,93	0,43	0,55	3
CN097	B06	B06	0,826	0,009	1,17	1,26	0,19	0,32	4
CN098	B14	-	-0,505	0,013	0,96	0,95	0,40	0,64	2
CN099	B09	B09	0,068	0,009	0,90	0,88	0,49	0,47	3
CN100	B09	B09	-0,28	0,009	0,97	0,97	0,42	0,55	3
CN101	B11	-	-0,475	0,013	0,90	0,86	0,46	0,63	2
CN102	B13	-	1,200	0,014	1,04	1,10	0,31	0,28	4
CN103	B07	B07	-1,027	0,008	0,92	0,91	0,42	0,72	2
CN104	B01	-	1,407	0,014	1,15	1,30	0,13	0,23	4
CN105	-	B17	-1,02	0,015	0,99	0,98	0,38	0,65	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CN106	B01	-	0,092	0,012	0,92	0,90	0,45	0,50	3
CN107	B07	B07	-0,462	0,008	0,94	0,91	0,43	0,60	2
CN108	-	B18	0,070	0,015	1,16	1,21	0,21	0,41	3
CN109	-	B18	-1,022	0,015	1,09	1,16	0,26	0,64	2
CN110	-	B18	-1,719	0,017	0,88	0,76	0,46	0,77	1
CN111	-	B18	-0,681	0,015	0,88	0,84	0,49	0,57	2
CN112	-	B18	-0,494	0,015	0,94	0,92	0,44	0,53	2
CN113	-	B16	-1,156	0,015	0,92	0,87	0,44	0,68	2
CN114	-	B16	-1,349	0,016	0,88	0,80	0,48	0,71	2
CN115	B01	-	-0,925	0,013	0,87	0,79	0,48	0,72	2
CN116	B01	-	-1,191	0,014	0,96	0,92	0,36	0,77	2
CN117	-	B16	-0,890	0,015	1,01	1,01	0,35	0,62	2
CN118	B07	B07	-2,113	0,011	0,88	0,68	0,42	0,88	1

Tabla 4.4 Dificultad y ajuste de los ítems al modelo Rasch, Lectura de 2.º grado de primaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT01	B05	B05	-0,644	0,009	0,88	0,61	0,41	0,90	2
CT02	B05	B05	-0,238	0,008	0,84	0,64	0,46	0,86	2
CT03	B05	B05	1,211	0,006	0,96	0,94	0,44	0,67	3
CT04	B18	B18	0,593	0,007	0,88	0,76	0,48	0,76	3
CT05	B18	B18	2,146	0,006	1,10	1,14	0,35	0,50	4
CT06	B07	B07	1,385	0,006	0,98	1,02	0,43	0,64	3
CT07	B07	B07	1,717	0,006	0,99	1,01	0,43	0,58	3
CT08	B18	B18	2,141	0,006	0,95	0,95	0,48	0,51	4
CT09	B07	B07	1,252	0,006	0,86	0,78	0,53	0,66	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT10	B18	B18	0,423	0,007	0,91	0,76	0,45	0,78	3
CT11	-	B26	0,540	0,016	0,88	0,78	0,48	0,68	3
CT12	-	B26	1,992	0,015	1,02	1,04	0,38	0,42	3
CT13	-	B26	1,688	0,015	1,00	1,00	0,40	0,48	3
CT14	B06	B06	0,759	0,007	0,94	0,90	0,43	0,74	3
CT15	B06	B06	2,732	0,006	1,14	1,21	0,32	0,40	4
CT16	B06	B06	2,118	0,006	0,95	0,92	0,48	0,51	4
CT17	B06	B06	1,839	0,006	1,07	1,12	0,36	0,56	3
CT18	-	B12	0,752	0,016	0,88	0,81	0,50	0,65	3
CT19	-	B12	1,200	0,015	0,84	0,79	0,54	0,57	3
CT20	-	B12	1,711	0,015	1,01	1,01	0,39	0,47	3
CT21	B20	B20	1,819	0,006	1,08	1,12	0,36	0,56	3
CT22	B20	B20	0,589	0,007	0,83	0,70	0,52	0,76	3
CT23	B20	B20	2,217	0,006	1,17	1,23	0,30	0,49	4
CT24	B20	B20	2,165	0,006	1,00	1,01	0,43	0,50	4
CT25	-	B13	1,265	0,015	0,90	0,86	0,49	0,56	3
CT26	-	B13	1,417	0,015	0,88	0,83	0,51	0,53	3
CT27	-	B13	0,871	0,016	0,85	0,78	0,52	0,63	3
CT28	-	B13	1,785	0,015	1,03	1,04	0,38	0,46	3
CT29	B05	B05	-0,487	0,009	0,85	0,59	0,44	0,88	2
CT30	B21	B21	2,376	0,008	1,04	1,06	0,40	0,46	4
CT31	B06	B06	1,166	0,006	0,97	1,03	0,42	0,68	3
CT32	B07	B07	0,202	0,007	0,85	0,70	0,48	0,81	3
CT33	B08	B08	2,566	0,008	1,09	1,13	0,36	0,43	4
CT34	B08	B08	2,971	0,009	1,09	1,15	0,36	0,36	4
CT35	B08	B08	2,070	0,008	0,99	1,00	0,44	0,52	4

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT36	B08	B08	2,449	0,008	1,22	1,29	0,26	0,45	4
CT37	B09	B09	2,839	0,009	1,16	1,25	0,30	0,39	4
CT38	B09	B09	1,999	0,008	1,03	1,03	0,41	0,54	3
CT39	B09	B09	2,095	0,008	0,96	0,93	0,47	0,52	4
CT40	B06	B06	1,980	0,006	0,95	0,93	0,48	0,54	3
CT41	B08	B08	1,721	0,008	0,98	0,95	0,44	0,58	3
CT42	B22	B22	2,820	0,009	1,26	1,41	0,21	0,39	4
CT43	B22	B22	2,360	0,008	0,99	1,00	0,44	0,47	4
CT44	B22	B22	2,601	0,008	1,22	1,30	0,26	0,42	4
CT45	B22	B22	2,205	0,008	1,00	0,99	0,44	0,50	4
CT46	B22	B22	1,816	0,008	1,12	1,17	0,32	0,57	3
CT47	B11	-	1,904	0,010	0,89	0,82	0,52	0,60	3
CT48	B11	-	2,217	0,010	0,87	0,82	0,54	0,54	4
CT49	B11	-	2,445	0,010	0,96	0,94	0,46	0,50	4
CT50	B11	-	1,627	0,010	0,88	0,79	0,51	0,65	3
CT51	B11	-	1,985	0,010	0,84	0,78	0,56	0,58	3
CT52	B09	B09	2,357	0,008	1,07	1,08	0,38	0,47	4
CT53	B09	B09	2,546	0,009	1,10	1,12	0,36	0,44	4
CT54	B10	-	3,288	0,010	1,04	1,12	0,38	0,35	4
CT55	B10	-	2,679	0,010	0,97	0,96	0,46	0,46	4
CT56	B10	-	2,321	0,010	0,82	0,78	0,58	0,52	4
CT57	B24	-	2,430	0,010	0,87	0,84	0,54	0,50	4
CT58	B24	-	3,434	0,011	1,03	1,07	0,39	0,33	4
CT59	B24	-	3,625	0,011	1,23	1,41	0,20	0,30	4
CT60	B24	-	3,842	0,011	1,10	1,29	0,30	0,27	4
CT61	B24	-	3,317	0,010	1,16	1,29	0,27	0,35	4

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT62	B23	-	2,887	0,010	0,98	1,00	0,44	0,42	4
CT63	B23	-	3,537	0,011	0,94	1,01	0,46	0,31	4
CT64	B23	-	3,266	0,010	1,13	1,21	0,31	0,35	4
CT65	B23	-	3,846	0,011	1,15	1,37	0,26	0,26	4
CT66	-	B01	-1,056	0,023	0,97	0,85	0,31	0,88	2
CT67	B16	B16	0,141	0,010	0,92	0,88	0,41	0,82	3
CT68	B17	B17	-0,108	0,011	0,96	0,90	0,36	0,85	2
CT69	B10	-	1,882	0,010	0,88	0,81	0,52	0,60	3
CT70	-	B26	-	-	-	-	-	-	-
CT71	-	B25	1,721	0,015	0,99	0,99	0,40	0,47	3
CT72	-	B25	1,690	0,015	1,12	1,15	0,28	0,47	3
CT73	-	B25	2,173	0,015	1,14	1,20	0,26	0,38	4
CT74	B08	B08	3,427	0,009	1,10	1,20	0,34	0,29	4
CT75	B21	B21	3,008	0,009	1,24	1,41	0,23	0,35	4
CT76	B21	B21	2,555	0,008	1,05	1,10	0,40	0,43	4
CT77	B21	B21	2,179	0,008	0,95	0,94	0,48	0,50	4
CT78	B21	B21	2,580	0,008	1,07	1,11	0,38	0,43	4
CT79	B21	B21	3,056	0,009	0,88	0,87	0,52	0,35	4
CT80	B19	B19	1,703	0,006	0,9	0,84	0,51	0,58	3
CT81	B19	B19	1,870	0,006	0,94	0,92	0,48	0,56	3
CT82	B19	B19	2,631	0,006	1,02	1,06	0,42	0,42	4
CT83	B19	B19	2,377	0,006	0,91	0,90	0,51	0,46	4
CT84	B19	B19	3,216	0,006	1,07	1,16	0,36	0,32	4
CT85	B19	B19	2,509	0,006	1,22	1,30	0,26	0,44	4
CT86	-	B02	-0,598	0,020	0,99	1,00	0,30	0,84	2
CT87	B03	B03	-0,565	0,013	0,94	0,84	0,35	0,89	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT88	-	B15	-0,261	0,018	0,88	0,74	0,45	0,80	2
CT89	B04	B04	-0,286	0,012	0,93	0,86	0,37	0,86	2
CT90	-	B14	-0,788	0,021	0,93	0,79	0,37	0,86	2

Tabla 4.5 Dificultad y ajuste de los ítems al modelo Rasch, Matemática de 2.º grado de primaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA01	B11	B11	-0,655	0,009	1,08	1,30	0,39	0,69	2
MA02	B11	B11	0,619	0,009	0,88	0,85	0,59	0,48	3
MA03	B11	B11	0,949	0,009	0,87	0,84	0,60	0,43	3
MA04	B10	B10	1,001	0,009	0,88	0,85	0,57	0,42	3
MA05	B09	B09	-0,292	0,009	1,00	1,13	0,46	0,64	2
MA06	-	B06	0,176	0,011	1,12	1,17	0,39	0,46	3
MA07	B10	B10	-0,662	0,009	0,92	0,89	0,48	0,70	2
MA08	B16	B16	0,386	0,009	0,97	0,98	0,52	0,52	3
MA09	B01	-	1,023	0,007	0,91	0,88	0,56	0,46	3
MA10	B10	B10	1,330	0,009	1,08	1,14	0,44	0,36	3
MA11	B09	B09	-0,791	0,009	0,79	0,71	0,57	0,71	2
MA12	-	B05	-0,122	0,011	1,04	1,03	0,44	0,51	2
MA13	B12	B12	-0,878	0,009	0,98	0,95	0,43	0,73	2
MA14	B12	B12	-0,347	0,009	0,96	0,92	0,48	0,64	2
MA15	B10	B10	0,449	0,009	0,94	0,94	0,52	0,51	3
MA16	B12	B12	0,161	0,009	1,00	0,97	0,48	0,56	3
MA17	-	B05	0,337	0,011	1,02	1,01	0,47	0,43	3
MA18	B01	-	0,649	0,007	1,19	1,26	0,37	0,52	3
MA19	B11	B11	0,416	0,009	1,06	1,10	0,47	0,52	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA20	B09	B09	2,545	0,011	1,10	1,45	0,40	0,21	4
MA21	B04	-	0,388	0,007	0,95	0,92	0,52	0,56	3
MA22	B13	B13	0,782	0,009	1,24	1,36	0,34	0,45	3
MA23	B02	-	-	-	-	-	-	-	-
MA24	B11	B11	-0,120	0,009	1,09	1,15	0,43	0,61	2
MA25	B10	B10	1,964	0,010	1,06	1,27	0,42	0,27	4
MA26	B04	-	-	-	-	-	-	-	-
MA27	B01	-	0,981	0,007	1,12	1,17	0,42	0,46	3
MA28	B12	B12	1,751	0,010	1,15	1,42	0,37	0,30	3
MA29	B09	B09	-	-	-	-	-	-	-
MA30	B16	B16	2,070	0,010	1,05	1,19	0,46	0,26	4
MA31	B12	B12	1,640	0,009	1,18	1,31	0,37	0,31	3
MA32	B11	B11	-0,115	0,009	0,91	0,90	0,54	0,61	2
MA33	B12	B12	0,134	0,009	1,17	1,27	0,36	0,56	3
MA34	B09	B09	2,087	0,010	0,90	1,05	0,54	0,26	4
MA35	B10	B10	-	-	-	-	-	-	-
MA36	-	B07	0,267	0,011	0,88	0,85	0,56	0,44	3
MA37	-	B07	-	-	-	-	-	-	-
MA38	B12	B12	0,513	0,009	1,22	1,33	0,34	0,50	3
MA39	B10	B10	1,028	0,009	1,10	1,16	0,42	0,41	3
MA40	-	B08	-2,722	0,016	1,00	1,22	0,27	0,88	2
MA41	B10	B10	-1,368	0,010	0,93	0,90	0,42	0,80	2
MA42	B09	B09	-1,269	0,010	1,03	1,28	0,37	0,78	2
MA43	B12	B12	0,090	0,009	0,76	0,69	0,63	0,57	2
MA44	B15	B15	0,614	0,009	0,88	0,87	0,57	0,48	3
MA45	-	B08	-1,390	0,012	1,01	1,11	0,37	0,73	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA46	-	B06	-1,145	0,012	0,94	0,86	0,46	0,69	2
MA47	-	B06	-0,198	0,011	0,84	0,81	0,58	0,53	2
MA48	B11	B11	0,084	0,009	0,83	0,75	0,60	0,57	2
MA49	B11	B11	2,338	0,010	1,15	1,42	0,39	0,23	4
MA50	B09	B09	1,282	0,009	0,96	1,00	0,54	0,38	3
MA51	-	B05	-1,416	0,012	0,91	0,87	0,45	0,73	2
MA52	B09	B09	-1,168	0,010	0,94	1,01	0,44	0,77	2
MA53	B04	-	2,596	0,009	1,00	1,10	0,46	0,22	4
MA54	-	B07	0,050	0,011	0,97	0,96	0,49	0,48	2
MA55	-	B05	-0,640	0,011	1,08	1,09	0,38	0,60	2
MA56	B14	B14	2,026	0,010	0,98	0,98	0,51	0,27	4
MA57	B02	-	0,837	0,007	0,86	0,81	0,59	0,49	3
MA58	B03	-	1,932	0,008	1,01	1,02	0,49	0,31	4
MA59	B02	-	0,862	0,007	1,21	1,31	0,36	0,48	3
MA60	B02	-	-1,358	0,009	0,98	0,99	0,38	0,82	2
MA61	B02	-	0,792	0,007	1,06	1,10	0,46	0,50	3
MA62	B13	B13	0,750	0,009	0,97	0,98	0,51	0,46	3
MA63	B04	-	0,385	0,007	0,87	0,83	0,57	0,56	3
MA64	B15	B15	-0,168	0,009	0,83	0,76	0,58	0,61	2
MA65	B14	B14	-0,582	0,009	1,00	1,06	0,44	0,69	2
MA66	B04	-	0,510	0,007	0,95	0,92	0,52	0,54	3
MA67	-	B06	-2,204	0,014	0,96	1,02	0,35	0,84	2
MA68	-	B07	-2,752	0,016	0,99	1,06	0,28	0,89	2
MA69	B03	-	0,775	0,007	0,91	0,89	0,56	0,50	3
MA70	B01	-	1,718	0,008	1,01	1,03	0,49	0,34	3
MA71	B03	-	1,507	0,008	0,96	1,03	0,52	0,38	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA72	B13	B13	0,501	0,009	1,01	1,02	0,48	0,50	3
MA73	-	B07	-0,892	0,011	1,10	1,12	0,36	0,65	2
MA74	-	B08	0,447	0,011	0,97	0,96	0,50	0,41	3
MA75	-	B06	0,005	0,011	0,96	0,95	0,50	0,49	2
MA76	B16	B16	-0,096	0,009	0,94	0,96	0,52	0,61	2
MA77	-	B06	0,069	0,011	0,83	0,78	0,59	0,48	2
MA78	B03	-	1,420	0,008	1,04	1,16	0,47	0,39	3
MA79	B03	-	0,883	0,007	1,01	1,04	0,49	0,48	3
MA80	B01	-	1,078	0,007	1,27	1,41	0,32	0,45	3
MA81	-	B07	-1,644	0,012	0,96	0,97	0,40	0,76	2
MA82	-	B05	-0,169	0,011	0,76	0,70	0,63	0,52	2
MA83	-	B08	-0,556	0,011	0,92	0,89	0,50	0,59	2
MA84	-	B05	-1,158	0,012	0,95	0,90	0,45	0,69	2
MA85	B14	B14	0,204	0,009	0,88	0,83	0,57	0,56	3
MA86	B02	-	1,876	0,008	1,10	1,23	0,42	0,32	4
MA87	B01	-	1,748	0,008	0,98	1,02	0,51	0,34	3
MA88	B03	-	-0,625	0,008	0,79	0,69	0,56	0,72	2
MA89	-	B08	-0,155	0,011	1,16	1,29	0,34	0,52	2
MA90	B15	B15	0,486	0,009	1,00	1,01	0,49	0,50	3
MA91	-	B08	-1,532	0,012	0,90	0,90	0,45	0,75	2
MA92	B04	-	-0,651	0,008	0,95	1,03	0,45	0,73	2

Tabla 4.6 Dificultad y ajuste de los ítems al modelo Rasch, Lectura de 4.º grado de primaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT01	B08	B08	1,123	0,012	1,07	1,25	0,37	0,28	4
CT02	B08	B08	1,095	0,012	1,18	1,44	0,28	0,29	4
CT03	B08	B08	0,962	0,012	1,00	1,11	0,44	0,31	4
CT04	B08	B08	-0,258	0,011	0,88	0,83	0,55	0,55	3
CT05	B08	B08	-1,260	0,012	0,94	0,90	0,44	0,74	2
CT06	B16	B16	-1,465	0,011	0,96	0,94	0,44	0,75	2
CT07	B16	B16	-0,046	0,010	1,10	1,15	0,40	0,48	3
CT08	B16	B16	-0,898	0,010	0,88	0,80	0,54	0,65	3
CT09	B16	B16	-0,573	0,010	1,08	1,13	0,40	0,59	3
CT10	B11	B11	-0,835	0,012	0,96	0,93	0,49	0,65	3
CT11	B16	B16	-1,363	0,011	1,01	1,14	0,39	0,74	2
CT12	B10	B10	-0,406	0,010	1,10	1,12	0,39	0,56	3
CT13	B10	B10	-0,270	0,010	1,03	1,02	0,45	0,53	3
CT14	B10	B10	-1,071	0,010	0,98	0,95	0,45	0,68	2
CT15	B10	B10	-0,467	0,010	1,08	1,08	0,40	0,57	3
CT16	B16	B16	-0,710	0,010	0,93	0,89	0,51	0,62	3
CT17	B10	B10	-1,603	0,011	0,96	0,89	0,43	0,77	2
CT18	B10	B10	0,160	0,010	1,09	1,13	0,40	0,44	4
CT19	B04	B04	-0,455	0,006	1,05	1,04	0,43	0,59	3
CT20	-	B15	-2,652	0,019	0,87	0,71	0,45	0,84	1
CT21	-	B15	-2,557	0,019	0,82	0,62	0,50	0,83	1
CT22	-	B15	-2,043	0,017	0,85	0,73	0,52	0,76	2
CT23	-	B15	-2,441	0,018	0,83	0,61	0,51	0,81	1
CT24	-	B15	-2,598	0,019	0,88	0,76	0,45	0,84	1

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT25	-	B09	-1,742	0,016	0,92	0,87	0,49	0,71	2
CT26	-	B09	-1,323	0,015	0,87	0,82	0,55	0,63	2
CT27	-	B09	-0,425	0,015	1,08	1,12	0,42	0,45	3
CT28	-	B09	-1,270	0,015	0,88	0,80	0,54	0,62	2
CT29	-	B09	-0,238	0,015	1,12	1,18	0,39	0,42	3
CT30	-	B09	-1,377	0,015	0,90	0,85	0,52	0,64	2
CT31	B12	-	0,275	0,009	1,04	1,06	0,41	0,49	4
CT32	B12	-	0,180	0,009	0,97	0,96	0,47	0,51	4
CT33	B12	-	-0,079	0,009	0,88	0,83	0,54	0,56	3
CT34	B12	-	1,107	0,010	1,15	1,39	0,29	0,32	4
CT35	B12	-	0,162	0,009	1,05	1,09	0,40	0,51	4
CT36	B03	B03	-1,060	0,007	0,88	0,78	0,53	0,70	2
CT37	B02	B02	-1,161	0,007	0,96	0,98	0,45	0,72	2
CT38	B12	-	-1,157	0,010	0,83	0,70	0,53	0,76	2
CT39	B11	B11	0,091	0,011	1,02	1,14	0,45	0,47	3
CT40	B11	B11	-1,091	0,012	1,05	1,1	0,42	0,70	2
CT41	B11	B11	-1,390	0,013	0,87	0,75	0,53	0,75	2
CT42	B11	B11	-0,900	0,012	0,98	0,94	0,48	0,66	3
CT43	B11	B11	-0,456	0,011	0,96	0,95	0,50	0,58	3
CT44	B14	B14	-1,392	0,013	0,81	0,65	0,54	0,76	2
CT45	B14	B14	-0,813	0,011	0,88	0,82	0,52	0,66	3
CT46	B14	B14	-0,278	0,011	0,93	0,91	0,51	0,55	3
CT47	B14	B14	-0,592	0,011	0,81	0,73	0,59	0,62	3
CT48	B14	B14	0,446	0,011	1,03	1,06	0,44	0,41	4
CT49	B14	B14	-0,059	0,011	1,25	1,35	0,27	0,51	3
CT50	-	B15	-2,380	0,018	0,85	0,69	0,50	0,81	1

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT51	B05	B05	-1,843	0,014	0,88	0,82	0,48	0,81	2
CT52	B05	B05	-1,212	0,012	0,94	0,87	0,49	0,72	2
CT53	B05	B05	-2,653	0,017	0,86	0,69	0,43	0,90	1
CT54	B05	B05	-2,762	0,018	0,92	1,00	0,37	0,91	1
CT55	B01	B01	-0,904	0,007	0,93	0,89	0,49	0,68	3
CT56	B01	B01	-1,208	0,007	0,93	0,83	0,48	0,73	2
CT57	B01	B01	-0,435	0,006	0,94	0,93	0,50	0,59	3
CT58	B01	B01	0,384	0,081	0,86	0,80	0,58	0,43	4
CT59	B05	B05	0,394	0,011	1,09	1,24	0,39	0,41	4
CT60	B05	B05	-2,364	0,016	0,84	0,68	0,47	0,87	1
CT61	B01	B01	0,450	0,006	1,19	1,29	0,32	0,41	4
CT62	B01	B01	-0,894	0,007	1,01	0,98	0,44	0,67	3
CT63	B06	-	0,447	0,009	0,94	0,94	0,50	0,46	4
CT64	B06	-	0,978	0,009	1,11	1,24	0,33	0,35	4
CT65	B06	-	0,349	0,009	1,13	1,17	0,34	0,48	4
CT66	B06	-	-0,336	0,009	1,01	0,99	0,43	0,62	3
CT67	B06	-	0,433	0,009	1,03	1,07	0,42	0,46	4
CT68	B06	-	0,467	0,009	1,04	1,08	0,41	0,45	4
CT69	B07	-	-0,200	0,013	1,01	0,97	0,43	0,60	3
CT70	B07	-	1,341	0,014	1,08	1,32	0,33	0,29	4
CT71	B07	-	0,241	0,013	1,00	1,02	0,44	0,50	4
CT72	B13	-	2,034	0,016	1,13	1,43	0,26	0,18	4
CT73	B13	-	0,666	0,013	1,07	1,14	0,39	0,41	4
CT74	B13	-	-0,700	0,014	1,00	0,96	0,42	0,69	3
CT75	B13	-	0,258	0,013	0,86	0,84	0,56	0,49	4
CT76	B13	-	1,010	0,014	1,19	1,38	0,27	0,34	4

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT77	B04	B04	2,613	0,128	0,90	0,64	0,41	0,10	4
CT78	B08	B08	0,370	0,011	1,01	1,04	0,45	0,42	4
CT79	B13	-	0,723	0,013	0,99	1,05	0,45	0,40	4
CT80	B07	-	0,148	0,013	0,99	0,98	0,45	0,52	4
CT81	B07	-	0,663	0,013	1,10	1,18	0,35	0,42	4
CT82	B07	-	0,079	0,013	1,07	1,09	0,38	0,54	3
CT83	B03	B03	-1,492	0,007	0,93	0,98	0,44	0,77	2
CT84	B03	B03	-0,396	0,006	0,87	0,83	0,56	0,58	3
CT85_1	B03	B03	-0,529	0,059	0,87	0,86	0,65	0,84	3
CT85_2	B03	B03	1,408	0,059	0,87	0,86	0,65	0,84	4
CT86	B02	B02	0,406	0,006	1,10	1,20	0,38	0,42	4
CT87	B02	B02	-1,089	0,007	0,86	0,79	0,53	0,71	2
CT88	B02	B02	0,522	0,007	1,18	1,29	0,33	0,39	4
CT89	B02	B02	1,092	0,087	0,98	0,94	0,46	0,29	4
CT90	B02	B02	-1,280	0,007	0,94	0,96	0,45	0,74	2
CT91	B04	B04	0,586	0,007	1,02	1,08	0,44	0,38	4
CT92	B04	B04	0,258	0,006	1,07	1,15	0,41	0,44	4
CT93	B04	B04	-0,334	0,006	1,16	1,30	0,34	0,56	3
CT94	B04	B04	-0,159	0,006	1,06	1,09	0,42	0,53	3
CT95	B03	B03	-0,993	0,007	0,84	0,75	0,56	0,69	2
CT96	B03	B03	-0,601	0,007	0,95	0,92	0,50	0,62	3

Tabla 4.7 Dificultad y ajuste de los ítems al modelo Rasch, Matemática de 4.º grado de primaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA001	-	B07	-1,559	0,012	1,07	1,09	0,43	0,60	2
MA002	B17	B17	-2,429	0,014	0,82	0,65	0,52	0,83	1
MA003	B01	-	-1,523	0,011	0,86	0,78	0,52	0,76	2
MA004	B18	B18	-1,774	0,013	0,93	0,85	0,50	0,75	2
MA005	B08	B08	-2,452	0,011	1,04	1,22	0,38	0,80	1
MA006	B13	B13	-1,222	0,012	0,96	0,92	0,50	0,67	2
MA007	-	B05	-1,250	0,012	1,03	1,03	0,47	0,56	2
MA008	B08	B08	-2,340	0,011	0,89	0,77	0,50	0,78	1
MA009	B17	B17	-0,945	0,012	1,03	1,01	0,48	0,63	3
MA010	B06	B06	-2,340	0,011	0,85	0,75	0,52	0,79	1
MA011	B08	B08	-2,175	0,010	0,93	0,84	0,49	0,76	1
MA012	-	B07	-2,209	0,013	0,91	0,83	0,50	0,71	1
MA013	B16	B16	-1,417	0,012	0,99	0,94	0,48	0,70	2
MA014	B14	B14	-1,573	0,012	0,88	0,78	0,54	0,73	2
MA015	B18	B18	-1,753	0,013	0,83	0,71	0,56	0,75	2
MA016	B04	-	-2,007	0,010	0,84	0,64	0,51	0,82	2
MA017	B18	B18	-1,368	0,012	1,03	1,05	0,46	0,70	2
MA018	B03	-	0,435	0,009	1,00	1,05	0,47	0,45	4
MA019	B17	B17	-2,084	0,014	0,92	0,85	0,48	0,79	1
MA020	B13	B13	-0,839	0,011	1,09	1,17	0,42	0,61	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA021	B16	B16	-0,225	0,011	1,30	1,48	0,28	0,50	3
MA022	B15	B15	-2,498	0,015	0,98	1,09	0,40	0,84	1
MA023	B15	B15	-1,902	0,013	0,95	0,98	0,47	0,77	2
MA024	B13	B13	-0,321	0,011	0,86	0,84	0,58	0,52	3
MA025	B13	B13	0,469	0,011	1,02	1,06	0,47	0,39	4
MA026	B14	B14	-0,704	0,012	1,01	1,02	0,49	0,60	3
MA027	-	B05	-1,180	0,012	1,01	1,03	0,48	0,54	2
MA028	B02	-	-2,467	0,011	1,01	1,07	0,34	0,87	1
MA029	B15	B15	1,211	0,012	1,08	1,40	0,40	0,29	4
MA030	B16	B16	-3,107	0,017	0,86	0,66	0,44	0,89	1
MA031	B17	B17	0,313	0,011	1,07	1,21	0,44	0,43	3
MA032	B18	B18	0,374	0,011	1,09	1,21	0,43	0,41	4
MA033	B16	B16	-1,736	0,013	0,89	0,82	0,52	0,74	2
MA034	B18	B18	0,065	0,011	1,09	1,24	0,43	0,46	3
MA035	B18	B18	-1,995	0,013	0,92	0,93	0,48	0,78	2
MA036	B14	B14	1,581	0,013	1,04	1,41	0,40	0,24	4
MA037	-	B05	-0,744	0,012	1,08	1,12	0,45	0,47	3
MA038	-	B07	-1,818	0,013	0,98	0,97	0,48	0,64	2
MA039	B17	B17	0,975	0,012	1,2	1,41	0,34	0,32	4
MA040	-	B05	-1,570	0,012	0,96	0,91	0,50	0,61	2
MA041	B14	B14	-1,037	0,012	0,98	0,98	0,50	0,65	3
MA042	B16	B16	-0,160	0,011	1,15	1,23	0,39	0,49	3
MA043	B04	-	1,649	0,009	1,09	1,44	0,34	0,26	4
MA044	B02	-	-0,192	0,008	0,98	0,98	0,49	0,56	3
MA045	B02	-	-	-	-	-	-	-	-
MA046	-	B07	-2,505	0,014	0,86	0,78	0,51	0,75	1

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA047	B01	-	-1,548	0,011	0,80	0,66	0,56	0,77	2
MA048	B08	B08	-0,178	0,009	0,89	0,90	0,58	0,46	3
MA049	B06	B06	-0,420	0,009	1,08	1,12	0,46	0,50	3
MA050	B02	-	0,360	0,008	1,10	1,17	0,40	0,46	4
MA051	B03	-	-0,355	0,009	1,25	1,30	0,30	0,58	3
MA052	B14	B14	-0,967	0,012	0,97	0,96	0,51	0,64	3
MA053	B01	-	0,739	0,009	1,21	1,42	0,30	0,40	4
MA054	B02	-	-	-	-	-	-	-	-
MA055	B01	-	0,313	0,009	1,02	1,05	0,46	0,47	3
MA056	B06	B06	-	-	-	-	-	-	-
MA057	B02	-	1,603	0,008	1,19	1,41	0,29	0,26	4
MA058	B03	-	-0,398	0,009	1,03	1,03	0,45	0,59	3
MA059	B08	B08	-0,393	0,009	1,00	1,01	0,51	0,49	3
MA060	B06	B06	-2,230	0,011	0,98	0,98	0,45	0,77	1
MA061	B04	-	-0,780	0,008	1,08	1,09	0,41	0,65	3
MA062	B16	B16	0,724	0,012	1,09	1,25	0,41	0,35	4
MA063	-	B05	-3,621	0,017	0,99	1,37	0,30	0,88	1
MA064	B04	-	-0,039	0,008	0,98	0,99	0,49	0,53	3
MA065	-	B05	-0,839	0,012	1,25	1,35	0,33	0,49	3
MA066	B08	B08	-0,776	0,009	1,05	1,07	0,48	0,56	3
MA067	B04	-	1,035	0,008	1,07	1,26	0,39	0,35	4
MA068	B04	-	-	-	-	-	-	-	-
MA069	B01	-	-0,661	0,010	0,96	0,93	0,50	0,64	3
MA070	B01	-	1,058	0,010	1,14	1,34	0,34	0,34	4
MA071	-	B07	-1,697	0,013	0,82	0,74	0,59	0,63	2
MA072_1	B10	B10	0,126	0,062	0,98	1,00	0,59	0,60	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA072_2	B10	B10	1,975	0,062	0,98	1,00	0,59	0,60	4
MA073	B01	-	-1,595	0,011	0,95	0,88	0,46	0,77	2
MA074	B03	-	0,483	0,009	1,03	1,06	0,45	0,44	4
MA075	B03	-	-0,642	0,010	0,84	0,77	0,58	0,63	3
MA076	B06	B06	-3,178	0,013	0,90	0,76	0,43	0,87	1
MA077	B06	B06	-3,049	0,012	0,99	1,04	0,37	0,86	1
MA078	B15	B15	0,366	0,011	1,14	1,33	0,40	0,42	4
MA079	B02	-	-0,851	0,008	1,01	1,04	0,44	0,67	3
MA080	B06	B06	-0,329	0,009	1,07	1,12	0,47	0,48	3
MA081	B03	-	-1,606	0,011	0,87	0,76	0,51	0,77	2
MA082	B19	B19	-0,476	0,007	0,94	0,92	0,54	0,56	3
MA083	B19	B19	-0,101	0,007	1,08	1,13	0,45	0,49	3
MA084	B19	B19	-0,451	0,007	0,89	0,85	0,57	0,55	3
MA085	B20	B20	-0,714	0,007	0,97	0,96	0,52	0,59	3
MA086	B20	B20	-1,297	0,007	0,88	0,83	0,55	0,68	2
MA087	B20	B20	-1,413	0,007	1,04	1,10	0,44	0,70	2
MA088	B03	-	-0,639	0,010	1,00	0,95	0,47	0,63	3
MA089	B04	-	-2,524	0,011	0,88	0,75	0,44	0,87	1
MA090_1	B11	B11	0,106	0,058	0,94	0,86	0,64	0,67	3
MA090_2	B11	B11	1,442	0,058	0,94	0,86	0,64	0,67	4
MA091_1	B12	B12	-1,653	0,058	1,05	1,30	0,59	1,44	2
MA091_2	B12	B12	-0,937	0,058	1,05	1,30	0,59	1,44	3
MA092	B09	B09	1,824	0,102	0,98	0,90	0,43	0,21	4
MA093	B15	B15	-0,639	0,011	0,86	0,79	0,59	0,58	3
MA094	B13	B13	-0,190	0,011	0,88	0,87	0,57	0,50	3
MA095	B15	B15	-2,918	0,016	0,95	0,93	0,40	0,87	1

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA096	B14	B14	-1,769	0,013	0,94	0,84	0,50	0,75	2
MA097	-	B07	-0,717	0,012	1,06	1,14	0,46	0,46	3
MA098	B13	B13	-2,149	0,013	1,05	1,13	0,38	0,80	1
MA099	B08	B08	-1,870	0,010	0,87	0,82	0,54	0,72	2
MA100	-	B07	-1,165	0,012	0,92	0,87	0,54	0,54	2
MA101	B17	B17	-1,155	0,012	0,82	0,74	0,60	0,66	2
MA102	-	B05	-1,915	0,013	1,00	1,06	0,45	0,66	2

Tabla 4.8 Dificultad y ajuste de los ítems al modelo Rasch, Lectura de 6.º grado de primaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT001	B17	B17	-1,316	0,073	0,91	0,95	0,43	0,88	1
CT002	B17	B17	0,361	0,050	1,05	1,11	0,47	0,60	3
CT003	B17	B17	-0,819	0,062	0,82	0,65	0,54	0,82	2
CT004	B17	B17	0,190	0,051	1,14	1,21	0,42	0,63	3
CT005	B17	B17	-0,874	0,063	1,07	1,32	0,37	0,83	2
CT006	B17	B17	-0,795	0,062	0,94	0,82	0,48	0,82	2
CT007	B08	B08	0,800	0,056	0,88	0,91	0,58	0,48	3
CT008	B09	B09	-0,972	0,072	0,95	0,81	0,45	0,84	1
CT009	B08	B08	0,253	0,057	1,13	1,17	0,42	0,61	3
CT010	B02	B02	-1,684	0,083	0,86	0,79	0,39	0,94	1
CT011	B02	B02	0,380	0,044	0,97	0,96	0,50	0,64	3
CT012	B02	B02	-1,060	0,064	1,01	0,88	0,37	0,89	1
CT013	B02	B02	-0,849	0,059	0,88	0,83	0,46	0,86	2
CT014	B02	B02	1,108	0,043	1,14	1,28	0,40	0,47	4
CT015	B02	B02	1,916	0,074	0,85	0,72	0,58	0,29	4

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT016	B05	B05	-0,445	0,050	1,09	1,35	0,37	0,78	2
CT017	B05	B05	-0,360	0,049	0,98	0,94	0,46	0,76	2
CT018	B05	B05	1,033	0,043	1,01	1,17	0,49	0,45	4
CT019	B05	B05	2,567	0,094	1,02	0,76	0,41	0,14	4
CT020	B06	B06	0,358	0,044	1,17	1,29	0,37	0,65	3
CT021	B06	B06	-0,318	0,050	0,88	0,82	0,51	0,78	2
CT022	B06	B06	1,343	0,043	1,06	1,21	0,44	0,41	4
CT023	B07	B07	-0,087	0,061	0,77	0,69	0,60	0,72	2
CT024	B07	B07	0,446	0,057	1,10	1,18	0,43	0,60	3
CT025	B07	B07	0,527	0,056	1,12	1,14	0,42	0,58	3
CT026	B07	B07	0,259	0,058	0,93	0,94	0,53	0,65	3
CT027	B13	-	-0,310	0,078	1,02	1,04	0,39	0,79	2
CT028	B13	-	-	-	-	-	-	-	-
CT029	B13	-	1,241	0,066	1,11	1,17	0,42	0,44	4
CT030	B13	-	-	-	-	-	-	-	-
CT031	B13	-	1,083	0,066	1,23	1,28	0,34	0,48	4
CT032	B05	B05	0,203	0,044	1,01	1,07	0,47	0,65	3
CT033	B05	B05	0,800	0,043	1,21	1,37	0,36	0,50	3
CT034	B06	B06	0,607	0,043	0,98	0,97	0,50	0,59	3
CT035	B06	B06	-	-	-	-	-	-	-
CT036	B01	B01	-1,395	0,068	0,88	0,68	0,43	0,90	1
CT037	B15	-	0,385	0,050	1,02	0,99	0,46	0,68	3
CT038	B15	-	-0,064	0,054	0,82	0,68	0,55	0,77	2
CT039	B15	-	-0,477	0,061	0,90	0,75	0,47	0,84	2
CT040	B03	B03	-0,689	0,071	1,01	0,91	0,42	0,82	2
CT041	B04	B04	-0,769	0,069	0,91	0,85	0,48	0,81	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT042	B04	B04	-1,345	0,084	0,95	0,82	0,41	0,89	1
CT043	B04	B04	-0,474	0,064	0,90	0,88	0,51	0,76	2
CT044	B04	B04	-0,131	0,089	0,90	0,84	0,55	0,66	2
CT045	B04	B04	0,437	0,056	1,14	1,19	0,42	0,57	3
CT046	B04	B04	-0,93	0,073	0,94	0,79	0,46	0,84	1
CT047	-	B18	-1,736	0,102	0,92	1,10	0,40	0,87	1
CT048	-	B18	0,070	0,072	1,00	1,02	0,50	0,53	2
CT049	-	B18	-0,881	0,080	0,92	0,82	0,51	0,74	2
CT050	-	B18	0,314	0,072	1,16	1,25	0,41	0,48	3
CT051	-	B18	-0,068	0,072	0,82	0,74	0,61	0,57	2
CT052	B12	-	0,298	0,049	0,92	0,88	0,49	0,68	3
CT053	B12	-	1,053	0,046	1,05	1,12	0,42	0,50	4
CT054	B12	-	0,659	0,047	1,04	1,04	0,42	0,60	3
CT055	B12	-	-0,279	0,056	0,98	0,94	0,40	0,80	2
CT056	B12	-	0,660	0,047	1,05	1,06	0,42	0,60	3
CT057	B12	-	0,747	0,046	1,00	1,02	0,46	0,58	3
CT058	B11	-	0,920	0,046	1,07	1,12	0,40	0,51	3
CT059	B11	-	0,516	0,047	0,95	0,92	0,48	0,61	3
CT060	B11	-	0,463	0,047	1,11	1,18	0,35	0,62	3
CT061	B01	B01	-1,130	0,062	1,03	1,38	0,33	0,88	1
CT062	B01	B01	-0,901	0,057	0,80	0,61	0,52	0,85	1
CT063_1	B01	B01	-1,025	0,052	0,91	0,95	0,64	1,34	1
CT063_2	B01	B01	0,789	0,052	0,91	0,95	0,64	1,34	3
CT064	B03	B03	0,559	0,056	1,08	1,09	0,45	0,58	3
CT065	B03	B03	-1,095	0,081	0,87	0,83	0,45	0,88	1
CT066	B03	B03	0,796	0,056	1,16	1,27	0,40	0,52	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT067	B03	B03	0,537	0,112	1,02	1,02	0,44	0,64	3
CT068	B03	B03	-0,934	0,077	1,02	1,23	0,37	0,86	1
CT069	B08	B08	0,396	0,056	1,02	1,00	0,49	0,58	3
CT070	B08	B08	0,144	0,058	0,90	0,85	0,55	0,64	2
CT071	B08	B08	0,702	0,056	0,87	0,83	0,59	0,51	3
CT072	B08	B08	1,921	0,100	1,11	1,26	0,39	0,22	4
CT073	B15	-	-0,915	0,071	0,87	0,65	0,43	0,89	1
CT074	B15	-	0,456	0,049	0,92	0,87	0,53	0,66	3
CT075	B15	-	-0,503	0,061	0,97	0,95	0,41	0,84	2
CT076	B09	B09	0,157	0,057	0,90	0,86	0,54	0,62	3
CT077	B09	B09	0,017	0,058	1,04	1,02	0,46	0,65	2
CT078	B09	B09	0,002	0,058	1,3	1,45	0,28	0,66	2
CT079	B09	B09	-0,086	0,059	1,00	0,92	0,48	0,67	2
CT080	B09	B09	0,852	0,055	1,03	1,19	0,45	0,45	3
CT081	B14	-	-0,982	0,070	0,91	0,73	0,40	0,89	1
CT082	B14	-	0,134	0,051	1,03	0,99	0,44	0,71	2
CT083	B14	-	-0,203	0,054	0,91	0,77	0,49	0,78	2
CT084	B14	-	0,396	0,049	1,05	1,02	0,45	0,65	3
CT085	-	B16	-0,175	0,073	1,01	0,98	0,50	0,58	2
CT086	-	B16	-1,074	0,084	0,82	0,63	0,57	0,77	1
CT087	-	B16	-1,368	0,090	0,87	0,80	0,52	0,81	1
CT088	B06	B06	2,301	0,081	0,94	0,82	0,50	0,22	4
CT089	B10	B10	-0,216	0,063	0,90	0,79	0,55	0,74	2
CT090	B10	B10	1,318	0,058	1,12	1,32	0,44	0,40	4
CT091	B10	B10	0,37	0,058	1,22	1,40	0,38	0,62	3
CT092	B10	B10	0,167	0,060	0,95	0,86	0,54	0,66	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
CT093	B10	B10	0,858	0,057	1,14	1,23	0,44	0,50	3
CT094	B13	-	0,707	0,066	0,98	0,99	0,49	0,57	3
CT095	B11	-	0,468	0,047	1,01	1,01	0,43	0,62	3
CT096	B07	B07	-0,452	0,066	0,88	0,77	0,52	0,79	2
CT097	-	B16	-0,183	0,073	0,96	0,90	0,54	0,58	2
CT098	-	B16	-0,126	0,073	0,86	0,83	0,59	0,57	2
CT099	-	B16	0,370	0,073	1,11	1,13	0,44	0,45	3
CT100	B01	B01	-1,263	0,065	0,90	0,72	0,43	0,89	1
CT101	B14	-	0,194	0,050	1,02	1,02	0,45	0,70	3
CT102	B14	-	0,314	0,049	1,01	0,98	0,46	0,67	3
CT103	B11	-	1,275	0,046	1,09	1,19	0,39	0,42	4
CT104	B11	-	0,332	0,048	0,95	0,93	0,47	0,66	3
CT105	-	B18	-0,931	0,081	0,80	0,64	0,57	0,75	1
CT106	B07	B07	2,436	0,112	0,97	0,67	0,49	0,17	4
CT107	B10	B10	-0,749	0,072	0,87	0,67	0,52	0,82	2
CT108	B01	B01	-0,931	0,058	0,94	0,85	0,44	0,85	1

Tabla 4.9 Dificultad y ajuste de los ítems al modelo Rasch, Matemática de 6.º grado de primaria

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA01	B17	B17	-2,19	0,078	0,85	0,67	0,37	0,93	1
MA02	B18	B18	-1,788	0,068	0,91	0,68	0,39	0,91	1
MA03	B20	B20	-0,4	0,061	0,89	0,82	0,53	0,72	2
MA04	B18	B18	-1,218	0,057	0,91	0,74	0,43	0,85	1
MA05	B17	B17	-0,234	0,046	1,10	1,30	0,36	0,71	2
MA06	B17	B17	0,373	0,043	0,85	0,81	0,56	0,60	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA07	B18	B18	-0,531	0,048	0,99	1,13	0,40	0,76	2
MA08	B20	B20	-0,703	0,065	0,86	0,90	0,52	0,77	2
MA09	B19	B19	-0,978	0,066	0,91	0,93	0,45	0,80	1
MA10	B17	B17	-1,053	0,054	0,96	0,88	0,41	0,83	1
MA11	B19	B19	0,304	0,055	0,86	0,79	0,56	0,58	2
MA12	B17	B17	-0,918	0,053	0,92	0,85	0,45	0,82	1
MA13	B18	B18	0,688	0,041	1,07	1,09	0,39	0,53	3
MA14	B19	B19	-1,176	0,069	0,87	0,77	0,47	0,83	1
MA15	B19	B19	0,198	0,055	1,06	1,08	0,42	0,60	2
MA16	B20	B20	1,732	0,059	1,04	1,29	0,43	0,32	4
MA17	B18	B18	0,406	0,042	1,23	1,38	0,26	0,59	3
MA18	B17	B17	-0,407	0,047	0,91	0,98	0,48	0,74	2
MA19	B20	B20	-1,049	0,070	0,96	1,18	0,42	0,82	1
MA20	B20	B20	0,745	0,055	1,05	1,13	0,45	0,51	3
MA21	B19	B19	1,598	0,058	1,20	1,45	0,29	0,32	4
MA22	B17	B17	-	-	-	-	-	-	-
MA23	B19	B19	1,401	0,056	1,16	1,33	0,33	0,36	3
MA24	B18	B18	-0,178	0,045	1,02	1,19	0,38	0,70	2
MA25	B17	B17	-0,623	0,049	1,04	1,11	0,39	0,78	2
MA26	B18	B18	0,590	0,041	1,00	1,01	0,44	0,55	3
MA27	B19	B19	-0,491	0,060	0,95	1,01	0,46	0,73	2
MA28	B18	B18	0,992	0,041	0,95	0,95	0,48	0,47	3
MA29	B19	B19	-0,700	0,062	0,87	0,73	0,51	0,76	2
MA30	B20	B20	1,347	0,056	1,14	1,24	0,39	0,39	3
MA31	B20	B20	-1,538	0,079	0,80	0,57	0,50	0,87	1
MA32	B17	B17	-0,578	0,049	0,89	0,80	0,49	0,77	2

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA33	B20	B20	-1,892	0,088	0,92	0,69	0,40	0,90	1
MA34	-	B07	-0,866	0,076	0,99	0,95	0,48	0,68	2
MA35	B02	B02	-0,016	0,041	0,93	0,89	0,52	0,65	2
MA36	-	B07	-	-	-	-	-	-	-
MA37	B03	-	1,083	0,046	0,97	0,96	0,47	0,48	3
MA38	B02	B02	0,694	0,039	1,01	1,02	0,48	0,52	3
MA39	B01	-	0,656	0,046	1,12	1,16	0,32	0,58	3
MA40	-	B08	-1,683	0,087	0,94	0,87	0,44	0,81	1
MA41	B02	B02	-0,786	0,046	0,89	0,81	0,50	0,78	2
MA42	-	B08	-1,520	0,084	1,02	1,03	0,40	0,79	1
MA43	B06	-	-0,133	0,049	0,97	1,01	0,39	0,73	2
MA44	B04	B04	0,758	0,039	1,07	1,13	0,43	0,50	3
MA45	-	B08	-1,690	0,087	0,91	0,81	0,46	0,81	1
MA46	-	B08	-0,838	0,075	0,95	0,87	0,50	0,68	2
MA47	-	B07	-1,049	0,078	0,87	0,71	0,55	0,71	1
MA48	-	B08	-1,177	0,078	1,03	1,01	0,42	0,74	1
MA49	B18	B18	-1,399	0,060	0,82	0,64	0,47	0,87	1
MA50	B20	B20	-1,146	0,071	0,90	1,04	0,47	0,83	1
MA51	B02	B02	0,428	0,039	0,95	0,95	0,52	0,57	3
MA52	B19	B19	-1,411	0,074	0,92	0,96	0,41	0,86	1
MA53	B05	-	0,377	0,047	0,99	1,00	0,42	0,62	3
MA54	-	B07	-1,389	0,082	0,89	0,72	0,51	0,77	1
MA55	B01	-	1,249	0,046	1,06	1,11	0,38	0,45	3
MA56	B04	B04	0,289	0,039	0,88	0,83	0,55	0,59	2
MA57_1	B13	-	0,629	0,051	0,86	0,84	0,64	0,88	3
MA57_2	B13	-	1,792	0,051	0,86	0,84	0,64	0,88	4

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA58	B05	-	1,421	0,046	0,93	0,91	0,50	0,41	4
MA59	B06	-	-0,282	0,051	0,99	0,94	0,37	0,75	2
MA60	B12	B12	0,716	0,059	0,83	0,75	0,60	0,49	3
MA61	B03	-	1,745	0,048	0,90	0,89	0,52	0,35	4
MA62	B05	-	-	-	-	-	-	-	-
MA63	-	B15	0,490	0,095	0,79	0,71	0,64	0,44	3
MA64	B11	-	2,533	0,093	0,83	0,78	0,54	0,21	4
MA65_1	B10	-	3,109	0,095	1,04	1,08	0,40	0,14	4
MA65_2	B10	-	4,063	0,095	1,04	1,08	0,40	0,14	4
MA66_1	B09	-	0,339	0,049	0,91	0,88	0,63	1,12	2
MA66_2	B09	-	1,119	0,049	0,91	0,88	0,63	1,12	3
MA67	B03	-	0,346	0,047	0,99	0,96	0,44	0,63	2
MA68	B05	-	1,817	0,049	1,00	1,10	0,44	0,33	4
MA69	B05	-	1,863	0,049	1,03	1,14	0,42	0,32	4
MA70_1	-	B16	-1,038	0,065	0,93	0,88	0,66	1,22	1
MA70_2	-	B16	0,267	0,065	0,93	0,88	0,66	1,22	2
MA71	B02	B02	0,812	0,039	1,03	1,10	0,47	0,49	3
MA72	B01	-	0,981	0,045	1,15	1,22	0,31	0,51	3
MA73	B03	-	1,637	0,047	1,22	1,42	0,27	0,37	4
MA74	B04	B04	0,463	0,039	1,06	1,08	0,44	0,55	3
MA75	-	B07	-0,844	0,076	1,03	1,02	0,45	0,68	2
MA76	B01	-	1,485	0,046	0,92	0,90	0,50	0,40	4
MA77_1	B14	B14	0,811	0,039	0,84	0,76	0,69	0,78	3
MA77_2	B14	B14	1,424	0,039	0,84	0,76	0,69	0,78	4
MA78	B06	-	-0,475	0,053	0,94	0,85	0,41	0,78	2
MA79	B01	-	1,268	0,046	1,01	1,04	0,43	0,45	3

Ítem	Bloque A	Bloque B	Medida	Error	Infit	Outfit	PTME	p	Nivel
MA80	B04	B04	-0,514	0,043	1,02	1,08	0,42	0,74	2
MA81	B06	-	1,031	0,045	1,06	1,08	0,37	0,48	3
MA82	B06	-	1,641	0,047	1,20	1,38	0,23	0,36	4
MA83	B03	-	1,030	0,046	0,96	0,97	0,48	0,49	3
MA84	B20	B20	1,325	0,056	1,13	1,27	0,39	0,40	3
MA85	B19	B19	1,061	0,055	1,16	1,40	0,33	0,42	3
MA86	B18	B18	0,716	0,041	0,95	1,00	0,48	0,52	3
MA87	-	B22	0,540	0,052	1,05	1,07	0,48	0,42	3
MA88	-	B22	0,265	0,051	0,99	1,00	0,52	0,47	2
MA89	-	B22	0,986	0,054	1,03	1,11	0,48	0,34	3
MA90	B21	-	0,594	0,026	0,99	0,97	0,44	0,58	3
MA91	B21	-	1,538	0,027	0,96	0,98	0,47	0,39	4
MA92	B21	-	1,591	0,027	1,10	1,21	0,36	0,38	4
MA93	B17	B17	1,365	0,043	1,16	1,34	0,37	0,40	3
MA94	B04	B04	1,092	0,039	0,97	0,97	0,50	0,43	3

4.3.2 Evidencias de validez vinculadas a la estructura interna

Para la EM 2022, se realizó un análisis de componentes principales de los residuos estandarizados luego de ajustar el modelo Rasch unidimensional (Linacre, 2023). Los resultados se presentan en las siguientes tablas.

Tabla 4.10 Análisis de unidimensionalidad de las medidas derivadas de las pruebas de 2.º grado de secundaria

Prueba	Varianza de la dimensión principal (%)	Varianza modelada (%)	Primer autovalor	Primer autovalor (%)
Lectura	30,3	30,6	1,72	1,0
Matemática	29,5	29,3	1,71	1,1
Ciencia y Tecnología	23,6	23,5	1,65	1,1

Tabla 4.11 Análisis de unidimensionalidad de las medidas derivadas de las pruebas de primaria

Prueba	Varianza de la dimensión principal (%)	Varianza modelada (%)	Primer autovalor	Primer autovalor (%)
Lectura, 2.º grado de primaria	27,1	27,2	2,43	2,0
Matemática, 2.º grado de primaria	32,0	31,5	1,93	1,5
Lectura, 4.º grado de primaria	28,0	37,9	2,11	1,6
Matemática, 4.º grado de primaria	32,2	32,0	1,73	1,2
Lectura, 6.º grado de primaria	31,4	31,6	1,50	1,0
Matemática, 6.º grado de primaria	33,2	32,8	1,40	1,0

Linacre (2023) afirma que, aunque 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, entonces no existen evidencias serias que atenten contra el supuesto de unidimensionalidad, por lo que 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 en las pruebas analizadas. Por ende, no hay evidencias contundentes en contra del supuesto de unidimensionalidad del conjunto de ítems incluidos en estos instrumentos de medición.

4.3.3 Confiabilidad y consistencia de la clasificación

Para expresar los resultados de estos análisis, se utilizan tanto el índice de confiabilidad de la separación de las personas (Rp) como el índice de separación de personas (Gp).

Tabla 4.12 Análisis de confiabilidad de las medidas derivadas de la aplicación de la EM 2022 de 2.º grado de secundaria

Prueba	Rp	Gp
Lectura, 2.º grado de secundaria	0,88	2,73
Matemática, 2.º grado de secundaria	0,87	2,63
Ciencia y Tecnología, 2.º grado de secundaria	0,77	1,85

Considerando el valor de Rp de la EM 2022 de 2.º grado de secundaria, se puede apreciar que la varianza de error en la prueba de Lectura es de 12 %, en la de Matemática es de 13 % y en la de Ciencia y Tecnología es de 23 %. Por lo tanto, es posible afirmar que las medidas derivadas de estas pruebas poseen adecuadas evidencias de confiabilidad.

Tabla 4.13 Análisis de confiabilidad de las medidas derivadas de las pruebas de primaria

Prueba	Rp	Gp
Lectura, 2.º grado de primaria	0,89	2,90
Matemática, 2.º grado de primaria	0,91	3,14
Lectura, 4.º grado de primaria	0,89	2,91
Matemática, 4.º grado de primaria	0,90	3,00
Lectura, 6.º grado de primaria	0,78	1,89
Matemática, 6.º grado de primaria	0,82	2,10

Del mismo modo, en la EM 2022, se puede apreciar que la varianza de error es, como máximo, de 22 % en la prueba de Lectura de 6.º grado de primaria y, como mínimo, de 9 % en la de Matemática de 2.º grado de primaria. Por tanto, es posible afirmar que las medidas derivadas de aplicar dichas pruebas poseen adecuadas evidencias de confiabilidad.

Como se señaló con anterioridad, otros indicadores importantes son los índices de precisión y consistencia de clasificación, cuyos resultados se presentan en las tablas que aparecen a continuación.

Tabla 4.14 Indicadores del análisis de precisión y consistencia de la clasificación usando las medidas derivadas de las pruebas de 2.º grado de secundaria

	Lectura		Matemática		Ciencia y Tecnología	
	Precisión	Consistencia	Precisión	Consistencia	Precisión	Consistencia
Nacional	0,73	0,63	0,75	0,66	0,69	0,58
Hombre	0,73	0,63	0,75	0,66	0,69	0,58
Mujer	0,73	0,63	0,75	0,67	0,69	0,58
Urbano	0,73	0,62	0,74	0,64	0,69	0,58
Rural	0,75	0,66	0,79	0,71	0,70	0,60
Estatal	0,73	0,63	0,76	0,67	0,69	0,58
No estatal	0,73	0,63	0,72	0,62	0,69	0,59
Alta habilidad	0,72	0,62	0,73	0,64	0,68	0,57
Baja habilidad	0,75	0,66	0,79	0,71	0,70	0,60

Tabla 4.15 Indicadores del análisis de precisión y consistencia de la clasificación usando las medidas derivadas de las pruebas de 2.º grado de primaria

	Lectura		Matemática	
	Precisión	Consistencia	Precisión	Consistencia
Nacional	0,89	0,84	0,89	0,85
Hombre	0,88	0,84	0,89	0,85
Mujer	0,89	0,84	0,89	0,85
Urbano	0,89	0,85	0,89	0,84
Rural	0,87	0,81	0,91	0,87
Estatal	0,88	0,83	0,90	0,86
No estatal	0,91	0,87	0,87	0,82
Polidocente Completo	0,89	0,84	0,89	0,84
Unidocente / Multigrado	0,86	0,81	0,92	0,89
Alta habilidad	0,89	0,85	0,88	0,84
Baja habilidad	0,87	0,81	0,91	0,88

Tabla 4.16 Indicadores del análisis de precisión y consistencia de la clasificación usando las medidas derivadas de las pruebas de 4.º grado de primaria

	Lectura		Matemática	
	Precisión	Consistencia	Precisión	Consistencia
Nacional	0,79	0,70	0,81	0,74
Hombre	0,79	0,70	0,82	0,74
Mujer	0,79	0,71	0,81	0,73
Urbano	0,79	0,71	0,81	0,73
Rural	0,78	0,69	0,83	0,76
Estatal	0,78	0,70	0,81	0,74
No estatal	0,81	0,74	0,81	0,73
Polidocente completo	0,79	0,71	0,81	0,73
Unidocente / multigrado	0,78	0,69	0,83	0,76
Alta habilidad	0,79	0,71	0,81	0,73
Baja habilidad	0,78	0,69	0,82	0,75

Tabla 4.17 Indicadores del análisis de precisión y consistencia de la clasificación usando las medidas derivadas de las pruebas de 6.º grado de primaria

	Lectura		Matemática	
	Precisión	Consistencia	Precisión	Consistencia
Nacional	0,64	0,54	0,71	0,61
Hombre	0,64	0,54	0,72	0,61
Mujer	0,64	0,53	0,71	0,60
Urbano	0,64	0,53	0,70	0,60
Rural	0,67	0,57	0,75	0,65
Estatal	0,64	0,53	0,71	0,61
No estatal	0,65	0,54	0,71	0,60
Polidocente completo	0,64	0,53	0,71	0,60
Unidocente / multigrado	0,69	0,59	0,75	0,66
Alta habilidad	0,64	0,53	0,70	0,59
Baja habilidad	0,67	0,57	0,75	0,65

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

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 et al., 2021). 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, ambas pruebas deben ser equiparadas, es decir, puestas en la misma escala. Esta conversión, en ocasiones, implica el uso de una constante aditiva y, en otras, además, el uso de una constante multiplicativa (Livingston, 2004).

Diversas síntesis teóricas que tratan sobre la equiparación de puntuaciones (Kolen y Brennan, 2014; 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, si se deseara equiparar dos pruebas distintas X e Y, el objetivo sería encontrar un sistema o función para convertir la métrica de X en Y.

Considerando la literatura sobre el tema (Ho y Osborn, 2005; Kolen y Brennan, 2014; Livingston, 2004; Navas, 1996; y Zhu, 1998), se puede señalar que existen 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 de 2.º grado de primaria en la misma métrica que el de 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 de la prueba. Lo esencial de este método es que, en cada grupo, se administra un conjunto de ítems o una prueba común que permite establecer la equivalencia entre las pruebas que se quieren equiparar.

Una vez obtenidos los datos, se utiliza 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, 2014), e implica el cálculo de dos constantes de equiparación a y b , de tal manera que surge la siguiente fórmula:

$$medida_{equiparada} = b + a \times medida_{actual}$$

Donde:

$$a = \frac{S_{medida_{anterior}}}{S_{medida_{actual}}}$$

$$b = \bar{X}_{medida_{anterior}} - a \times \bar{X}_{medida_{actual}}$$

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 constante¹ independientemente del grupo poblacional en el que se 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 un análisis de FDI según la metodología propuesta por Bond et al. (2021). De esta manera, solo se utilizan en la equiparación los ítems sin FDI a lo largo del tiempo.

Desde la tabla 4.18 hasta la 4.28, se presentan los resultados de la equiparación por ítems comunes de las pruebas aplicadas en la EM 2022. Dichos resultados nos permiten señalar que este proceso de equiparación ha sido realizado siguiendo rigurosos criterios técnicos, que sustentan la comparación entre los resultados de las evaluaciones realizadas en años diferentes.

¹ En realidad, la dificultad del ítem debe mantenerse dentro de ciertos márgenes de error al comparar su dificultad en dos poblaciones diferentes.

Tabla 4.18 Análisis FDI de las medidas de la EM 2022 de Lectura de 2.º grado de secundaria usadas en la equiparación de 2019 con 2022

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida trans.	Medida	Error
CT01	2,238	0,008	1,588	0,011	2,230	0,008	0,014
CT02	1,849	0,007	1,091	0,010	1,813	0,036	0,012
CT03	1,476	0,007	0,526	0,010	1,339	0,137	0,012
CT011	0,471	0,007	-0,421	0,011	0,544	-0,073	0,013
CT012	0,058	0,008	-1,442	0,017	-0,312	0,370	0,019
CT014	0,531	0,007	-0,178	0,014	0,748	-0,218	0,016
CT015	1,812	0,007	1,311	0,014	1,998	-0,186	0,016
CT016	0,509	0,007	-0,362	0,014	0,594	-0,085	0,016
CT017	0,666	0,007	0,045	0,013	0,935	-0,269	0,015
CT018	1,763	0,007	1,272	0,013	1,965	-0,202	0,015
CT020	0,387	0,003	-0,406	0,007	0,557	-0,170	0,008
CT021	-1,073	0,005	-2,068	0,010	-0,838	-0,235	0,011
CT022	-0,830	0,004	-1,687	0,009	-0,518	-0,312	0,010
CT023	0,746	0,003	0,068	0,006	0,955	-0,208	0,007
CT024	1,604	0,003	1,290	0,006	1,980	-0,375	0,007
CT028	2,768	0,009	2,288	0,012	2,817	-0,049	0,015
CT029	1,898	0,007	1,179	0,010	1,887	0,012	0,012
CT030	1,984	0,007	1,294	0,010	1,983	0,001	0,012
CT031	2,180	0,008	1,708	0,011	2,331	-0,151	0,014
CT032	2,142	0,008	1,407	0,011	2,078	0,064	0,014
CT038	0,528	0,007	-0,304	0,006	0,642	-0,115	0,009
CT039	1,669	0,007	0,819	0,006	1,585	0,084	0,009
CT040	0,959	0,007	-0,054	0,006	0,852	0,107	0,009

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida trans.	Medida	Error
CT041	1,475	0,007	0,580	0,006	1,384	0,091	0,009
CT048	0,064	0,008	-1,204	0,013	-0,113	0,177	0,015
CT049	0,246	0,008	-1,232	0,013	-0,136	0,382	0,015
CT050	1,753	0,007	0,759	0,010	1,534	0,218	0,012
CT056	0,417	0,007	-0,634	0,011	0,366	0,051	0,013
CT108	1,536	0,007	0,550	0,006	1,359	0,177	0,009
CT110	1,692	0,007	0,623	0,006	1,420	0,272	0,009
CT112	1,756	0,007	0,726	0,006	1,507	0,249	0,009
CT116	0,572	0,007	-0,247	0,014	0,690	-0,119	0,016
CT117	0,563	0,007	-0,903	0,015	0,140	0,423	0,017
CT118	0,480	0,007	-0,387	0,014	0,573	-0,093	0,016
CT119	0,204	0,008	-0,890	0,015	0,151	0,053	0,017
CT120	1,063	0,007	0,260	0,013	1,116	-0,052	0,015

Tabla 4.19 Análisis FDI de las medidas de los ítems de Matemática de 2.º grado de secundaria usadas en la equiparación de 2019 con 2022

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA1	1,298	0,007	0,269	0,013	1,133	0,165	0,015
MA2	2,160	0,003	1,359	0,011	2,199	-0,039	0,011
MA3	0,416	0,008	-0,761	0,011	0,126	0,291	0,014
MA4	0,980	0,003	-0,138	0,010	0,735	0,245	0,010
MA5	0,764	0,007	-0,080	0,013	0,791	-0,028	0,015
MA8	2,069	0,008	1,169	0,014	2,013	0,056	0,016
MA9	0,859	0,003	0,148	0,010	1,014	-0,155	0,010

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA13	-0,869	0,005	-1,818	0,018	-0,908	0,039	0,019
MA14	1,572	0,007	0,581	0,013	1,438	0,134	0,015
MA15	1,092	0,007	0,198	0,013	1,063	0,029	0,015
MA16	2,337	0,008	1,229	0,014	2,072	0,265	0,016
MA18	2,019	0,008	1,055	0,013	1,901	0,117	0,015
MA22	2,268	0,008	1,577	0,011	2,412	-0,144	0,014
MA25	2,581	0,008	1,834	0,015	2,663	-0,082	0,017
MA28	1,900	0,008	1,165	0,011	2,009	-0,109	0,014
MA29	-0,369	0,004	-1,183	0,015	-0,287	-0,082	0,016
MA30	2,376	0,008	1,535	0,014	2,371	0,005	0,016
MA32	0,746	0,008	-0,034	0,013	0,836	-0,091	0,015
MA33	-0,445	0,004	-1,162	0,015	-0,267	-0,179	0,016
MA36	0,830	0,097	-0,097	0,013	0,775	0,055	0,098
MA43	1,462	0,093	0,610	0,013	1,466	-0,004	0,094
MA49	1,340	0,090	0,678	0,010	1,533	-0,193	0,091
MA68	0,709	0,003	-0,068	0,013	0,803	-0,094	0,013
MA71	0,731	0,007	-0,521	0,014	0,360	0,371	0,016
MA72	-0,225	0,009	-1,210	0,016	-0,314	0,089	0,018
MA73	-0,332	0,004	-0,993	0,012	-0,101	-0,231	0,013
MA74	1,148	0,003	0,285	0,013	1,148	-0,001	0,013
MA84	0,326	0,008	-0,393	0,014	0,485	-0,159	0,016
MA85	1,851	0,007	0,898	0,013	1,748	0,103	0,015
MA91	0,987	0,007	0,467	0,013	1,326	-0,339	0,015
MA95	-0,291	0,004	-1,242	0,012	-0,345	0,053	0,013
MA96	0,719	0,007	0,034	0,010	0,903	-0,184	0,012

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA98	0,460	0,008	-0,330	0,014	0,547	-0,087	0,016
MA100	1,098	0,007	0,206	0,013	1,071	0,027	0,015
MA102	-0,563	0,004	-1,621	0,017	-0,716	0,153	0,017

Tabla 4.20 Análisis FDI de las medidas de los ítems de Ciencia y Tecnología de 2.º grado de secundaria usadas en la equiparación de 2019 con 2022

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
CN01	-0,193	0,005	0,117	0,013	-0,078	-0,115	0,014
CN04	0,028	0,005	0,145	0,013	-0,051	0,080	0,014
CN05	0,736	0,005	0,970	0,014	0,737	-0,001	0,015
CN08	0,470	0,005	0,719	0,014	0,497	-0,027	0,015
CN13	0,427	0,005	0,633	0,013	0,415	0,012	0,014
CN14	-2,024	0,008	-2,334	0,021	-2,420	0,396	0,022
CN15	-1,135	0,006	-1,021	0,015	-1,166	0,030	0,016
CN16	1,186	0,006	1,439	0,015	1,185	0,001	0,016
CN18	-0,353	0,005	-0,016	0,013	-0,205	-0,148	0,014
CN20	-0,332	0,005	-0,192	0,014	-0,373	0,041	0,015
CN22	-1,115	0,006	-0,794	0,014	-0,949	-0,167	0,015
CN23	0,303	0,005	0,362	0,013	0,156	0,147	0,014
CN25	-0,432	0,005	-0,521	0,014	-0,688	0,255	0,015
CN26	0,679	0,005	0,470	0,013	0,259	0,420	0,014
CN29	0,674	0,005	0,861	0,014	0,633	0,041	0,015
CN30	-0,675	0,006	-0,470	0,014	-0,639	-0,036	0,015

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
CN31	-1,932	0,007	-1,765	0,018	-1,877	-0,055	0,019
CN32	-0,075	0,005	0,451	0,013	0,241	-0,316	0,014
CN33	-1,576	0,007	-1,078	0,015	-1,220	-0,356	0,017
CN34	-1,452	0,006	-1,370	0,016	-1,499	0,048	0,017
CN35	-0,335	0,005	-0,229	0,013	-0,409	0,074	0,014
CN40	-0,689	0,005	-0,368	0,014	-0,542	-0,147	0,015
CN41	-1,557	0,006	-1,535	0,017	-1,657	0,100	0,018
CN42	-2,143	0,008	-2,358	0,021	-2,443	0,300	0,022
CN43	-1,720	0,007	-1,211	0,015	-1,347	-0,373	0,017
CN48	-0,819	0,006	-0,469	0,014	-0,638	-0,181	0,015
CN50	-0,079	0,005	0,052	0,013	-0,140	0,061	0,014
CN51	1,058	0,006	1,261	0,015	1,015	0,043	0,016
CN52	-1,178	0,006	-0,990	0,015	-1,136	-0,042	0,016
CN55	-0,771	0,005	-0,522	0,014	-0,689	-0,082	0,015

Tabla 4.21 Análisis FDI de las medidas de los ítems de Lectura de 2.º grado de primaria usadas en la equiparación de 2019 con 2022

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
CT01	-0,674	0,010	-2,208	0,009	-0,644	-0,031	0,013
CT02	-0,274	0,009	-1,844	0,008	-0,238	-0,036	0,012
CT03	1,115	0,006	-0,544	0,006	1,212	-0,097	0,008
CT04	0,391	0,007	-1,099	0,007	0,593	-0,202	0,010
CT05	2,015	0,005	0,294	0,006	2,146	-0,131	0,008
CT06	1,675	0,017	-0,388	0,006	1,385	0,289	0,018

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
CT07	2,060	0,017	-0,091	0,006	1,717	0,343	0,018
CT08	1,730	0,006	0,290	0,006	2,141	-0,412	0,008
CT10	0,298	0,007	-1,251	0,007	0,423	-0,125	0,010
CT15	2,723	0,008	0,820	0,006	2,732	-0,009	0,010
CT16	2,365	0,008	0,269	0,006	2,118	0,248	0,010
CT17	1,839	0,008	0,019	0,006	1,839	-0,001	0,010
CT21	1,947	0,008	0,001	0,006	1,819	0,128	0,010
CT22	0,906	0,009	-1,102	0,007	0,589	0,317	0,011
CT23	2,384	0,008	0,358	0,006	2,217	0,167	0,010
CT24	2,537	0,008	0,311	0,006	2,165	0,373	0,010
CT29	-0,565	0,010	-2,067	0,009	-0,486	-0,079	0,013
CT37	2,865	0,008	0,916	0,009	2,839	0,026	0,012
CT38	1,884	0,008	0,162	0,008	1,999	-0,115	0,011
CT39	2,190	0,008	0,248	0,008	2,094	0,095	0,011
CT42	2,425	0,006	0,899	0,009	2,820	-0,395	0,011
CT44	2,518	0,006	0,702	0,008	2,601	-0,083	0,010
CT52	2,459	0,008	0,483	0,008	2,356	0,103	0,011
CT53	2,262	0,008	0,653	0,009	2,546	-0,284	0,012
CT76	2,430	0,008	0,661	0,008	2,555	-0,125	0,011
CT78	2,530	0,008	0,683	0,008	2,579	-0,049	0,011
CT79	2,773	0,008	1,110	0,009	3,055	-0,283	0,012
CT80	1,878	0,008	-0,103	0,006	1,703	0,175	0,010
CT82	2,601	0,008	0,729	0,006	2,631	-0,030	0,010
CT83	2,597	0,008	0,501	0,006	2,377	0,221	0,010

Tabla 4.22 Análisis FDI de las medidas de los ítems de 2.º grado de primaria usadas en la equiparación de 2019 con 2022

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA01	-0,593	0,017	-0,832	0,009	-0,655	0,061	0,019
MA02	0,660	0,006	0,341	0,009	0,620	0,040	0,011
MA03	1,038	0,006	0,645	0,009	0,950	0,088	0,011
MA04	0,725	0,006	0,692	0,009	1,001	-0,276	0,011
MA05	-0,294	0,006	-0,498	0,009	-0,292	-0,003	0,011
MA10	1,525	0,006	0,995	0,009	1,330	0,195	0,011
MA11	-0,699	0,017	-0,958	0,009	-0,792	0,092	0,019
MA13	-1,041	0,018	-1,038	0,009	-0,878	-0,163	0,020
MA15	0,411	0,017	0,184	0,009	0,449	-0,038	0,019
MA16	0,087	0,006	-0,081	0,009	0,161	-0,074	0,011
MA22	0,710	0,006	0,491	0,009	0,783	-0,073	0,011
MA24	-0,297	0,017	-0,340	0,009	-0,120	-0,177	0,019
MA25	1,841	0,009	1,579	0,010	1,965	-0,124	0,013
MA30	2,175	0,009	1,677	0,010	2,071	0,103	0,013
MA31	1,880	0,009	1,281	0,009	1,641	0,239	0,013
MA32	-0,054	0,017	-0,335	0,009	-0,115	0,061	0,019
MA33	0,247	0,006	-0,106	0,009	0,134	0,113	0,011
MA39	0,860	0,006	0,717	0,009	1,028	-0,168	0,011
MA41	-1,288	0,018	-1,489	0,010	-1,368	0,081	0,021
MA42	-1,065	0,018	-1,398	0,010	-1,270	0,205	0,021
MA50	1,161	0,009	0,951	0,009	1,282	-0,122	0,013
MA52	-1,229	0,018	-1,305	0,010	-1,169	-0,061	0,021

Tabla 4.23 Análisis FDI de las medidas de los ítems de Lectura de 4.º grado de primaria usadas en la equiparación de 2019 con 2022

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
CT01	1,307	0,012	1,643	0,012	1,123	0,184	0,017
CT02	1,080	0,012	1,614	0,012	1,095	-0,015	0,017
CT03	1,166	0,012	1,474	0,012	0,962	0,204	0,017
CT04	-0,356	0,011	0,190	0,011	-0,257	-0,099	0,016
CT05	-1,283	0,013	-0,865	0,012	-1,259	-0,024	0,018
CT37	-1,027	0,007	-0,761	0,007	-1,160	0,133	0,010
CT39	-0,225	0,006	0,557	0,011	0,091	-0,316	0,013
CT40	-0,841	0,007	-0,687	0,012	-1,090	0,249	0,014
CT42	-0,914	0,007	-0,486	0,012	-0,899	-0,015	0,014
CT44	-1,769	0,014	-1,004	0,013	-1,391	-0,378	0,019
CT45	-1,112	0,012	-0,395	0,011	-0,813	-0,299	0,016
CT46	-0,470	0,011	0,168	0,011	-0,278	-0,192	0,016
CT47	-0,762	0,012	-0,162	0,011	-0,591	-0,171	0,016
CT48	0,069	0,011	0,931	0,011	0,447	-0,378	0,016
CT49	0,032	0,011	0,399	0,011	-0,059	0,091	0,016
CT51	-1,744	0,014	-1,479	0,014	-1,842	0,098	0,020
CT52	-1,013	0,012	-0,815	0,012	-1,211	0,198	0,017
CT53	-2,383	0,017	-2,332	0,017	-2,652	0,269	0,024
CT54	-2,420	0,017	-2,446	0,018	-2,760	0,340	0,025
CT55	-0,887	0,007	-0,491	0,007	-0,904	0,017	0,010
CT56	-1,463	0,007	-0,810	0,007	-1,207	-0,256	0,010
CT57	-0,448	0,006	0,003	0,006	-0,435	-0,013	0,008
CT58	0,761	0,031	0,865	0,081	0,384	0,377	0,087

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equipada	Medida	Error
CT59	0,583	0,011	0,876	0,011	0,394	0,189	0,016
CT60	-2,068	0,015	-2,027	0,016	-2,362	0,294	0,022
CT61	0,317	0,006	0,935	0,006	0,450	-0,133	0,008
CT62	-0,838	0,007	-0,480	0,007	-0,893	0,055	0,010
CT86	0,334	0,006	0,888	0,006	0,406	-0,072	0,008
CT87	-1,475	0,008	-0,685	0,007	-1,088	-0,387	0,011
CT88	0,612	0,007	1,011	0,007	0,522	0,090	0,010
CT90	-1,306	0,007	-0,886	0,007	-1,279	-0,027	0,010
CT91	0,928	0,012	1,078	0,007	0,586	0,342	0,014
CT92	0,191	0,011	0,733	0,006	0,258	-0,067	0,013
CT93	-0,659	0,011	0,109	0,006	-0,334	-0,325	0,013
CT94	-0,120	0,011	0,294	0,006	-0,158	0,038	0,013

Tabla 4.24 Análisis FDI de las medidas de los ítems de Matemática de 4.º grado de primaria usadas en la equiparación de 2019 con 2022

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equipada	Medida	Error
MA002	-2,225	0,009	-1,377	0,014	-2,429	0,204	0,017
MA004	-1,351	0,008	-0,782	0,013	-1,774	0,423	0,015
MA009	-1,246	0,008	-0,028	0,012	-0,944	-0,302	0,014
MA013	-1,506	0,008	-0,457	0,012	-1,416	-0,090	0,014
MA014	-1,546	0,014	-0,599	0,012	-1,573	0,027	0,018
MA015	-1,747	0,015	-0,763	0,013	-1,753	0,006	0,020
MA020	-0,815	0,012	0,068	0,011	-0,839	0,024	0,016
MA021	-0,036	0,007	0,626	0,011	-0,225	0,189	0,013

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA022	-2,791	0,011	-1,440	0,015	-2,498	-0,293	0,019
MA023	-1,616	0,014	-0,898	0,013	-1,902	0,286	0,019
MA024	-0,201	0,007	0,539	0,011	-0,321	0,120	0,013
MA025	0,590	0,007	1,257	0,011	0,469	0,121	0,013
MA026	-0,951	0,013	0,191	0,012	-0,703	-0,248	0,018
MA029	1,244	0,012	1,932	0,012	1,212	0,032	0,017
MA030	-2,973	0,012	-1,994	0,017	-3,107	0,134	0,021
MA031	0,309	0,007	1,115	0,011	0,313	-0,004	0,013
MA032	0,700	0,007	1,171	0,011	0,375	0,325	0,013
MA033	-1,775	0,008	-0,747	0,013	-1,735	-0,040	0,015
MA034	-0,250	0,007	0,890	0,011	0,066	-0,316	0,013
MA035	-2,086	0,009	-0,983	0,013	-1,995	-0,091	0,016
MA039	0,767	0,007	1,717	0,012	0,976	-0,209	0,014
MA041	-0,723	0,012	-0,113	0,012	-1,038	0,315	0,017
MA093	-0,682	0,007	0,250	0,011	-0,639	-0,043	0,013
MA094	-0,473	0,012	0,658	0,011	-0,190	-0,283	0,016
MA096	-1,889	0,009	-0,777	0,013	-1,768	-0,121	0,016
MA098	-2,239	0,016	-1,123	0,013	-2,149	-0,090	0,021
MA101	-1,231	0,008	-0,219	0,012	-1,155	-0,076	0,014

Tabla 4.25 Análisis FDI de las medidas de los ítems de Lectura de 6.º grado de primaria usadas en el escalamiento en 4.º grado de primaria

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
CT001	-1,465	0,011	-1,801	0,073	-1,315	-0,150	0,074

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
CT003	-0,898	0,010	-1,153	0,062	-0,818	-0,080	0,063
CT006	-0,710	0,010	-1,121	0,062	-0,794	0,084	0,063
CT036	-1,060	0,007	-1,905	0,068	-1,394	0,334	0,068
CT041	-0,904	0,007	-1,087	0,069	-0,768	-0,136	0,069
CT042	-1,208	0,007	-1,840	0,084	-1,345	0,137	0,084
CT043	-0,435	0,006	-0,702	0,064	-0,473	0,038	0,064
CT045	0,450	0,006	0,487	0,056	0,437	0,013	0,056
CT046	-0,894	0,007	-1,298	0,073	-0,929	0,035	0,073
CT061	-1,492	0,007	-1,559	0,062	-1,129	-0,363	0,062
CT064	0,406	0,006	0,646	0,056	0,559	-0,153	0,056
CT065	-1,089	0,007	-1,513	0,081	-1,094	0,005	0,081
CT066	0,522	0,007	0,956	0,056	0,796	-0,274	0,056
CT068	-1,280	0,007	-1,303	0,077	-0,933	-0,347	0,077
CT076	0,586	0,007	0,122	0,057	0,158	0,428	0,057
CT077	0,258	0,006	-0,061	0,058	0,018	0,240	0,058
CT078	-0,334	0,006	-0,081	0,058	0,002	-0,336	0,058
CT079	-0,159	0,006	-0,196	0,059	-0,086	-0,073	0,059
CT100	-0,993	0,007	-1,732	0,065	-1,262	0,269	0,065
CT108	-0,601	0,007	-1,299	0,058	-0,930	0,329	0,058

Tabla 4.26 Análisis FDI de las medidas de los ítems de Matemática de 6.º grado de primaria usadas en el escalamiento en 4.º grado de primaria

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA02	-1,774	0,013	-2,127	0,068	-1,789	0,015	0,069

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA04	-1,417	0,012	-1,519	0,057	-1,219	-0,198	0,058
MA06	0,435	0,009	0,179	0,043	0,372	0,063	0,044
MA07	-0,839	0,011	-0,786	0,048	-0,532	-0,307	0,049
MA08	-0,321	0,011	-0,969	0,065	-0,704	0,383	0,066
MA09	-0,704	0,012	-1,263	0,066	-0,979	0,275	0,067
MA10	-1,180	0,012	-1,343	0,054	-1,054	-0,126	0,055
MA11	0,313	0,011	0,106	0,055	0,304	0,009	0,056
MA12	-0,744	0,012	-1,199	0,053	-0,919	0,175	0,054
MA13	0,975	0,012	0,516	0,041	0,688	0,287	0,043
MA14	-1,037	0,012	-1,474	0,069	-1,177	0,140	0,070
MA15	-0,160	0,011	-0,008	0,055	0,197	-0,357	0,056
MA16	1,649	0,009	1,630	0,059	1,733	-0,084	0,060
MA18	-0,420	0,009	-0,653	0,047	-0,407	-0,013	0,048
MA19	-0,967	0,012	-1,338	0,070	-1,050	0,083	0,071
MA20	0,739	0,009	0,576	0,055	0,745	-0,006	0,056
MA23	1,603	0,008	1,276	0,056	1,401	0,202	0,057
MA24	-0,393	0,009	-0,409	0,045	-0,179	-0,214	0,046
MA25	-0,780	0,008	-0,884	0,049	-0,624	-0,156	0,050
MA26	0,724	0,012	0,411	0,041	0,590	0,134	0,043
MA27	-0,776	0,009	-0,743	0,060	-0,492	-0,284	0,061
MA28	1,035	0,008	0,840	0,041	0,992	0,043	0,042
MA29	-0,661	0,010	-0,966	0,062	-0,701	0,040	0,063
MA30	1,058	0,010	1,219	0,056	1,347	-0,289	0,057
MA31	-1,697	0,013	-1,860	0,079	-1,539	-0,158	0,080
MA33	-1,595	0,011	-2,238	0,088	-1,893	0,298	0,089

Ítem	2019		2022		EQP	FDI	
	Medida	Error	Medida	Error	Medida equiparada	Medida	Error
MA40	-1,769	0,013	-2,015	0,087	-1,684	-0,085	0,088
MA47	-1,165	0,012	-1,338	0,078	-1,050	-0,115	0,079
MA49	-1,155	0,012	-1,712	0,060	-1,400	0,245	0,061

Tabla 4.27 Constantes de equiparación de la EM 2022 de 2.º grado de secundaria

Prueba	a	b
Lectura	0,839	0,898
Matemática	0,978	0,870
Ciencia y Tecnología	0,956	-0,190

Tabla 4.28 Constantes de equiparación de la EM 2022 de 2.º, 4.º y 6.º grado de primaria

Prueba	a	b
Lectura, 2.º grado de primaria	1,115	1,818
Matemática, 2.º grado de primaria	1,086	0,249
Lectura, 4.º grado de primaria	0,950	-0,438
Matemática, 4.º grado de primaria	1,100	-0,914
Lectura, 6.º grado de primaria	0,766	0,064
Matemática, 6.º grado de primaria	0,937	0,205

Nota: El último reporte de 6.º grado de primaria en las áreas de Lectura y Matemática data de la EM 2013. Por ello, para el 2022, se ha realizado una transformación de la medida de los ítems distinta a las de los otros grados. Esta se conoce como escalamiento vertical (véase Kolen y Brennan, 2014) y consistió en escalar la métrica de 6.º grado de primaria a la de 4.º grado de primaria, previa equiparación de 4.º grado de primaria con el año 2019 (la cual mantiene un escrutinio histórico con la escala primaria).

4.5 Niveles de logro y preparación de resultados

El diseño de las evaluaciones permite estimar, a partir de su desempeño en las pruebas, lo que los estudiantes saben y hacen respecto de lo que deberían saber y hacer. De acuerdo con ello, la interpretación de los resultados de la EM está referida a criterios o estándares de logro (Muñiz, 1998). Para hacerlo, 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 determina qué es lo que un estudiante mínimamente competente debe poder 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, con la 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 las pruebas de Lectura y Matemática en 6.º grado de primaria.

Para establecer puntos de corte, se utiliza el método Bookmark (Cizek y Bunch, 2007; Karantonis y Sireci, 2006; Lewis et al., 1996), el cual 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, implica colocar marcas (tantas como cortes se hayan preestablecido) en un cuadernillo de ítems ordenado por dificultad (la cual es 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, determinen, 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 especialistas (para el caso, 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, reciben un reporte de discrepancias en el que se señala qué tan distintos son sus juicios 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 los datos (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 esta 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 de 2004 (Ministerio de Educación del Perú, 2005), se establece que un estudiante está en un nivel de logro si tiene una probabilidad mayor o igual a 0,62 de 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, la cual es medio logit superior a la medida de dificultad del ítem en cuestión. Una vez hecha esta adición para cada estudiante evaluado, se transforman los valores a medida 500 (M500) y se redondean a dos decimales. Finalmente, se asigna a cada estudiante a un determinado nivel de logro según su M500.

En las tablas 4.29 y 4.30, se presentan los puntos de corte (en M500) de las evaluaciones muestrales del 2022.

Tabla 4.29 Medidas (en M500) que marcan los puntos de corte de los alumnos en la EM 2022 de 2.º grado de secundaria

Área	< Nivel 1 vs. Nivel 1	Nivel 1 vs. Nivel 2	Nivel 2 vs Nivel 3
Lectura	505,14	580,61	641,25
Matemática	519,67	595,96	649,38
Ciencia y Tecnología	374,60	509,58	628,48

Tabla 4.30 Medida (en M500) que marcan los puntos de corte de los alumnos en las EM 2022 de 2.º, 4.º y 6.º grado de primaria

Área	< Nivel 1 vs. Nivel 1	Nivel 1 vs. Nivel 2	Nivel 2 vs Nivel 3
Lectura, 2.º grado de primaria	458,39	583,66	-
Matemática, 2.º grado de primaria	512,22	639,21	-
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, 6.º grado de primaria	445,46	522,17	579,14
Matemática, 6.º grado de primaria	432,32	526,61	603,41

Finalmente, es importante señalar que las medidas de habilidad de los estudiantes, así como los puntos de corte antes mencionados, fueron transformadas linealmente ($Y = a + bX$), de tal manera que, en la primera evaluación que se realizó en cada grado y área, la media aritmética fue 500 y la desviación estándar fue 100. Para ello, se usaron las constantes que aparecen en las tablas 4.31 y 4.32.

Tabla 4.31 Valores de a y b utilizados para la transformación lineal de las medidas de la EM de 2.º grado de secundaria

Área	a	b
Lectura	474,896	73,414
Matemática	463,157	73,781
Ciencia y Tecnología	496,120	134,513

Tabla 4.32 Valores de a y b utilizados para la transformación lineal de las medidas de la EM 2022 de 2.º, 4.º y 6.º grado de primaria

Área	a	b
Lectura, 2.º grado de primaria	423,614	63,585
Matemática, 2.º grado de primaria	468,321	72,44
Lectura, 4.º grado de primaria	474,896	73,414
Matemática, 4.º grado de primaria	463,157	73,781
Lectura, 6.º grado de primaria	474,896	73,414
Matemática, 6.º grado de primaria	463,157	73,781

4.6 Resultados del análisis psicométrico de habilidades socioemocionales

En relación con el análisis psicométrico de las escalas de habilidades socioemocionales (HSE), se aplicó el modelo de escala de valoración de Rasch (Rasch Rating Scale Model; Andrich, 1978) en su versión unidimensional para las escalas de autorregulación conductual y responsabilidad, y en su versión multidimensional para las escalas de autoeficacia, autonomía y relación, empatía, y toma de decisiones.

Las medidas de habilidad de las personas fueron estimadas con la metodología de valores plausibles. Se emplearon cinco de ellos para cada persona evaluada, ya que se ha demostrado la superioridad de esta metodología para reportar estimaciones de la población (Wu, 2005). La estimación de estos valores plausibles consideró una regresión latente en la cual se usaron como covariables el sexo de los estudiantes, el tipo de gestión de la IE y el área geográfica de la IE. Además de las escalas

de HSE, se incluyó una escala para evaluar el sesgo por deseabilidad social, la cual se analizó a través del modelo Rasch para ítems dicotómicos (Rasch, 1960). Posteriormente, se analizaron las correlaciones entre las medidas de HSE y las de deseabilidad social. Dado que, en general, estas correlaciones fueron pequeñas, se infiere que las respuestas de los estudiantes no fueron influenciadas marcadamente por su deseabilidad social.

4.6.1 Calibración de los ítems

En las tablas 4.33 y 4.34, se presentan los resultados de la calibración de los ítems de cada una de las escalas de HSE. La columna denominada “Medida” representa la dificultad del ítem. Así, una menor medida implica que el ítem es más fácil de endosar (es decir, es más fácil que los estudiantes respondan marcando alguna de las categorías de respuesta más altas), mientras que una mayor medida implica que el ítem es más difícil de endosar (es decir, es más difícil que los estudiantes respondan marcando alguna de las categorías de respuesta más altas). Asimismo, la dificultad de cada ítem viene acompañada de un error que representa qué tan precisa ha sido su calibración. Así, mientras más pequeño es el error, más precisa ha sido la calibración de la dificultad del ítem (y viceversa). Como se puede observar en ambas tablas, los errores de calibración son muy pequeños, lo que indica que la calibración de la dificultad fue bastante precisa. Por último, en ambas tablas, se observa que el infit y el outfit de todos los ítems de las escalas de HSE utilizadas en la EM 2022 mostraron valores dentro del rango aceptable, por lo que se concluye que existe evidencia de un buen ajuste de los datos al modelo Rasch.

Tabla 4.33 Calibración de los ítems de las escalas de HSE aplicadas en 6.º grado de primaria

Escala	Subescala	Ítem	Medida	Error	Infit	Outfit
		HSEAEF6P001	-0,587	0,011	0,73	0,74
		HSEAEF6P003	-0,775	0,011	0,88	0,88
		HSEAEF6P005	-1,003	0,012	0,72	0,73
Autoeficacia	Autoeficacia académica	HSEAEF6P010	0,024	0,010	1,08	1,09
		HSEAEF6P014	-0,647	0,011	0,70	0,71
		HSEAEF6P015	-0,769	0,011	0,93	0,92
		HSEAEF6P017	-0,278	0,010	0,88	0,88

Escola	Subescola	Ítem	Medida	Error	Infit	Outfit	
Autoeficacia	Autoeficacia emocional	HSEAEF6P008	-0,055	0,011	0,95	0,95	
		HSEAEF6P009	0,000	0,011	0,91	0,91	
		HSEAEF6P016	0,075	0,011	0,98	0,98	
		HSEAEF6P018	-0,061	0,011	1,18	1,18	
	Autoeficacia social	HSEAEF6P002	-0,915	0,011	1,15	1,13	
		HSEAEF6P004	0,072	0,010	1,12	1,14	
		HSEAEF6P007	0,232	0,011	1,30	1,30	
		HSEAEF6P011	0,381	0,011	1,26	1,25	
		HSEAEF6P012	-0,688	0,011	1,06	1,06	
			HSEAEF6P013	0,152	0,011	1,08	1,08
	Autonomía y relación	Autonomía	HSEAYR6P001	-0,115	0,010	0,99	1,01
			HSEAYR6P003	-0,770	0,010	0,99	1,00
			HSEAYR6P004	-1,150	0,010	0,93	0,92
HSEAYR6P007			-1,033	0,010	0,87	0,87	
HSEAYR6P010			-0,707	0,010	1,08	1,08	
HSEAYR6P011			-0,925	0,010	0,95	0,94	
Relación		HSEAYR6P002	-1,396	0,010	1,06	1,06	
		HSEAYR6P005	-1,309	0,010	1,14	1,13	
		HSEAYR6P006	-1,448	0,010	0,99	0,96	
		HSEAYR6P008	-1,132	0,010	1,11	1,12	
		HSEAYR6P009	-1,789	0,020	1,00	0,95	
		HSEAYR6P012	-1,904	0,020	1,20	1,09	
Autorregulación conductual	-	HSEARC6P001	-1,318	0,010	1,01	0,97	
		HSEARC6P002	-0,161	0,010	1,02	1,04	
		HSEARC6P003	-0,457	0,010	1,02	1,03	

Escala	Subescala	Ítem	Medida	Error	Infit	Outfit		
Autorregulación conductual		HSEARC6P004	-0,253	0,010	0,98	0,99		
		HSEARC6P005	-0,562	0,010	1,17	1,21		
		HSEARC6P006	-1,135	0,010	0,85	0,84		
		HSEARC6P007	-0,815	0,010	0,98	0,96		
		HSEARC6P008	-0,476	0,010	0,94	0,96		
		HSEARC6P009	-0,064	0,010	1,03	1,06		
		HSEARC6P010	-0,380	0,010	0,97	0,99		
		HSEARC6P011	-0,937	0,010	1,32	1,34		
		HSEARC6P012	-1,164	0,010	0,85	0,85		
		HSEARC6P013	-0,900	0,010	1,00	0,99		
		HSEARC6P014	-1,012	0,010	1,06	1,06		
		HSEARC6P015	-1,811	0,020	1,12	1,03		
		Empatía	Empatía cognitiva	HSEEMP6P001	-0,379	0,010	0,94	0,96
				HSEEMP6P004	-0,422	0,010	1,08	1,11
				HSEEMP6P005	-0,481	0,010	1,19	1,21
HSEEMP6P006	-0,732			0,010	1,12	1,11		
HSEEMP6P008	-0,343			0,010	1,00	1,01		
HSEEMP6P010	-0,860			0,010	0,97	0,94		
HSEEMP6P012	-0,429			0,010	0,81	0,80		
HSEEMP6P015	-0,515		0,010	1,06	1,06			
Disposición empática			HSEEMP6P002	-1,100	0,010	1,34	1,35	
			HSEEMP6P003	-1,647	0,020	1,17	1,10	
			HSEEMP6P007	-0,739	0,010	0,99	1,06	
			HSEEMP6P009	-1,301	0,010	0,91	0,90	
			HSEEMP6P011	-1,050	0,010	0,97	0,96	

Escala	Subescala	Ítem	Medida	Error	Infit	Outfit
Empatía	Disposición empática	HSEEMP6P013	-1,008	0,010	0,87	0,90
		HSEEMP6P014	-1,245	0,010	0,80	0,80
		HSEEMP6P016	-0,901	0,010	1,07	1,07
Responsabilidad	-	HSERSP6P001	-1,357	0,010	0,83	0,86
		HSERSP6P002	-1,508	0,010	1,14	1,16
		HSERSP6P003	-1,644	0,010	0,93	0,91
		HSERSP6P004	-1,286	0,010	1,10	1,08
		HSERSP6P005	-1,568	0,010	0,93	0,93
		HSERSP6P006	-1,273	0,010	1,08	1,07
		HSERSP6P007	-1,604	0,010	0,98	0,94
Toma de decisiones	Estilo vigilante	HSETDD6P003	-1,122	0,010	1,04	1,05
		HSETDD6P009	-1,459	0,010	1,00	0,96
		HSETDD6P011	-0,589	0,010	0,95	0,95
		HSETDD6P012	-0,953	0,010	0,96	0,96
		HSETDD6P013	-0,941	0,010	0,88	0,90
	Estilo hipervigilante	HSETDD6P001	-0,081	0,010	0,88	0,89
		HSETDD6P006	-0,140	0,010	1,08	1,08
		HSETDD6P007	0,020	0,010	0,91	0,91
		HSETDD6P015	-0,272	0,010	0,98	0,99
	Estilo procrastinador	HSETDD6P002	-0,221	0,010	0,82	0,82
		HSETDD6P004	-0,164	0,010	1,04	1,06
		HSETDD6P008	0,082	0,010	0,98	0,98
		HSETDD6P016	0,215	0,010	0,97	0,96
	Estilo de transferencia	HSETDD6P005	0,446	0,010	1,32	1,30
		HSETDD6P010	0,408	0,010	0,98	1,00
		HSETDD6P014	0,374	0,010	1,06	1,05
		HSETDD6P017	0,386	0,010	1,12	1,12

Tabla 4.34 Calibración de los ítems de las escalas de HSE aplicadas en 2.º grado de secundaria

Escala	Subescala	Ítem	Medida	Error	Infit	Outfit
Autoeficacia	Autoeficacia académica	HSEAEF6P001	-0,576	0,003	0,67	0,69
		HSEAEF6P003	-0,589	0,003	0,80	0,80
		HSEAEF6P005	-1,077	0,003	0,69	0,70
		HSEAEF6P010	-0,069	0,003	0,98	0,99
		HSEAEF6P014	-0,624	0,003	0,69	0,70
		HSEAEF6P015	-0,515	0,003	0,88	0,88
		HSEAEF6P017	-0,201	0,003	0,83	0,83
	Autoeficacia emocional	HSEAEF6P008	0,001	0,003	0,93	0,92
		HSEAEF6P009	-0,012	0,003	0,87	0,87
		HSEAEF6P016	0,094	0,003	1,02	1,02
		HSEAEF6P018	0,007	0,003	1,14	1,14
	Autoeficacia social	HSEAEF6P002	-0,767	0,003	1,19	1,18
		HSEAEF6P004	0,126	0,003	1,11	1,13
		HSEAEF6P007	0,253	0,003	1,32	1,33
		HSEAEF6P011	0,586	0,003	1,28	1,26
		HSEAEF6P012	-0,588	0,003	1,03	1,03
		HSEAEF6P013	0,112	0,003	1,02	1,02
	Autonomía y relación	Autonomía	HSEAYR6P001	-0,286	0,003	1,02
HSEAYR6P003			-0,963	0,004	0,99	1,00
HSEAYR6P004			-1,493	0,004	0,92	0,89
HSEAYR6P007			-1,349	0,004	0,87	0,86
HSEAYR6P010			-0,974	0,004	1,18	1,18
HSEAYR6P011			-1,242	0,004	0,93	0,92

Escala	Subescala	Ítem	Medida	Error	Infit	Outfit
Autonomía y relación	Relación	HSEAYR6P002	-1,497	0,004	1,12	1,13
		HSEAYR6P005	-1,527	0,004	1,08	1,07
		HSEAYR6P006	-1,582	0,004	0,95	0,92
		HSEAYR6P008	-1,336	0,004	1,12	1,13
		HSEAYR6P009	-1,925	0,004	0,93	0,87
		HSEAYR6P012	-2,005	0,004	1,18	1,11
Autorregulación conductual	-	HSEARC6P001	-1,546	0,004	0,98	0,94
		HSEARC6P002	-0,307	0,003	1,11	1,17
		HSEARC6P003	-0,556	0,003	1,00	1,02
		HSEARC6P004	-0,442	0,003	1,00	1,03
		HSEARC6P005	-0,650	0,003	1,11	1,15
		HSEARC6P006	-1,256	0,004	0,82	0,81
		HSEARC6P007	-0,998	0,004	0,97	0,96
		HSEARC6P008	-0,717	0,003	0,99	1,01
		HSEARC6P009	-0,280	0,003	1,04	1,10
		HSEARC6P010	-0,547	0,003	1,01	1,05
		HSEARC6P011	-1,214	0,004	1,31	1,33
		HSEARC6P012	-1,384	0,004	0,83	0,81
		HSEARC6P013	-1,176	0,004	1,01	1,00
		HSEARC6P014	-1,371	0,004	1,04	1,03
		HSEARC6P015	-1,964	0,004	1,07	1,00
Empatía	Empatía cognitiva	HSEEMP6P001	-0,565	0,003	1,00	1,03
		HSEEMP6P004	-0,707	0,003	1,17	1,22
		HSEEMP6P005	-0,667	0,003	1,26	1,30
		HSEEMP6P006	-0,945	0,004	1,17	1,15
		HSEEMP6P008	-0,284	0,003	1,01	1,05

Escala	Subescala	Ítem	Medida	Error	Infit	Outfit	
Empatía	Empatía cognitiva	HSEEMP6P010	-0,929	0,004	0,95	0,93	
		HSEEMP6P012	-0,438	0,003	0,79	0,80	
		HSEEMP6P015	-0,827	0,003	1,02	1,01	
	Disposición empática	HSEEMP6P002	-1,146	0,004	1,35	1,39	
		HSEEMP6P003	-1,638	0,004	1,14	1,10	
		HSEEMP6P007	-0,789	0,003	0,91	0,97	
		HSEEMP6P009	-1,324	0,004	0,88	0,86	
		HSEEMP6P011	-0,979	0,004	0,97	0,97	
		HSEEMP6P013	-1,064	0,004	0,84	0,85	
		HSEEMP6P014	-1,265	0,004	0,75	0,74	
		HSEEMP6P016	-0,957	0,004	1,07	1,07	
		Responsabilidad	-	HSERSP6P001	-1,762	0,004	0,96
	HSERSP6P002		-2,122	0,004	1,11	1,12	
	HSERSP6P003		-2,271	0,004	0,87	0,86	
	HSERSP6P004		-1,891	0,004	1,11	1,08	
	HSERSP6P005		-2,285	0,004	0,95	0,94	
HSERSP6P006	-2,128		0,004	1,07	1,06		
HSERSP6P007	-2,199		0,004	0,96	0,93		
Toma de decisiones	Estilo vigilante	HSETDD6P003	-1,292	0,003	0,98	0,99	
		HSETDD6P009	-1,601	0,004	1,00	0,98	
		HSETDD6P011	-0,754	0,003	0,88	0,89	
		HSETDD6P012	-1,165	0,003	0,87	0,86	
		HSETDD6P013	-1,241	0,003	0,81	0,82	
	Estilo hipervigilante	HSETDD6P001	-0,087	0,003	0,93	0,94	
		HSETDD6P006	-0,157	0,003	1,11	1,10	
		HSETDD6P007	-0,046	0,003	0,89	0,89	
		HSETDD6P015	-0,291	0,003	0,99	0,99	

Escala	Subescala	Ítem	Medida	Error	Infit	Outfit
Toma de decisiones	Estilo procrastinador	HSETDD6P002	-0,259	0,003	0,82	0,83
		HSETDD6P004	-0,333	0,003	1,09	1,10
		HSETDD6P008	0,145	0,003	0,97	0,97
		HSETDD6P016	0,344	0,003	1,02	1,02
	Estilo de transferencia	HSETDD6P005	0,317	0,003	1,32	1,30
		HSETDD6P010	0,429	0,003	0,99	1,01
		HSETDD6P014	0,512	0,003	1,12	1,11
		HSETDD6P017	0,411	0,003	1,19	1,19

4.6.2 Estructura de umbrales

Una característica adicional del modelo de escala de valoración de Rasch (respecto del modelo para datos dicotómicos) es que, además de reportar las estimaciones de habilidad de las personas y de dificultad de cada ítem, provee una estructura de umbrales que es común para todos los ítems. Los umbrales de Andrich representan los puntos en que las categorías de respuesta adyacentes son igualmente probables. Así, cuando un estudiante tiene una medida promedio que excede el valor del umbral de dos categorías adyacentes, tiene una mayor probabilidad de ser observado en la categoría más alta (encima del umbral). Es importante mencionar que se espera que los umbrales estén ordenados de manera creciente, ya que las observaciones en las categorías más altas deberían ser producidas por medidas más altas (Linacre, 1999). Esto quiere decir que las medidas promedio por categoría, para cada conjunto empírico de observaciones, deberían avanzar monotónicamente hacia arriba en la escala de valoración, ya que de otro modo su significado sería incierto y, consecuentemente, cualquier medida derivada sería de dudosa utilidad (Linacre, 1999). Asimismo, se recomienda que la distancia mínima entre los umbrales sea de por lo menos 1,4 logits, aunque esto no necesariamente es un requerimiento para construir medidas válidas e inferencialmente útiles a partir de las observaciones de una escala de valoración (Linacre, 1999).

Tabla 4.35 Estructura de umbrales de las escalas de HSE aplicadas en 6.º grado de primaria

Escala	Primer umbral		Segundo umbral		Tercer umbral		Cuarto umbral	
	Medida	Error	Medida	Error	Medida	Error	Medida	Error
Autoeficacia	-1,282	0,006	-0,331	0,005	0,501	0,006	1,112	-
Autonomía y relación	-1,063	0,007	-0,703	0,007	0,061	0,006	1,705	-
Autorregulación conductual	-1,106	0,006	-0,459	0,006	0,082	0,006	1,483	-
Empatía	-0,817	0,006	-0,616	0,006	-0,108	0,005	1,541	-
Responsabilidad	-0,857	0,009	-1,127	0,008	0,785	0,009	1,199	-
Toma de decisiones	-0,607	0,006	-0,798	0,005	0,666	0,006	0,739	-

Tabla 4.36 Estructura de umbrales de las escalas de HSE aplicadas en 2.º grado de secundaria

Escala	Primer umbral		Segundo umbral		Tercer umbral		Cuarto umbral	
	Medida	Error	Medida	Error	Medida	Error	Medida	Error
Autoeficacia	-1,728	0,002	-0,590	0,001	0,632	0,002	1,686	-
Autonomía y relación	-1,764	0,002	-0,832	0,002	-0,075	0,002	2,671	-
Autorregulación conductual	-1,839	0,002	-0,573	0,002	0,000	0,001	2,412	-
Empatía	-1,381	0,002	-0,709	0,002	-0,178	0,001	2,268	-
Responsabilidad	-1,690	0,003	-1,625	0,002	0,870	0,002	2,445	-
Toma de decisiones	-1,314	0,002	-1,042	0,001	0,835	0,002	1,521	-

4.6.3 Evidencias de validez vinculadas a la estructura interna

En las tablas 4.37 y 4.38, se presentan los resultados de este análisis. A partir de ello, se concluye que todas las escalas y subescalas de HSE utilizadas en la EM 2022 cumplen con el requerimiento de unidimensionalidad y cada una de ellas mide fundamentalmente una sola variable latente.

Tabla 4.37 Análisis de unidimensionalidad de las medidas de HSE derivadas de las escalas aplicadas en 6.º grado de primaria

Escala	Subescala	Medidas Rasch		ACPR	
		Autovalor	Varianza explicada (%)	Autovalor	Varianza explicada (%)
Autoeficacia	Autoeficacia académica	5,0	41,9	1,6	13,0
	Autoeficacia emocional	2,7	34,8	1,5	19,1
	Autoeficacia social	4,2	40,9	1,5	14,8
Autonomía y relación	Autonomía	4,9	45,0	1,4	12,8
	Relación	4,8	44,4	1,3	11,7
Autorregulación conductual	-	8,7	36,6	2,0	8,5
Empatía	Empatía cognitiva	5,3	40,0	1,4	10,8
	Disposición empática	4,4	35,3	1,8	14,8
Responsabilidad	-	3,7	34,6	1,5	13,6
Toma de decisiones	Estilo vigilante	3,5	41,1	1,3	15,7
	Estilo hipervigilante	2,8	41,0	1,5	21,5
	Estilo procrastinador	2,9	42,0	1,5	21,5
	Estilo de transferencia	2,6	39,5	1,7	25,8

Tabla 4.38 Análisis de unidimensionalidad de las medidas de HSE derivadas de las escalas aplicadas en 2.º grado de secundaria

Escala	Subescala	Medidas Rasch		ACPR	
		Autovalor	Varianza explicada (%)	Autovalor	Varianza explicada (%)
Autoeficacia	Autoeficacia académica	6,1	46,7	1,6	12,6
	Autoeficacia emocional	3,4	40,4	1,6	19,2
	Autoeficacia social	4,9	44,8	1,5	13,6
Autonomía y relación	Autonomía	6,1	50,3	1,5	12,7
	Relación	6,3	51,3	1,4	11,1
Autorregulación conductual	-	10,8	41,9	2,2	8,3
Empatía	Empatía cognitiva	6,6	45,4	1,5	10,4
	Disposición empática	5,5	40,8	1,8	13,7
Responsabilidad	-	6,3	47,2	1,3	10,0
Toma de decisiones	Estilo vigilante	4,6	48,0	1,5	15,7
	Estilo hipervigilante	3,6	47,3	1,6	20,4
	Estilo procrastinador	3,7	47,8	1,5	19,3
	Estilo de transferencia	3,3	45,2	1,8	24,2

4.6.4 Confiabilidad

La confiabilidad se calculó utilizando la estimación EAP (expected a posteriori) de las medidas de las personas. En relación con las escalas de HSE aplicadas en 6.º grado de primaria, la mayoría mostró medidas con una confiabilidad por encima del punto de corte ($\leq 0,70$) o muy cercano a este (0,68 y 0,69). Solo la medida derivada de la subescala del estilo vigilante (toma de decisiones) mostró una confiabilidad más alejada del punto de corte esperado (0,63), lo que se podría deber a la baja variabilidad de esta habilidad en la muestra y a la poca cantidad de ítems considerados para

medirla (5 ítems). Una baja confiabilidad implica una menor precisión en la medición de esta habilidad. Sin embargo, dado que la confiabilidad de esta medida no está demasiado alejada del punto de corte, no se desestimó su uso en la EM 2022. Con todo, estos resultados deberían interpretarse con una mayor cautela. Estos resultados aparecen en la tabla 4.39. En cuanto a las escalas de HSE aplicadas en 2.º grado de secundaria, todas las medidas derivadas de aplicarlas mostraron una confiabilidad por encima del punto de corte. Estos resultados aparecen en la tabla 4.40.

Tabla 4.39 Confiabilidad de las medidas de HSE derivadas de las escalas aplicadas en 6.º grado de primaria

Escala	Subescala	Confiabilidad
Autoeficacia	Autoeficacia académica	0,71
	Autoeficacia emocional	0,71
	Autoeficacia social	0,71
Autonomía y relación	Autonomía	0,75
	Relación	0,79
Autorregulación conductual	-	0,83
Empatía	Empatía cognitiva	0,80
	Disposición empática	0,82
Responsabilidad	-	0,69
Toma de decisiones	Estilo vigilante	0,63
	Estilo hipervigilante	0,69
	Estilo procrastinador	0,68
	Estilo de transferencia	0,70

Tabla 4.40 Confiabilidad de las medidas de HSE derivadas de aplicar las escalas en 2.º grado de secundaria

Escala	Subescala	Confiabilidad
Autoeficacia	Autoeficacia académica	0,77
	Autoeficacia emocional	0,75
	Autoeficacia social	0,78
Autonomía y relación	Autonomía	0,82
	Relación	0,85
Autorregulación conductual	-	0,89
Empatía	Empatía cognitiva	0,85
	Disposición empática	0,87
Responsabilidad	-	0,82
Toma de decisiones	Estilo vigilante	0,72
	Estilo hipervigilante	0,79
	Estilo procrastinador	0,78
	Estilo de transferencia	0,77

Referencias

Referencias

- American Educational Research Association, American Psychological Association y National Council of Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43, 561–573. <https://doi.org/10.1007/BF02293814>
- Bond, T. G., Yan, Z. y Heene, M. (2021). *Applying the Rasch model. Fundamental measurement in the human Sciences* (4a ed.). Routledge
- Browning, D. (2003). *Common definitions: Adaptations, accommodations, modifications*. http://www.pent.ca.gov/acc/commondefinitions_accom-mod.pdf
- 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*. SAGE.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>
- Embretson, S., y Reise, S. (2000). *Item Response Theory for psychologists*. Lawrence Erlbaum Associates.
- Fox, C. (1999). An introduction to the Partial Credit model for developing nursing assessments. *Journal of Nursing Education*, 38(8), 340-346. <https://doi.org/10.3928/0148-4834-19991101-04>
- Hambleton, R., Swaminathan, H. y Rogers, J. (1991). *Fundamentals of Item Response Theory*. 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. <https://doi.org/10.7275/68dy-z131>
- Ingebo, G. (1997). *Probability in the measure of achievement*. MESA Press.
- 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 la Annual Meeting of the National Council on Education in Measurement, San Francisco, CA.

- Kolen, M., y Brennan, R. (2014). *Test equating, scaling and linking. Methods and practices* (3a ed.). 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>
- 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.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22, 5-55. https://legacy.voteview.com/pdf/Likert_1932.pdf
- 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. <https://www.rasch.org/rmt/rmt151g.htm>
- Linacre, J. M. (2023). *A user's guide to WINSTEPS, MINISTEP Rasch model computer programs*. <https://www.winsteps.com/a/Winsteps-Manual.pdf>
- Livingston, S. A. (2004). *Equating test scores (without IRT)*. Educational Testing Service. <https://www.ets.org/Media/Research/pdf/LIVINGSTON2ed.pdf>
- Lohr, S. (2021). *Samplig. Design and analysis*. CRC Press.
- 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. & 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*. <http://repositorio.minedu.gob.pe/handle/123456789/1134>
- Ministerio de Educación. (2009). *Evaluación Censal de Estudiantes (ECE). Segundo grado de primaria y cuarto grado de primaria de IE EIB. Marco de trabajo*. <http://repositorio.minedu.gob.pe/handle/123456789/1106>
- Ministerio de Educación. (2015). *Marco de fundamentación de las pruebas de rendimiento de la Evaluación Censal de Estudiantes de 2do grado de secundaria 2015*. <http://repositorio.minedu.gob.pe/handle/123456789/5055>
- Moreno, R., Martínez, R. J. y Muñoz, J. (2004). Directrices para la construcción de ítems de elección múltiple. *Psicothema*, 16(3), 490-497. <https://www.psicothema.com/pi?pii=3023>

- Muñiz, J. (1998). *Teoría clásica de los tests* (2nd ed.). Pirámide.
- Navas, M. (1996). Equiparación de puntuaciones. En J. Muñiz (Ed.), *Psicometría*, (pp. 293-370). Universitas.
- Nunnally, J., y Bernstein, I. (1995). *Teoría psicométrica* (3rd ed.). McGraw-Hill.
- Paz, M. (1996). Validez. En J. Muñiz (Ed.), *Psicometría* (pp. 49-103). Universitas.
- Rasch, G. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests*. Nielsen & Lydiche.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4(3), 207-230. <https://doi.org/10.2307/1164671>
- Schulz, E. M. (1990). Functional assessment of fit. En J. M. Linacre (Ed.), *Rasch Measurement Transactions, 1990* (p. 82). 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 score reports. En J. M. Linacre (Ed.), *Rasch Measurement Transactions Part 1, 1995* (pp.33-35). MESA Press.
- Suen, H. (1990). *The psychometric process*. Pennsylvania State University.
- Wilson, M. (2005). *Constructing measures. An item response modeling approach*. Lawrence Erlbaum Associates.
- Wind, S., y Hua, C. (2022). *Rasch measurement theory analysis in R*. Chapman & Hall.
- Wright, B. D. (1999). Model selection: Rating Scale Model (RSM) or Partial Credit Model (PCM)? *Rasch Measurement Transactions*, 12(3), 641-642. <https://www.rasch.org/rmt/rmt1231.htm>
- Wright, B. D., y Linacre, J. M. (1994). Reasonable mean-square fit values. En J. M. Linacre (Ed.), *Rasch Measurement Transactions, 1994* (p. 370). MESA Press
- Wright, B. D., y Stone, M. (1998). *Diseño de mejores pruebas*. CENEVAL.
- Wright, B. D., y Masters, G. (1982). *Rating scale analysis*. MESA Press.
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2-3), 114-128. <https://doi.org/10.1016/j.stueduc.2005.05.005>
- 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>

Anexos

Anexo A Tablas de especificaciones de la Evaluación Muestral 2022, Lectura, Matemática, Ciencia y Tecnología

2.º grado de primaria

Tabla A.1 Lectura. Distribución de ítems por capacidades evaluadas en la prueba

Capacidades	Cantidad de ítems	Porcentaje
Lee palabras y oraciones	6	6,7 %
Obtiene información del texto escrito	31	34,4 %
Infiere e interpreta información del texto	39	43,3 %
Reflexiona y evalúa la forma, el contenido y el contexto del texto	14	15,6 %
Total	90	100 %

Tabla A.2 Lectura. Distribución de textos presentes en la prueba por tipo y género textual

Tipo de texto	Género	Correlativo del texto	Cantidad por tipo de texto	Porcentaje por tipo de texto
Narrativo	Cuento	1	4	15,4 %
		2		
		3		
		4		
Instructivo	Manual de procedimientos	5	2	7,7 %
	Texto de recomendaciones	6		
Descriptivo	Artículo enciclopédico	7	6	23,1 %
		8		
		9		
		10		
		11		
		12		
Argumentativo	Afiche	13	4	15,4 %
		14		
		15		
		16		
No tipificable Palabra y oración	Nota	17	10	38,5 %
		18		
		19		
		20		
		21		
		22		
Palabra y oración	Palabra y oración	23	8	30,8 %
		24		
		25		
		26		
		26		
Total			26	100 %

Tabla A.3 Lectura. Distribución de ítems por tipos de textos presentes en la prueba

Tipo de texto	Género	Correlativo del texto	Cantidad de ítems por cada texto	Cantidad de ítems por tipo de texto	Porcentaje de ítems por tipo de texto
Narrativo	Cuento	1	5	18	20,0 %
		2	5		
	Anécdota	3	4		
		4	4		
Instructivo	Manual de procedimientos	5	6	11	12,2 %
	Texto de recomendaciones	6	5		
Descriptivo	Artículo enciclopédico	7	4	28	31,1 %
		8	6		
		9	6		
		10	3		
		11	5		
	Cuadro comparativo	12	4		
Argumentativo	Afiche	13	4	18	20,0 %
		14	6		
		15	4		
	Artículo de opinión	16	4		
No tipificable Palabra y oración	Nota	17	4	15	16,7 %
		18	3		
		19	1		
	Palabra y oración	20	1		
		21	1		
		22	1		
		23	1		
		24	1		
		25	1		
26	1				
Total			90	90	100 %

Tabla A.4 Matemática. Distribución de ítems por competencias evaluadas en la prueba

Competencias	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	92	100 %
Resuelve problemas de regularidad, equivalencia y cambio	0	0 %
Resuelve problemas de forma, movimiento y localización	0	0 %
Resuelve problemas de gestión de datos e incertidumbre	0	0 %
Total	92	100 %

Tabla A.5 Matemática. Distribución de ítems por capacidades evaluadas en la prueba

Competencia	Capacidades	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	Traduce cantidades a expresiones numéricas.	46	50,0 %
	Comunica su comprensión sobre los números y las operaciones.	18	19,6 %
	Usa estrategias y procedimientos de estimación y cálculo.	14	15,2 %
	Argumenta afirmaciones sobre las relaciones numéricas y las operaciones.	14	15,2 %
Total		102	100 %

*El total ha sido redondeado a un decimal. Por ello, la suma total en esta tabla no es exactamente 100 %.

Tabla A.6 Matemática. Distribución de ítems por contextos evaluados en la prueba

Contextos	Cantidad de ítems	Porcentaje
Intramatemático	83	90,2 %
Extramatemático	9	9,8 %
Total	92	100 %

4.º grado de primaria

Tabla A.7 Lectura. Distribución de ítems por capacidades evaluadas en la prueba

Capacidades	Cantidad de ítems	Porcentaje
Obtiene información del texto escrito	28	29,1 %
Infiere e interpreta información del texto	45	46,9 %
Reflexiona y evalúa la forma, el contenido y el contexto del texto	23	24,0 %
Total	96	100 %

Tabla A.8 Lectura. Distribución de textos presentes en la prueba por tipo y género textual

Tipo de texto	Género	Correlativo del texto	Cantidad por tipo de texto	Porcentaje por tipo de texto
Narrativo	Cuento	1	5	29 %
	Historieta	2		
	Noticia	3		
	Anécdota	4		
	Anécdota	5		
Descriptivo	Artículo enciclopédico	6	5	29 %
	Artículo periodístico	7		
	Infografía	8		
		9		
Instructivo	Manual de procedimientos	10	2	12 %
		11		
Argumentativo	Infografía	12	3	18 %
	Afiche	13		
	Artículo de opinión	14		
Expositivo	Artículo de divulgación	15	2	12 %
		16		
		17		
Total			17	100 %

Tabla A.9 Lectura. Distribución de ítems por tipos de textos presentes en la prueba

Tipo de texto	Género	Correlativo del texto	Cantidad de ítems por cada texto	Cantidad de ítems por tipo de texto	Porcentaje de ítems por tipo de texto
Narrativo	Cuento	1	6	24	25,0 %
	Historieta	2	6		
	Noticia	3	6		
	Anécdota	4	3		
		5	3		
Descriptivo	Artículo enciclopédico	6	6	30	31,0 %
		7	6		
	Artículo periodístico	8	6		
		9	6		
Instructivo	Infografía	10	6	12	12,5 %
	Manual de procedimientos	11	6		
Argumentativo		12	6	18	19,0 %
	Infografía	13	6		
	Afiche	14	6		
Expositivo	Artículo de opinión	15	6	12	12,5 %
	Artículo de divulgación	16	6		
		17	6		
Total				96	100 %

Tabla A.10 Matemática. Distribución de ítems por competencias evaluadas en la prueba

Competencias	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	38	37,3 %
Resuelve problemas de regularidad, equivalencia y cambio	27	26,5 %
Resuelve problemas de forma, movimiento y localización	19	18,6 %
Resuelve problemas de gestión de datos e incertidumbre	18	17,6 %
Total	102	100 %

Tabla A.11 Matemática. Distribución de ítems por capacidades evaluadas en la prueba

Competencia	Capacidades	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	Traduce cantidades a expresiones numéricas.	20	19,6 %
	Comunica su comprensión sobre los números y las operaciones.	13	12,7 %
	Usa estrategias y procedimientos de estimación y cálculo.	3	2,9 %
	Argumenta afirmaciones sobre las relaciones numéricas y las operaciones.	2	2,0 %
Resuelve problemas de regularidad, equivalencia y cambio	Traduce datos y condiciones a expresiones algebraicas y gráficas.	4	3,9 %
	Comunica su comprensión sobre las relaciones algebraicas.	3	2,9 %
	Usa estrategias y procedimientos para encontrar equivalencias y reglas generales.	14	13,7 %
	Argumenta afirmaciones sobre las relaciones numéricas y las operaciones.	6	5,9 %
Resuelve problemas de forma, movimiento y localización	Modela objetos con formas geométricas y sus transformaciones.	4	3,9 %
	Comunica su comprensión sobre las formas y relaciones geométricas.	10	9,8 %
	Usa estrategias y procedimientos para orientarse en el espacio.	2	2,0 %
	Argumenta afirmaciones sobre relaciones geométricas.	3	2,9 %
Resuelve problemas de gestión de datos e incertidumbre	Representa datos con gráficos y medidas estadísticas o probabilísticas.	1	1,0 %
	Comunica su comprensión de los conceptos estadísticos y probabilísticos.	10	9,8 %
	Usa estrategias y procedimientos para recopilar y procesar datos.	3	2,9 %
	Sustenta conclusiones o decisiones con base en la información obtenida.	4	3,9 %
Total		102	100 %

*El total ha sido redondeado a un decimal. Por ello, la suma total en esta tabla no es exactamente 100 %.

Tabla A.12 Matemática. Distribución de ítems por contextos evaluados en la prueba

Contextos	Cantidad de ítems	Porcentaje
Intramatemático	88	86,3 %
Extramatemático	14	13,7 %
Total	102	100 %

6.º grado de primaria**Tabla A.13 Lectura. Distribución de ítems por capacidades evaluadas en la prueba**

Capacidades	Cantidad de ítems	Porcentaje
Obtiene información del texto escrito	24	22,2 %
Infiere e interpreta información del texto	56	51,9 %
Reflexiona y evalúa la forma, el contenido y el contexto del texto	28	25,9 %
Total	108	100 %

Tabla A.14 Lectura. Distribución de textos presentes en la prueba por tipo y género textual

Tipo de texto	Género	Correlativo del texto	Cantidad por tipo de texto	Porcentaje por tipo de texto
Narrativo		1	4	22,2 %
		2		
		3		
		4		
Instructivo	Manual de procedimientos	5	3	16,7 %
		6		
		7		
Descriptivo	Artículo enciclopédico	8	4	22,2 %
		9		
		10		
		11		
	Descripción de personaje			

Tipo de texto	Género	Correlativo del texto	Cantidad por tipo de texto	Porcentaje por tipo de texto
Argumentativo	Artículo de opinión	12	4	22,2 %
		13		
		14		
	Carta	15		
Expositivo	Artículo enciclopédico	16	3	16,7 %
		17		
		18		
	Artículo de divulgación	18		
TOTAL			18	100 %

Tabla A.15 Lectura. Distribución de ítems por tipos de textos presentes en la prueba

Tipo de texto	Género	Correlativo del texto	Cantidad de ítems por cada texto	Cantidad de ítems por tipo de texto	Porcentaje de ítems por tipo de texto
Narrativo	Cuento	1	6	24	22,2 %
		2	6		
		3	6		
	Historieta	4	6		
Instructivo	Manual de procedimientos	5	6	18	16,7 %
		6	6		
		7	6		
	Artículo enciclopédico	8	6		
Descriptivo	Artículo enciclopédico	9	6	24	22,2 %
		10	6		
	Descripción de personaje	11	6		
Argumentativo	Artículo de opinión	12	6	24	22,2 %
		13	6		
		14	6		
	Carta	15	6		

Tipo de texto	Género	Correlativo del texto	Cantidad de ítems por cada texto	Cantidad de ítems por tipo de texto	Porcentaje de ítems por tipo de texto
Expositivo	Artículo enciclopédico	16	6	18	16,7 %
	Artículo de divulgación	17	6		
		18	6		
Total			108	108	100 %

Tabla A.16 Matemática. Distribución de ítems por competencias evaluadas en la prueba

Competencias	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	33	35,1 %
Resuelve problemas de regularidad, equivalencia y cambio	22	23,4 %
Resuelve problemas de forma, movimiento y localización	20	21,3 %
Resuelve problemas de gestión de datos e incertidumbre	19	20,2 %
Total	94	100 %

Tabla A.17 Matemática. Distribución de ítems por capacidades evaluadas en la prueba

Competencia	Capacidades	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	Traduce cantidades a expresiones numéricas.	19	20,2 %
	Comunica su comprensión sobre los números y las operaciones.	10	10,6 %
	Usa estrategias y procedimientos de estimación y cálculo.	0	0 %
	Argumenta afirmaciones sobre las relaciones numéricas y las operaciones.	4	4,3 %
Resuelve problemas de regularidad, equivalencia y cambio	Traduce datos y condiciones a expresiones algebraicas y gráficas.	5	5,3 %
	Comunica su comprensión sobre las relaciones algebraicas.	3	3,2 %
	Usa estrategias y procedimientos para encontrar equivalencias y reglas generales.	9	9,6 %
	Argumenta afirmaciones sobre las relaciones numéricas y las operaciones.	5	5,3 %

Competencia	Capacidades	Cantidad de ítems	Porcentaje
Resuelve problemas de forma, movimiento y localización	Modela objetos con formas geométricas y sus transformaciones.	7	7,4 %
	Comunica su comprensión sobre las formas y relaciones geométricas.	7	7,4 %
	Usa estrategias y procedimientos para orientarse en el espacio.	5	5,3 %
	Argumenta afirmaciones sobre relaciones geométricas.	1	1,0 %
Resuelve problemas de gestión de datos e incertidumbre	Representa datos con gráficos y medidas estadísticas o probabilísticas.	1	1,0 %
	Comunica su comprensión de los conceptos estadísticos y probabilísticos.	9	9,6 %
	Usa estrategias y procedimientos para recopilar y procesar datos.	3	3,2 %
	Sustenta conclusiones o decisiones con base en la información obtenida.	6	6,4 %
Total		94	100 %

*El total ha sido redondeado a un decimal. Por ello, la suma total en esta tabla no es exactamente 100 %.

Tabla A.18 Matemática. Distribución de ítems por contextos evaluados en la prueba

Contextos	Cantidad de ítems	Porcentaje
Intramatemático	7	7,4 %
Extramatemático	87	92,6 %
Total	94	100 %

2.º grado de secundaria

Tabla A.19 Lectura. Distribución de ítems por capacidades evaluadas en la prueba

Capacidades	Cantidad de ítems	Porcentaje
Obtiene información del texto escrito	25	20,5 %
Infiere e interpreta información del texto	58	47,5 %
Reflexiona y evalúa la forma, el contenido y el contexto del texto	39	32,0 %
Total	122	100 %

Tabla A.20 Lectura. Distribución de textos presentes en la prueba por tipo y género textual

Tipo de texto	Género	Correlativo del texto	Cantidad de ítems por cada texto	Porcentaje por tipo de texto
Narrativo	Noticia	1	3	15,0 %
	Cuento	2		
	Crónica	3		
Instructivo	Texto de recomendaciones	4	3	15,0 %
	Manual de procedimientos	5		
		6		
Descriptivo	Artículo enciclopédico	7	3	15,0 %
		8		
	Infografía	9		
Argumentativo	Artículo de opinión	10	6	30,0 %
		11		
		12		
	13			
	Reseña	14		
	Infografía	15		
Expositivo	Artículo de divulgación	16	5	25,0 %
		17		
		18		
	19			
	Afiche	20		
Total			20	100 %

Tabla A.21 Lectura. Distribución de ítems por tipos de textos presentes en la prueba

Tipo de texto	Género	Correlativo del texto	Cantidad de ítems por cada texto	Cantidad de ítems por tipo de texto	Porcentaje por tipo de texto
Narrativo	Noticia	1	6	18	14,8 %
	Cuento	2	6		
	Crónica	3	6		
Instructivo	Texto de recomendaciones	4	6	19	15,6 %
	Manual de procedimientos	5	7		
		6	6		

Tipo de texto	Género	Correlativo del texto	Cantidad de ítems por cada texto	Cantidad de ítems por tipo de texto	Porcentaje por tipo de texto
Descriptivo	Artículo enciclopédico	7	6	18	14,8 %
		8	6		
	Infografía	9	6		
Argumentativo		10	6	36	29,5 %
	Artículo de opinión	11	6		
		12	6		
		13	6		
	Reseña	14	6		
	Infografía	15	6		
Expositivo		16	7	31	25,4 %
	Artículo de divulgación	17	6		
		18	6		
		19	6		
	Afiche	20	6		
Total			122	122	100 %

Tabla A.22 Matemática. Distribución de ítems por competencias evaluadas en la prueba

Competencias	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	34	27,9 %
Resuelve problemas de regularidad, equivalencia y cambio	33	27,0 %
Resuelve problemas de forma, movimiento y localización	28	23,0 %
Resuelve problemas de gestión de datos e incertidumbre	27	22,1 %
Total	122	100,0 %

Tabla A.23 Matemática. Distribución de ítems por capacidades evaluadas en la prueba

Competencia	Capacidades	Cantidad de ítems	Porcentaje
Resuelve problemas de cantidad	Traduce cantidades a expresiones numéricas.	18	14,8 %
	Comunica su comprensión sobre los números y las operaciones.	11	9,0 %
	Usa estrategias y procedimientos de estimación y cálculo.	2	1,6 %
	Argumenta afirmaciones sobre las relaciones numéricas y las operaciones.	3	2,5 %
Resuelve problemas de regularidad, equivalencia y cambio	Traduce datos y condiciones a expresiones algebraicas y gráficas.	12	9,8 %
	Comunica su comprensión sobre las relaciones algebraicas.	6	4,9 %
	Usa estrategias y procedimientos para encontrar equivalencias y reglas generales.	10	8,2 %
	Argumenta afirmaciones sobre las relaciones numéricas y las operaciones.	5	4,1 %
Resuelve problemas de forma, movimiento y localización	Modela objetos con formas geométricas y sus transformaciones.	5	4,1 %
	Comunica su comprensión sobre las formas y relaciones geométricas.	8	6,6 %
	Usa estrategias y procedimientos para orientarse en el espacio.	10	8,2 %
	Argumenta afirmaciones sobre relaciones geométricas.	5	4,1 %
Resuelve problemas de gestión de datos e incertidumbre	Representa datos con gráficos y medidas estadísticas o probabilísticas.	4	3,3 %
	Comunica su comprensión de los conceptos estadísticos y probabilísticos.	11	9,0 %
	Usa estrategias y procedimientos para recopilar y procesar datos.	4	3,3 %
	Sustenta conclusiones o decisiones con base en la información obtenida.	8	6,6 %
Total		122	100,0 %

*El total ha sido redondeado a un decimal. Por ello, la suma total en esta tabla no es exactamente 100 %.

Tabla A.24 Matemática. Distribución de ítems por contextos evaluados en la prueba

Contextos	Cantidad de ítems	Porcentaje
Intramatemático	8	6,6 %
Extramatemático	114	93,4 %
Total	122	100,0 %

Tabla A.25 Ciencia y Tecnología. Distribución de ítems por competencias evaluadas en la prueba

Competencia	Cantidad	%
Indaga mediante métodos científicos para construir sus conocimientos.	38	32,2 %
Explica el mundo físico basándose en conocimientos sobre los seres vivos, materia y energía, biodiversidad, Tierra y universo.	64	54,2 %
Diseña y construye soluciones tecnológicas para resolver problemas de su entorno.	16	13,6 %
Total	118	100,0 %

Tabla A.26 Ciencia y Tecnología. Distribución de ítems por capacidades evaluadas en la prueba

Competencia	Capacidad	Cantidad	%
Indaga mediante métodos científicos para construir sus conocimientos.	Problematiza situaciones para hacer indagación.	14	36,8 %
	Diseña estrategias para hacer indagación.	9	23,7 %
	Genera y registra datos o información.	2	5,3 %
	Analiza datos e información.	12	31,6 %
	Evalúa y comunica el proceso y resultados de su indagación.	1	2,6 %
	Subtotal	38	100 %
Explica el mundo físico basándose en conocimientos sobre los seres vivos, materia y energía, biodiversidad, Tierra y universo.	Comprende y usa conocimientos sobre los seres vivos, materia y energía, biodiversidad, Tierra y universo.	59	92,2 %
	Evalúa las implicancias del saber y del quehacer científico y tecnológico.	5	7,8 %
	Subtotal	64	100 %

Competencia	Capacidad	Cantidad	%
Diseña y construye soluciones tecnológicas para resolver problemas de su entorno.	Determina una alternativa de solución tecnológica.	1	6,3 %
	Diseña la alternativa de solución tecnológica.	9	56,3 %
	Implementa la alternativa de solución tecnológica.	5	31,3 %
	Evalúa y comunica el funcionamiento y los impactos de su alternativa de solución tecnológica.	1	6,3 %
	Subtotal	16	100 %
Total		118	100 %

Tabla A.27 Ciencia y Tecnología. Distribución de ítems por conocimientos evaluados en la prueba

Conocimiento	Cantidad	%
Seres vivos	52	44,1 %
Materia y energía	38	32,2 %
Biodiversidad, Tierra y universo	28	23,7 %
Total	118	100,0 %

Tabla A.28 Ciencia y Tecnología. Distribución de ítems por contextos evaluados en la prueba

Contexto	Cantidad	%
Personal	56	47,5 %
Local/nacional	35	29,7 %
Global	27	22,9 %
Total	118	100,0 %

Anexo B Escalas, fuentes de información, cantidad de ítems y formato de instrumentos HSE

Escalas	Fuente	Autor / año	Subescala	Cantidad de ítems	Formato	Niveles	Alternativas
1. Autorregulación conductual	Cuestionario de Autorregulación (SRQ siglas en inglés)	Pichardo et al. (2014)	-	15	Likert	5	Totalmente en desacuerdo - En desacuerdo – Ni de acuerdo ni en desacuerdo - De acuerdo – Totalmente de acuerdo
2. Autoeficacia	Cuestionario de Autoeficacia para niños (SEQ-C siglas en inglés)	Muris (2001)	<ul style="list-style-type: none"> • Académica • Emocional • Social Total	7 4 6 18	Likert	5	Nada – Un poco – Regular - Mucho - Muchísimo
3. Empatía	Elaboración propia (UMC)	UMC (2019, 2020)	<ul style="list-style-type: none"> • Empatía cognitiva • Disposición empática Total	8 8 16	Likert	5	Totalmente en desacuerdo - En desacuerdo – Ni de acuerdo ni en desacuerdo - De acuerdo – Totalmente de acuerdo
4. Responsabilidad	Cuestionario "Big Five" de la Personalidad para niños (BFQ-C - siglas en inglés).	Barbaranelli et al. (2003)	-	7	Likert	5	Casi nunca – Pocas veces – Algunas veces - Muchas veces Casi siempre
5. Deseabilidad social	Escala de Deseabilidad Social para niños - CSD siglas en inglés	Miller et al. (2004)	-	9			SI - No
6. Tomas de Decisiones	Melbourne Decision Making Questionnaire (DMQ-II)	Mann et al. (1997)	Vigilancia Hipervigilancia Procrastinación Transfancia Total	5 4 4 4 17	Likert	5	Totalmente en desacuerdo - En desacuerdo – Ni de acuerdo ni en desacuerdo - De acuerdo – Totalmente de acuerdo
7. Autonomía y Relación	Cuestionario de autonomía creado por Noom, Dekovic y Meeus	Kagitcibasi (2017)	Autonomía Relación Total	6 6 12	Likert	5	Totalmente en desacuerdo - En desacuerdo – Ni de acuerdo ni en desacuerdo - De acuerdo – Totalmente de acuerdo

Referencias para los instrumentos de HSE

- Barbaranelli, C., Caprara, G., Rabasca, A., y Pastorelli, C. (2003). A questionnaire for measuring the Big Five in late childhood. *Personality and Individual Differences, 34*(4), 645-664.
- Kagjicibasi, C. (2017). *Family, self, and human development across cultures*. Routledge.
- Mann, L., Burnett, P., Radford, M., y Ford, S. (1997). The Melbourne decision making questionnaire: An instrument for measuring patterns for coping with decisional conflict. *Journal of Behavioral Decision Making, 10*(1), 1-19.
- Muris, P. (2001). A brief questionnaire for measuring self-efficacy in youths. *Journal of Psychopathology and Behavioral Assessment, 23*(3), 145-149.
- Pichardo, C., Justicia, F., de la Fuente, J., Martínez-Vicente, J. M., y Berbén, A. B. G. (2014). Factor structure of the self-regulation questionnaire (SRQ) at spanish universities. *Spanish Journal of Psychology, 17*, 1-8.
- Miller, P. H., Baxter, S. D., Hitchcock, D. B., Royer, J. A., Smith, A. F., & Guinn, C. H. (2014). Test-retest reliability of a short form of the Children's Social Desirability Scale for nutrition and health-related research. *Journal of Nutrition Education and Behavior, 46*(5), 423-428.

Ministerio de Educación

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

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