

Agencia de
Calidad de la
Educación



Evaluaciones Nacionales
Estandarizadas **2017**

INFORME TÉCNICO
EVALUACIONES NACIONALES
ESTANDARIZADAS 2017

Esta publicación usa criterios de lenguaje inclusivo tales como núcleos femeninos y no solo masculinos, según sentido particular [ej., madres, padres y apoderados], integración de género en nomenclaturas específicas [ej., director(a)] o empleo del femenino en usos históricamente masculinos [ej., ciudadanía en vez de ciudadanos], entre otros. Sin embargo, para evitar la saturación gráfica y léxica, que dificulta la comprensión y limita la fluidez de lo expresado, y en consonancia con la norma de la Real Academia Española, se usará el masculino sin marcar la oposición de géneros en la mayoría de los nombres y determinantes que el texto provea [ej., los docentes], según su formato e intención comunicativa.

**EVALUACIONES NACIONALES
ESTANDARIZADAS 2017**

Agencia de Calidad de la Educación
División de Evaluación de Logros de Aprendizaje
Morandé 360
Santiago de Chile, 2019
www.agenciaeducacion.cl

ISBN: 978-956-9484-15-5

PRESENTACIÓN

En el marco del Sistema de Aseguramiento de la Calidad (SAC), la Agencia de Calidad de la Educación tiene la misión de evaluar, informar y orientar a la comunidad educativa para contribuir al mejoramiento de la calidad y equidad del sistema educativo nacional.

Una de las herramientas para cumplir con esa labor son las Evaluaciones Nacionales estandarizadas. Dentro de estas, se encuentra el Simce, evaluación externa que tiene como principal propósito contribuir al mejoramiento de la calidad y equidad de la educación informando sobre los logros de aprendizaje de los estudiantes, según el currículo nacional. Además, se incluyen los Estudios Nacionales Muestrales, que entregan información a nivel país sobre el aprendizaje de los estudiantes en áreas y niveles que se consideran relevantes y que no requieren ser evaluados de manera censal. Por último, la información obtenida respecto de los aprendizajes se complementa con la aplicación de los Cuestionarios de Calidad y Contexto, que entregan información sobre el contexto escolar y social en el que aprenden a diario los estudiantes.

Los informes técnicos comenzaron a publicarse el 2013, como parte del esfuerzo por informar a la comunidad técnica y académica de una manera transparente y efectiva. Estos informes describen el proceso de construcción, distribución, envío, aplicación, procesamiento, análisis y comunicación de resultados de las pruebas Simce aplicadas el 2012.

Dejar registro de los procesos internos, procedimientos y contenidos disponibles, promueve una cultura de transparencia y compromiso con el aprendizaje continuo. Desde la perspectiva interna, elaborar este informe implica una oportunidad de revisión y sistematización de la información de los macroprocesos. Para la comunidad en general: investigadores, académicos, profesionales de la educación y apoderados, provee una oportunidad para interpretar, participar y colaborar en el avance hacia una educación de calidad que promueva la equidad de oportunidades para un desarrollo pleno de los estudiantes.

El presente informe se divide en dos partes. En la primera se entrega información sobre las evaluaciones censales Simce y en la segunda, sobre los estudios muestrales nacionales. Esperamos que este informe sea de utilidad para la comunidad técnica interesada en la medición de la calidad de la educación, y permita a la Agencia consolidar y validar el rol de la evaluación externa en el sistema educativo.

Contenido

EVALUACIONES SIMCE

Introducción	13
Capítulo I: Diseño y construcción de las pruebas Simce 2017	17
1.1 Definición del marco de referencia y formulación de especificaciones técnicas.....	17
1.2 Elaboración y validación de las preguntas.....	25
1.3 Ensamblaje y diagramación de las pruebas.....	26
Capítulo II: Logística y aplicación de las pruebas Simce 2017	31
2.1 Empadronamiento de establecimientos y estudiantes.....	32
2.2 Impresión, mecanizado y distribución del material de aplicación	34
2.3 Aplicación de las pruebas.....	35
2.4 Entrega del material a procesamiento de datos.....	42
2.5 Atención a la comunidad escolar.....	42
Capítulo III: Procesamiento y análisis de datos 2017	45
3.1 Captura, validación y consolidación de la base de datos de pruebas y cuestionarios.....	45
3.2 Corrección de preguntas de respuesta abierta.....	50
3.3 Análisis psicométricos de las pruebas.....	52
3.4 Análisis de datos agregados.....	60
Capítulo IV: Comunicación de resultados educativos 2017	65
4.1 Informes de resultados.....	66
4.2 Otros documentos comunicacionales.....	69
4.3 Síntesis de productos comunicacionales Simce 2017	71

ESTUDIOS NACIONALES MUESTRALES

Introducción	75
Capítulo I: Visión panorámica de los Estudios Nacionales Muestrales 2017	79
Capítulo II: Descripción de los Estudios Nacionales Muestrales 2017	85
1.1 Lectura 2.º básico.....	85
1.2 Formación Ciudadana 8.º básico	87
1.3 Idioma Extranjero Inglés III medio.....	95
Referencias bibliográficas	103
Anexos	105

Tablas

Tabla 1.1.1: Matriz de evaluación de las pruebas Simce Lectura aplicadas el 2017. Ejes de habilidades.....	18
Tabla 1.1.2: Matriz de evaluación de las pruebas Simce Matemática aplicadas el 2017. Ejes de habilidades	19
Tabla 1.1.3: Matriz de evaluación de la prueba Simce Matemática aplicada el 2017. Nivel 4° básico. Ejes temáticos.....	20
Tabla 1.1.4: Matriz de evaluación de la prueba Simce Matemática aplicada el 2017. Nivel 8° básico. Ejes temáticos.....	21
Tabla 1.1.5: Matriz de evaluación de la prueba Simce Matemática aplicada el 2017. Nivel II medio. Ejes temáticos.....	21
Tabla 1.1.6: Matriz de evaluación de la prueba Simce Historia, Geografía y Ciencias Sociales aplicada el 2017. Nivel II medio. Ejes de habilidades.....	22
Tabla 1.1.7: Matriz de evaluación de la prueba Simce Historia, Geografía y Ciencias Sociales aplicada el 2017. Nivel II medio. Ejes temáticos.....	23
Tabla 1.1.8: Matriz de evaluación de la prueba Simce Ciencias Naturales aplicada el 2017. Nivel 8.º básico. Ejes de habilidades.....	24
Tabla 1.1.9: Matriz de evaluación de la prueba Simce Ciencias Naturales aplicada el 2017. Nivel 8.º básico. Ejes temáticos.....	25
Tabla 1.1.10: Formas aplicadas en las pruebas Simce censales 2017 según asignatura y nivel.....	27
Tabla 1.2.1: Categorías de establecimientos según códigos SIGE evaluados con las pruebas Simce 2017	32
Tabla 1.2.2: Establecimientos y estudiantes evaluados en la aplicación censal de las pruebas Simce 2017	36
Tabla 1.3.1: Códigos de la variable de identificación por nivel escolar de la aplicación 2017	46
Tabla 1.3.2: Categorías de la variable de participación de estudiantes en la aplicación 2017	46
Tabla 1.3.3: Tipos de documentos capturados en 2017 por nivel escolar y prueba aplicada.....	48
Tabla 1.3.4: Distribución de correctores de preguntas abiertas pruebas censales Simce 2017	51
Tabla 1.3.5: Cantidad de preguntas abiertas pruebas censales Simce 2017	51
Tabla 1.3.6: Indicadores 2017 para preguntas de respuesta cerrada según Teoría Clásica e IRT.....	53

Tabla 1.3.7: Indicadores 2017 para preguntas de respuesta abierta según Teoría Clásica e IRT.....	54
Tabla 1.3.8: Puntajes de corte para la clasificación 2017 según Estándares de Aprendizaje.....	58
Tabla 1.3.9: Agrupación de análisis agregados de las pruebas Simce 2017.....	61
Tabla 1.4.1: Documentos comunicacionales de la aplicación Simce 2017.....	71
Tabla 2.1.1: Establecimientos y estudiantes evaluados en la aplicación de los Estudios Nacionales 2017.....	79
Tabla 2.1.2: Tipos de documentos capturados en 2017 por nivel escolar y prueba muestral aplicada.....	80
Tabla 2.1.3: Cantidad de preguntas abiertas pruebas muestrales 2017.....	81
Tabla 2.1.4: Distribución de correctores de preguntas abiertas pruebas muestrales 2017.....	81
Tabla 2.2.1: Matriz de evaluación del Estudio Nacional de Lectura, 2.º básico ejes de habilidades.....	85
Tabla 2.2.2: Puntajes de corte para la clasificación 2017 según Estándares de Aprendizaje Nivel 2.º básico.....	87
Tabla 2.2.3: Dominio de contenidos. Temas de la evaluación muestral Formación Ciudadana 2017.....	90
Tabla 2.2.4: Clasificación de actitudes. Temas de la evaluación muestral Formación Ciudadana 2017.....	93
Tabla 2.2.5: Tipología del dominio afectivo-comportamental Compromiso.....	94
Tabla 2.2.6: Habilidades de comprensión lectora y comprensión auditiva. Estudio Nacional de Inglés 2017.....	93
Tabla A1: Estudiantes según nivel escolar en las PE2017 y porcentaje de la muestra teórica que representan.....	106
Tabla A2: Código y descripción de los campos de identificación de los ítems de las PE2017.....	108
Tabla A3: Código y descripción de los campos de datos métricos de los ítems de las PE2017.....	109
Figuras	
Figura 1.3.1: Ejemplo de curva empírica del ítem.....	59



**Evaluaciones
Simce**

Introducción

Las pruebas Simce tienen carácter censal y evalúan el aprendizaje del currículo nacional vigente de todos los estudiantes del país, en diferentes asignaturas y niveles escolares, con el propósito de contribuir al proceso de mejoramiento de la calidad y equidad de la educación.

La información obtenida a través de los resultados de estas pruebas se pone a disposición de los establecimientos con el objetivo de que analicen y reflexionen sobre los aprendizajes alcanzados por sus estudiantes. Para ello, se informa tanto el puntaje promedio obtenido, como la distribución de estudiantes en cada nivel de los Estándares de Aprendizaje definidos por el Ministerio de Educación¹, de manera que sea posible identificar desafíos y fortalezas que contribuyan al mejoramiento continuo de los aprendizajes.

De acuerdo con el Plan de Evaluaciones Nacionales e Internacionales² 2016–2020, en el año 2017 se aplicaron un total de ocho pruebas censales, comprendiendo los niveles de 4.º y 8.º de Educación Básica y II de Educación Media. Estas abordaron los contenidos curriculares de las asignaturas de Lenguaje y Comunicación: Lectura³, Matemática, Ciencias Naturales e Historia, Geografía y Ciencias Sociales, tal como se muestra en el siguiente cuadro:

Pruebas Simce censales aplicadas en 2017

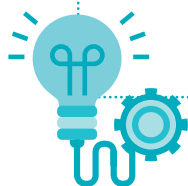
Pruebas Simce / Niveles evaluados	4.º básico	8.º básico	II medio
Lectura	✓	✓	✓
Matemática	✓	✓	✓
Historia, Geografía y Ciencias Sociales	✓	✓	✓
Ciencias Naturales	✓	✓	✓

¹ La información sobre el proceso de elaboración de los Estándares de Aprendizaje se encuentra en los documentos con los fundamentos de los respectivos estándares en <https://www.curriculumnacional.cl/portal/Documentos-Curriculares/Estandares-e-indicadores-de-calidad/>

² El Plan de Evaluaciones Nacionales e Internacionales define un programa que permite medir el grado de cumplimiento de los Objetivos de Aprendizaje definidos en las Bases Curriculares de Educación Básica y Media. En él se consideran tanto las evaluaciones nacionales como las internacionales. El Plan detalla, como mínimo, las áreas curriculares que son objeto de evaluación, los grados de Educación Básica y Media que son medidos, la periodicidad de la evaluación, y las principales desagregaciones y modos de informar los resultados.

³ Lengua y Literatura de acuerdo a las Bases Curriculares vigentes a la fecha de aplicación en 8.º básico.

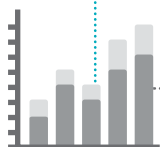
El ciclo evaluativo de las pruebas Simce 2017 constó de cuatro etapas que dan origen a los capítulos de esta sección. La primera etapa comprende el diseño y construcción de los instrumentos, incluyendo la definición del marco de referencia y la formulación de especificaciones técnicas. Asimismo, supone la elaboración y validación de las preguntas; la selección definitiva de los ítems que formarían parte de la aplicación censal; los procesos requeridos para establecer comparabilidad entre las mediciones de distintos años; y el ensamblaje y diagramación de los cuadernillos. En el segundo capítulo se describe la etapa que contempla las operaciones de campo, incluyendo los procesos de empadronamiento, imprenta, distribución y aplicación de las evaluaciones. En el tercer capítulo se aborda el procesamiento y análisis de los datos, desde la captura de las respuestas hasta los análisis psicométricos y agregados que se realizan con los resultados. Finalmente, en el cuarto capítulo se describe el proceso de comunicación de resultados de las pruebas.



Capítulo I:
Diseño y construcción de las
pruebas Simce 2017



Capítulo II:
Logística y aplicación de las
pruebas Simce 2017



Capítulo III:
Procesamiento y análisis
de datos 2017



Capítulo IV:
Comunicación de resultados
educativos 2017

Capítulo I: Diseño y construcción de las pruebas Simce 2017

La construcción de las pruebas Simce censales se lleva a cabo mediante un proceso que dura aproximadamente un año y medio. En él se resguarda el cumplimiento de estándares internacionales, como los definidos en el documento *Standards for Educational and Psychological Testing* (AERA, APA y NCME, 2014).

Este proceso se inicia con la definición del marco de referencia y las especificaciones técnicas, continúa con la elaboración y validación de las preguntas, y culmina con el armado y diagramación de las formas definitivas de las pruebas que se aplican a los estudiantes.

En este capítulo se describe cada una de las etapas de construcción de las pruebas censales Simce 2017.

1.1 Definición del marco de referencia y formulación de especificaciones técnicas

El primer paso en la construcción de las pruebas Simce consiste en definir el marco de referencia; es decir, los documentos que delimitan los constructos a evaluar y que, de acuerdo con los estándares internacionales (AERA, APA y NCME, 2014; Case, Jorgensen y Zucker, 2004), otorgan validez a las evaluaciones.

Considerando que las pruebas Simce son referidas al Currículum Nacional, el marco de referencia está dado por este. En las pruebas 2017, el marco de referencia lo constituyen las Bases Curriculares, establecidas mediante los decretos N.º 439/2012 (Mineduc, 2011) para 4.º básico y N.º 614/2013 (Mineduc, 2013) para 8.º básico⁴. Para las pruebas de II medio, el marco de referencia es el marco curricular ajustado de 2009, derivado de los decretos supremos de educación, números 254 y 256 (Mineduc, 2009b).

A partir de los marcos de referencia, se generan las especificaciones técnicas de las pruebas. Estas establecen las características específicas de diseño de los instrumentos y buscan garantizar cobertura y comparabilidad en la medición de los aprendizajes curriculares. Las especificaciones técnicas se elaboran por los equipos de la Agencia de Calidad con la colaboración de la Unidad de Currículum y Evaluación (UCE) del Ministerio de Educación.

Las especificaciones técnicas de las pruebas distinguen y definen ejes temáticos (o de contenidos) y ejes de habilidades (o de dominios cognitivos) y, además, establecen las matrices de evaluación teóricas que orientan la construcción de las pruebas, señalando los porcentajes de preguntas que se consideran deseables para representar cada uno de los ejes en la prueba. Estos porcentajes se definen a partir de un análisis de énfasis curriculares que realizan en conjunto profesionales de la Agencia y de la UCE.

⁴ El 2017 corresponde al primer año de vigencia de las Bases Curriculares de 8.º básico.

A continuación, se describen las evaluaciones Simce aplicadas el 2017, las que fueron construidas sobre la base de las especificaciones técnicas⁵.

1.1.1 Pruebas de Lectura

En todos los niveles evaluados en las pruebas de Lectura 2017 se distinguen tres ejes de habilidades: Localizar, Interpretar y relacionar, y Reflexionar, cuya matriz de evaluación se muestra en la Tabla 1.1.1.

El eje **Localizar** agrupa las tareas de lectura que el estudiante debe emplear para operar con elementos explícitos de los textos, tales como identificar, discriminar y extraer información explícita de las lecturas.

El eje **Interpretar y relacionar** incluye las tareas de lectura relativas a construir significados implícitos estableciendo conexiones entre elementos explícitos. En este eje se encuentran las habilidades de inferir información, interpretar lenguaje figurado y reconocer relaciones causales, entre otras.

El eje **Reflexionar** implica las tareas de lectura relativas a la capacidad de confrontar distintos aspectos del texto (formales o de contenido) con conocimientos previos, experiencia personal, conocimiento de mundo y lecturas anteriores, permitiendo al lector examinar el texto y evaluarlo críticamente. En este eje se incluye, por ejemplo, la capacidad de emitir opiniones sobre lo leído y de determinar el aporte de elementos gráficos al sentido de un texto.

Los ejes antes señalados se evalúan a partir de textos literarios y no literarios. Los primeros se entienden como aquellos que tienen una intención estética de base y constituyen un testimonio de la experiencia humana, que abre un espacio a la reflexión de aspectos personales, valóricos y sociales; los segundos se definen como aquellos que tienen una intención comunicativa focalizada en la conexión entre el autor y el lector, es decir, en la transmisión de un mensaje determinado. Adicionalmente, en cada uno de estos tipos de textos se distinguen subcategorías: en los textos literarios se incluyen textos narrativos, líricos y dramáticos; y en los no literarios se distinguen textos informativos, persuasivos e instruccionales. Los textos informativos, a su vez, se presentan en tres tipos de formatos: continuo (formado por oraciones que se organizan en párrafos), discontinuo (formado por una serie de listas, gráficos o diagramas, entre otros), y mixto (conformado por textos de formato continuo y discontinuo).

Tabla 1.1.1

Matriz de evaluación de las pruebas Simce Lectura aplicadas el 2017. Ejes de habilidades.

Ejes de habilidades	4.° básico	8.° básico	II Medio
Localizar	32 %	23 %	18 %
Interpretar y relacionar	52 %	52 %	58 %
Reflexionar	16 %	25 %	23 %

⁵ En esta descripción se muestran las matrices finales de las evaluaciones, es decir, las resultantes luego del análisis psicométrico.

1.1.2 Pruebas de Matemática

En las pruebas Simce de Matemática 2017 de todos los niveles evaluados, se distinguen tres ejes de habilidades: Conocimiento, Aplicación y Razonamiento, cuya matriz de evaluación se muestra en la Tabla 1.1.2.

El eje **Conocimiento** comprende el conocimiento factual de conceptos, propiedades y procedimientos matemáticos, así como sus representaciones simbólicas. En este eje se incluyen, entre otras habilidades, las de recordar definiciones, fórmulas y teoremas, métodos de cálculo y procedimientos de resolución de ecuaciones; reconocer propiedades; localizar datos; y clasificar elementos según propiedades dadas.

El eje **Aplicación** reúne las habilidades relativas a la selección de estrategias adecuadas para resolver un problema y la aplicación de operaciones, procedimientos, etc. Supone, asimismo, la capacidad de generar representaciones equivalentes para comunicar cantidades, operaciones y relaciones dadas, y de representar e interpretar información y datos en diagramas, tablas y gráficos.

El eje **Razonamiento** aborda el razonamiento intuitivo e inductivo necesario para descubrir patrones y regularidades, para realizar generalizaciones y para buscar estrategias, representaciones y modelos que permitan resolver problemas no rutinarios. Supone también la habilidad de realizar deducciones basadas en reglas y supuestos específicos, de diferenciar afirmaciones erróneas de verdaderas, de argumentar la validez de un resultado obtenido y de comunicar ideas y conclusiones.

Tabla 1.1.2

Matriz de evaluación de las pruebas Simce Matemática aplicadas el 2017. Ejes de habilidades.

Ejes de habilidades	4.º básico	8.º básico	II Medio
Conocimiento	49 %	46 %	33 %
Aplicación	39 %	36 %	53 %
Razonamiento	12 %	18 %	14 %

En lo que respecta a los ejes temáticos, estos se desprenden de los ejes presentes en las Bases Curriculares vigentes. Así, en las pruebas de Matemática de 4.º básico, cuya matriz de evaluación se muestra en la Tabla 1.1.3, se distinguen cinco ejes temáticos: Números y operaciones, Patrones y álgebra, Geometría, Medición, y Datos y probabilidades.

El eje **Números y operaciones** reúne conceptos básicos de números y algoritmos, tales como el sistema posicional, los algoritmos de la adición, sustracción, multiplicación y división de números naturales, y las propiedades, operaciones y establecimiento de conexiones entre Números y operaciones; Patrones y álgebra; Geometría; Medición; y Datos y probabilidades.

El eje **Patrones y álgebra** aborda las relaciones entre números, formas, objetos y conceptos; el cambio de una cantidad en relación con otra; las ecuaciones e inecuaciones; las reglas que generan los patrones; el transporte de los patrones de una forma de representación a otra; y el modelado de situaciones sencillas que impliquen las cuatro operaciones.

El eje **Geometría** abarca el reconocimiento, la visualización y la comparación de figuras de dos y tres dimensiones; el conocimiento de sus características y propiedades; y el reconocimiento y la descripción de su movimiento a través de la reflexión, la traslación y la rotación. Incluye asimismo conceptos relativos a la estructura del espacio, los elementos del entorno y los sistemas de coordenadas informales.

El eje **Medición** reúne la medición del tiempo y cuantificación y medición de objetos (ancho, largo, alto, peso, perímetro, superficie, volumen, ángulos, etc.), con unidades estandarizadas y no estandarizadas. Supone también la discriminación de las unidades de medición apropiadas al contexto y el reconocimiento de herramientas específicas de medición, en virtud del objeto a ser medido.

El eje **Datos y probabilidades** abarca el registro, clasificación y lectura de información dispuesta en tablas y gráficos presentes en la vida cotidiana, así como la construcción de tablas y gráficos, y la formulación de conclusiones a partir de los datos. Además, el eje incorpora cálculo e interpretación de la media aritmética y conceptos básicos sobre probabilidades, como la capacidad de describir la posibilidad de ocurrencia de un evento aleatorio y la comparación de probabilidades de distintos eventos, sin cálculo.

Tabla 1.1.3

Matriz de evaluación de la prueba Simce Matemática aplicada el 2017. Nivel 4.º básico, ejes temáticos.

Ejes temáticos	4.º básico
Números y operaciones	50 %
Patrones y álgebra	7 %
Geometría	16 %
Medición	15 %
Datos y probabilidades	12 %

Por su parte, en la prueba de Matemática de 8.º básico, cuya matriz de evaluación se muestra en la Tabla 1.1.4, se distinguen cuatro ejes temáticos: Números; Álgebra y funciones; Geometría; y Probabilidad y estadística.

El eje **Números** incluye números enteros y racionales y las operaciones entre ellos; el trabajo con potencias de base natural y su uso en la notación científica; el concepto de porcentaje, raíces cuadradas de números naturales y la resolución de problemas que involucran variaciones porcentuales.

El eje **Álgebra y funciones** abarca la comprensión del lenguaje algebraico; la escritura, representación y uso de expresiones algebraicas; reducción y operatoria con las mismas; proporciones directas e inversas; inecuaciones, ecuaciones lineales, funciones (lineal y afín) y modelamiento usando lenguaje algebraico.

El eje **Geometría** abarca la imaginación espacial; el conocimiento de las propiedades y características de diferentes figuras geométricas; la identificación de puntos en el plano cartesiano; área, perímetro y diámetro de círculo; área de triángulos, paralelogramos y trapecios, prismas rectos y cilindros; volumen de prismas rectos y cilindros y el uso del teorema de Pitágoras en la resolución de problemas. También las traslaciones, rotaciones y reflexiones, y la composición de rotaciones, traslaciones y reflexiones en el plano cartesiano.

El eje **Probabilidad y estadística** abarca el trabajo a partir de datos estadísticos, incluyendo las posibles interpretaciones erróneas de un gráfico y de las posibles manipulaciones que se pueden hacer con los datos. También la estimación de porcentajes; representación de datos; medidas de tendencia central y rango; probabilidades; comparación de frecuencias relativas. Además, medidas de posición, percentiles y cuartiles; y principio combinatorio multiplicativo.

Tabla 1.1.4

Matriz de evaluación de la prueba Simce Matemática aplicada el 2017. Nivel 8.º básico, ejes temáticos.

Ejes temáticos	8.º básico
Números	30 %
Geometría	23 %
Álgebra y funciones	29 %
Probabilidad y estadística	18 %

Por último, en las pruebas de Matemática de II medio, cuya matriz de evaluación se muestra en la Tabla 1.1.5, se distinguen cuatro ejes temáticos: Números; Geometría; Álgebra y funciones; y Datos y azar.

El eje **Números** agrupa cantidades y números; operaciones aritméticas (adición, sustracción, multiplicación y división); potencias; raíces y propiedades; y relaciones de los conjuntos numéricos (números naturales, enteros, racionales y reales).

El eje **Geometría** abarca la imaginación espacial; el conocimiento de objetos, figuras y cuerpos geométricos básicos y sus propiedades; las relaciones entre formas de dos y tres dimensiones; la construcción y transformación de figuras; la medición de figuras planas y cuerpos; y las transformaciones isométricas, los vectores y la geometría cartesiana.

El eje **Álgebra y funciones** reúne símbolos de representación y operación de cantidades; expresiones de relaciones generales y abstractas de aritmética y medición; lenguaje matemático abstracto; ecuaciones y sistemas de ecuaciones; y el concepto de función, tipos de funciones y su representación gráfica.

El eje **Datos y azar** aborda diversos tipos de tablas y gráficos; los conceptos de muestra y población; el análisis de datos, utilizando herramientas como las medidas de tendencia central, de posición y dispersión, y la teoría de probabilidades, con todos los conceptos asociados a ella como, por ejemplo, espacio muestral, Modelo de Laplace, principios aditivo y multiplicativo, y la Ley de los Grandes Números.

Tabla 1.1.5

Matriz de evaluación de la prueba Simce Matemática aplicada el 2017. Nivel II medio, ejes temáticos.

Ejes temáticos	II medio
Números	27 %
Geometría	23 %
Álgebra y funciones	29 %
Datos y azar	21 %

1.1.3 Prueba de Historia, Geografía y Ciencias Sociales

En la prueba de Historia, Geografía y Ciencias Sociales de II medio 2017, cuya matriz de evaluación se muestra en la Tabla 1.1.6, se distinguen tres ejes de habilidades: Conocimiento y comprensión, Aplicación y Razonamiento.

El eje **Conocimiento y comprensión** implica la capacidad del estudiante de recordar, definir, describir y comprender conceptos, hechos, fenómenos y procesos históricos, geográficos y sociales; así como de extraer información de diversas fuentes propias de las disciplinas sociales.

El eje **Aplicación** involucra las habilidades relativas al uso de conceptos y conocimientos históricos, geográficos y sociales, para situar elementos en el tiempo y en el espacio, clasificar y secuenciar información o eventos y transferirlos a situaciones nuevas. Asimismo, supone la capacidad de interpretar información presentada en diversas fuentes.

El eje **Razonamiento** apunta a diversas tareas de pensamiento crítico respecto de los procesos y fenómenos de la realidad histórica, geográfica y social. Supone, entre otras habilidades, las de comparar, analizar y evaluar información referida a las Ciencias Sociales, y las de utilizar y problematizar conceptos y situaciones.

Tabla 1.1.6

Matriz de evaluación de la prueba Simce Historia, Geografía y Ciencias Sociales aplicada el 2017. Nivel II medio, ejes de habilidades.

Ejes temáticos	II medio
Conocimientos y Comprensión	54 %
Aplicación	16 %
Razonamiento	30 %

En lo que respecta a los ejes temáticos de la prueba Simce de Historia, Geografía y Ciencias Sociales 2017, cuya matriz de evaluación se muestra en la Tabla 1.1.7, se distinguen: Sociedad en perspectiva histórica, Espacio geográfico, y Democracia y desarrollo.

El eje **Sociedad en perspectiva histórica** incluye contenidos referidos a la historia mundial del siglo XX y a la historia de Chile, desde el periodo colonial hasta fines del siglo XIX. La aproximación a los diversos procesos históricos se orienta hacia la distinción de características propias de las distintas etapas y procesos, y hacia su comprensión desde diversas dimensiones e interpretaciones, propiciando una visión de conjunto (no fragmentada) del desarrollo histórico. En particular, los conocimientos sobre historia mundial del siglo XX incluyen las guerras mundiales, el periodo de entreguerras y el impacto de las guerras en el ordenamiento mundial y en la creciente valoración de la paz y los derechos humanos. En relación con la historia de Chile, se incluye una visión panorámica, distinguiendo elementos de continuidad y cambio, abordando asimismo temas como la permanencia de las culturas indígenas, el proceso de mestizaje, el legado colonial y la conformación del estado-nación y su evolución política, social, económica, cultural y territorial a lo largo del siglo XIX.

El eje **Espacio geográfico** abarca los contenidos referidos a los procesos de urbanización del siglo XX, considerando migración campo-ciudad (efecto en la composición urbano-rural en distintas regiones a nivel mundial); ventajas y desventajas de la vida en las ciudades; y geografía de la población mundial (diversidad cultural, dinámica y estructura, y problemas e impacto del desarrollo tecnológico, entre otros).

El eje **Democracia y desarrollo** incluye contenidos referidos al proceso de globalización e internacionalización de la economía en el mundo actual, considerando sus principales características, sus falencias (como las desigualdades o contrastes que este proceso genera), sus logros y su impacto en la valoración de los derechos humanos. También aborda las características de la llamada sociedad de la información, esto es: interconectividad global; nuevas tecnologías de información y comunicación; problemas y desafíos provocados por la tensión entre la difusión de patrones culturales a nivel planetario, y permanencia de las identidades locales.

Tabla 1.1.7

Matriz de evaluación de la prueba Simce Historia, Geografía y Ciencias Sociales aplicada el 2017. Nivel II medio, ejes temáticos.

Ejes temáticos	II medio
Sociedad en perspectiva histórica	66 %
Espacio geográfico	18 %
Democracia y desarrollo	16 %

1.1.4 Pruebas de Ciencias Naturales

En la prueba de Ciencias Naturales de 8.º básico aplicada en 2017, cuya matriz de evaluación se muestra en la Tabla 1.1.8, se distinguen tres ejes de habilidades: Conocimiento, Aplicación y Razonamiento.

El eje **Conocimiento** incluye habilidades referidas a la comprensión de hechos, relaciones, conceptos y procesos propios de las Ciencias Naturales, tales como recordar, reconocer y ejemplificar afirmaciones, hechos y conceptos, y dominar vocabulario científico (símbolos y unidades).

El eje **Aplicación** supone la aplicación del conocimiento científico en situaciones sencillas, lo que implica habilidades como comparar, contrastar, clasificar e interpretar información relacionada con algún concepto o principio de la ciencia, así como organizar información para comunicarla, y explicar y relacionar información obtenida de diagramas o modelos.

El eje **Razonamiento** refiere al uso del razonamiento científico para resolver problemas, extraer conclusiones, tomar decisiones fundamentadas y abordar situaciones nuevas.

Tabla 1.1.8

Matriz de evaluación de la prueba Simce Ciencias Naturales aplicada el 2017. Nivel 8.º básico, ejes de habilidades.

Ejes de habilidades	8.º básico
Conocimiento	32 %
Aplicación	38 %
Razonamiento	30 %

Sobre los ejes temáticos de la prueba Simce Ciencias Naturales 2017, cuya matriz de evaluación se muestra en la Tabla 1.1.9, cabe señalar que en 8.º básico se presenta Ciencias Naturales como una sola asignatura que organiza sus Objetivos de Aprendizaje en tres ejes representativos de disciplinas científicas: Biología, Física y Química.

La organización en estos ejes se centra en mostrar al estudiante la distinción y particularidad de cada una de las disciplinas en forma concreta, consecuente con la propuesta de algunas grandes ideas de la ciencia, permitiendo una mayor apropiación de los distintos contenidos que aborda cada una, y proyectando una visión científica holística de la realidad al estudiante.

De acuerdo con lo anterior, en el eje de **Biología** se espera que los estudiantes logren desarrollar habilidades y conocimientos científicos sobre el cuerpo humano, sus estructuras, y los procesos relacionados con su ciclo de vida y su adecuado funcionamiento. Asimismo, se realiza un estudio profundo sobre las células, relacionándolas con la diversidad y la evolución genética transmitida de una generación a otra.

En el eje de **Física**, se tratan temas generales de astronomía, algunos aspectos básicos de geofísica y de clima y tiempo atmosférico, con foco en el comportamiento dinámico de la Tierra, desarrollando así conductas de prevención frente a fenómenos físicos a gran escala. Por otro lado, se busca que los estudiantes comprendan el origen y evolución del Universo, desde los primeros modelos que han explicado su forma y dinámica, hasta el conocimiento científico que hoy en día se tiene de este.

Se espera, además, que logren comprender las leyes físicas de movimiento planteadas por Isaac Newton, así como la conservación de la energía, y el uso responsable de la energía como recurso energético no renovable. Finalmente, el eje pretende que los alumnos puedan explicar fenómenos auditivos y luminosos que se perciben a través de la audición y visión humana, mediante la descripción del modelo ondulatorio.

En el eje de **Química** se espera desarrollar una conceptualización sobre el hecho de que toda la materia del Universo está compuesta por átomos, sus características y cómo estas partículas interactúan formando nuevas sustancias. Asimismo, se espera una comprensión del desarrollo de conocimiento científico relacionado con el mundo microscópico y la interacción de la materia inerte, los seres vivos y el entorno, analizando el comportamiento de los diferentes estados de la materia (sólido, líquido y gaseoso), comprendiendo a la vez los aspectos químicos asociados al desarrollo de la vida en la Tierra, para así poder relacionar el aspecto abiótico y biótico con su composición atómica y molecular. Finalmente, este eje evidencia cómo el conocimiento científico-químico contribuye a mejorar la calidad de vida de las personas y cómo aportar desde este ámbito al cuidado responsable del medio.

Tabla 1.1.9

Matriz de evaluación de la prueba Simce Ciencias Naturales aplicada el 2017. Nivel 8.º básico, ejes temáticos.

Ejes temáticos	8.º básico
Biología	40 %
Química	28 %
Física	32 %

1.2 Elaboración y validación de las preguntas

Las preguntas de las pruebas Simce se construyen a partir de Indicadores de Evaluación (IE) que corresponden a una operacionalización de los ejes descritos en el apartado anterior.

En las pruebas Simce se incluyen dos formatos de preguntas: de opción múltiple (“preguntas cerradas”) y de respuesta abierta (“preguntas abiertas”). La elección de estos formatos está determinada por el contenido y el comportamiento cognitivo que se busca evaluar (Haladyna y Rodríguez, 2013).

Las pruebas Simce se construyen mayoritariamente sobre la base de preguntas de opción múltiple, pues este formato permite reportar información de la mayoría de los constructos a evaluar en forma efectiva y eficiente, asegurando validez, confiabilidad y objetividad del instrumento en su totalidad (Downing y Haladyna, 2006). Las preguntas de respuesta abierta se utilizan fundamentalmente para evaluar aquellos aprendizajes que no son susceptibles de ser medidos con preguntas de respuesta cerrada y que, en términos generales, corresponden a habilidades cognitivas de orden superior, para las que el formato abierto permite mayor fidelidad al constructo medido (Haladyna y Rodríguez, 2013).

Las preguntas cerradas se construyen a partir de un enunciado para el cual existe una única opción correcta (clave) y dos o tres distractores (opciones incorrectas), dependiendo de la variedad de errores plausibles que sea posible recoger. Los distractores se elaboran sobre la base de errores comunes que cometen los estudiantes que no han logrado los aprendizajes (Haladyna y Rodríguez, 2013). En algunos casos, las preguntas se encuentran contextualizadas con información que sirve de base para la selección de la respuesta, tales como textos, gráficos, situaciones problemáticas, etc.

Las preguntas de respuesta abierta, por su parte, se construyen a partir de enunciados (“estímulos”) que plantean una tarea o interrogante que debe ser desarrollada por el estudiante. Al igual que en las preguntas cerradas, algunas preguntas abiertas también se encuentran contextualizadas con información de base para el desarrollo de la tarea. Las pautas de este formato de preguntas son de dos tipos: algunas corresponden a “pautas de respuesta única”, es decir, documentos que establecen la o las respuestas correctas posibles, y otras corresponden a rúbricas, esto es, documentos que establecen descripciones de desempeño, dispuestas en una escala, las que se elaboran sobre la base de aspectos que se observarán en la ejecución de una tarea o producto.

El proceso de elaboración de las preguntas de las pruebas Simce 2017 se llevó a cabo mediante una modalidad mixta: por licitación durante el año 2015 y elaboración interna durante el año 2016. Al finalizar el proceso de elaboración, se realizaron paneles de revisión de expertos que, mediante criterios de elaboración, revisaron cada ítem y fueron verificando que cumpla con los criterios de elaboración, los cuales están alineados con el Diseño Universal⁶. Posteriormente, los equipos de la UCE revisaron estos ítems para validar su alineamiento curricular.

Una vez terminados los procesos anteriores, los ítems fueron distribuidos en cuadernillos para su aplicación experimental conducente a su validación.

Las pruebas experimentales fueron aplicadas el año anterior a la aplicación de la prueba Simce definitiva a una muestra representativa de la población evaluable. El análisis de estas pruebas permite obtener información cuantitativa y cualitativa de los ítems para determinar si poseen la suficiente calidad para conformar una prueba Simce definitiva. En el caso de las preguntas abiertas, la información permite validar las pautas y rúbricas, y elaborar los manuales de corrección. En el caso de las preguntas de opción múltiple, el proceso permite obtener información sobre la calidad de los enunciados y las opciones, así como la dificultad que estas tienen.

Tomando como base el resultado de los análisis cuantitativos y cualitativos de las pruebas experimentales, se selecciona un conjunto de preguntas que, sumadas a otras provenientes de aplicaciones previas, conforman las pruebas definitivas de aplicación censal. Todas estas preguntas deben cumplir con los requisitos de representatividad y cobertura curricular, cobertura del rango de habilidades de los estudiantes y alineación a los Estándares de Aprendizaje⁷, en los niveles en que estos existan.

1.3 Ensamblaje y diagramación de las pruebas

El diseño de las pruebas Simce definitivas de aplicación censal busca obtener información de los aprendizajes alcanzados por todos los estudiantes, de acuerdo con el currículo vigente. Por lo anterior, la prueba está compuesta por varias formas equivalentes (cuadernillos), las que en su conjunto cubren todos los ejes de habilidades y temáticos, representando una amplia gama de los Objetivos de Aprendizaje del Currículum Nacional.

En la Tabla 1.1.10 se muestra la cantidad de formas que posee cada una de las pruebas Simce aplicadas el 2017.

⁶ El diseño universal es un enfoque hacia el diseño de pruebas que busca maximizar la accesibilidad para todos los examinandos previstos.

El diseño universal, requiere que los desarrolladores de la prueba sean claros sobre los constructos sometidos a medición, incluyendo el objetivo de la evaluación, el fin para el que se usarán los puntajes, las inferencias que se harán a partir de los puntajes, y las características de los individuos examinados y los subgrupos de la población prevista de la prueba que podrían influir en el acceso.

⁷ Los Estándares de Aprendizaje están contenidos en los decretos supremos de educación del Ministerio de Educación de Chile, N.º 129/2013, N.º 81/2014 y N.º 178/2015 y se encuentran disponibles en: <https://www.curriculumnacional.cl/614/w3-search.html>

Tabla 1.1.10*Formas aplicadas en las pruebas Simce censales 2017 según asignatura y nivel.*

Pruebas Simce / Niveles evaluados	4.º básico	8.º básico	II medio
Lectura	3 formas equivalentes 1 forma extendida 1 forma NEE DVT* 1 forma NEE DVP* 1 forma NEE DA*	2 formas equivalentes 2 formas estudio puente 1 forma NEE DVT 1 forma NEE DVP 1 forma NEE DA	2 formas equivalentes
Matemática	4 formas equivalentes 1 forma extendida 1 forma NEE DVP 1 forma NEE DA	2 formas equivalentes 3 formas estudio puente 1 forma NEE DVP 1 forma NEE DA	2 formas equivalentes 1 forma NEE DA
Historia, Geografía y Ciencias Sociales			2 formas equivalentes
Ciencias Naturales		2 formas equivalentes 2 formas estudio puente	

*DVT: discapacidad visual total

*DVP: discapacidad visual parcial

*DA: discapacidad auditiva

Uno de los elementos centrales en el ensamblaje de las pruebas definitivas de aplicación censal consiste en establecer equiparaciones, tanto entre las formas de una misma prueba, como entre pruebas aplicadas en distintos años. Sin las equiparaciones no sería posible comparar los resultados.

En virtud de lo anterior, la primera acción del armado consiste en seleccionar el conjunto de ítems de equiparación entre años (*equating* entre años) para cada prueba. Estos conjuntos conforman subpruebas de las mediciones de origen, esto es, una representación del total de ítems de cada evaluación, en términos de constructo y características psicométricas. Estos ítems de *equating* corresponden a no menos del 20% de las preguntas de la prueba de la medición anterior.

La selección de los ítems de *equating* entre años se realiza en función de una serie de criterios que aseguren su calidad, tales como que cuenten con parámetros IRT adecuados, que no presenten problemas de construcción, que no estén publicados (ni asociados a un estímulo publicado) y que, en su conjunto, abarquen el rango de dificultad de la prueba y sean representativos de la matriz teórica de la prueba respectiva del año, y de la prueba definitiva del ciclo anterior.

Una vez validados los ítems de *equating*, la etapa siguiente consiste en posicionarlos. Para ello se consideran dos criterios: (a) mantener el mismo ordenamiento relativo y (b) mantener una posición relativa similar a la que tuvieron los ítems en la prueba definitiva previa, en las distintas formas, asegurando en este último caso, que no exista una distancia superior a ± 5 posiciones respecto de la ubicación original en las distintas formas⁸. Así, el conjunto de *equating* entre años se distribuye en los distintos cuadernillos para asegurar la comparabilidad entre años y un subconjunto de ellos se repite en las distintas formas de la nueva prueba para asegurar la comparabilidad entre las mismas.

⁸ En el caso de los ítems asociados a un estímulo común (testlet), la distancia puede tener un leve desplazamiento adicional a las ± 5 posiciones, dado que deben ser ubicados en torno al estímulo que les sirve de base.

La tercera etapa del armado consiste en posicionar los ítems de equiparación entre formas, en función de los mismos criterios anteriores, es decir, asegurando que no exista inversión de posiciones entre las distintas formas y que no exista una distancia mayor a ± 5 posiciones en la ubicación del ítem en las distintas formas de la prueba. Los ítems que permiten equiparar las formas representan aproximadamente 20 % del total de preguntas de cada forma.

Una vez posicionadas las preguntas de equiparación entre formas y años, se distribuyen y posicionan los restantes ítems, atendiendo a los siguientes criterios:

- El posicionamiento general debe atender a la lógica interna de cada disciplina. Así, por ejemplo, en el caso de las pruebas de Lectura, el ordenamiento está supeditado a los textos para la lectura, y en el caso de las pruebas de Ciencias Naturales y de Historia, Geografía y Ciencias Sociales, la secuenciación debe realizarse en función del contenido, según recomienda el informe ACER (2013).
- Cada forma debe tener similar carga cognitiva, en términos de presencia de estímulos y figuras que deban procesarse de manera simultánea.
- Los primeros ítems de cada forma deben ser sencillos y motivadores, y los de mayor dificultad o más complejos deben estar en posiciones intermedias.
- Los ítems de respuesta abierta no deben ubicarse en las posiciones finales.
- Las claves (letras de opciones correctas) deben estar distribuidas de manera homogénea al interior de cada prueba y se debe asegurar que no existan patrones.
- Si hubiera ítems dependientes o encadenados (es decir, que contengan información útil para responder correctamente otras preguntas), se deben ubicar en formas distintas, para garantizar la independencia local de las preguntas.

Una vez finalizado el posicionamiento de las preguntas en cada forma, se evalúa que estas exhiban una función de información y una curva característica, similares entre sí y, en caso de detectar falencias en este ámbito, se realizan los ajustes pertinentes⁹.

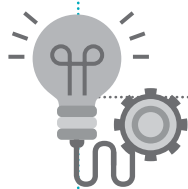
La última acción de la construcción de los instrumentos consiste en la diagramación de los cuadernillos. Las preguntas de las pruebas Simce se presentan en cuadernillos impresos, cuyas portadas contienen las instrucciones generales de la aplicación. Adicionalmente, se diseñan y elaboran hojas de respuesta donde los estudiantes deben marcar la opción que consideren correcta. Las preguntas abiertas se responden en el mismo cuadernillo de prueba, en 4.º y 8.º básico¹⁰, y en cuadernillos de respuesta separados, en II medio.

La diagramación de las pruebas Simce se realiza en función de los criterios y procedimientos que aseguran los aspectos técnicos del diseño y la comparabilidad interanual de las evaluaciones.

Luego de la diagramación de las pruebas, estas pasan por múltiples revisiones y controles de calidad por parte de los equipos internos de la Agencia.

⁹ Para los cursos con menos de diez estudiantes por nivel escolar, se elabora una prueba extendida. Esta prueba contempla todas las formas regulares, siempre y cuando, en su conjunto, no excedan una determinada cantidad de ítems. Si el conjunto de formas regulares excede la cantidad límite de ítems, se analizan los datos psicométricos y se decide qué ítems se pueden excluir de la prueba extendida, sin que se altere la función de información de la prueba regular. La prueba Simce 2017 que requirió contar con una versión extendida fue la de 4.º básico y el límite máximo se estableció en 80 ítems.

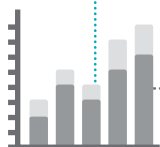
¹⁰ Desde esta aplicación 8.º básico no posee cuadernillo de respuesta.



Capítulo I:
Diseño y construcción de las
pruebas Simce 2017



Capítulo II:
Logística y aplicación de las
pruebas Simce 2017



Capítulo III:
Procesamiento y análisis
de datos 2017



Capítulo IV:
Comunicación de resultados
educativos 2017

Capítulo II: Logística y aplicación de las pruebas Simce 2017

La aplicación Simce involucra procesos operativos y logísticos que son igualmente relevantes para la calidad de los resultados obtenidos que la fase de diseño y construcción descrita en el capítulo anterior.

Estos procesos de aplicación de los instrumentos se llevan a cabo en función de instrucciones, manuales y protocolos de actuación que se basan en los criterios de estandarización internacional definidos en el documento *Standards for Educational and Psychological Testing* (AERA, APA y NCME, 2014). Tales documentos tienen por objeto ofrecer condiciones equivalentes a todos los estudiantes al momento de rendir la prueba; además, permiten asegurar confidencialidad de la información y resguardo de los materiales aplicados.

Considerando la extensión del territorio nacional y la cantidad de pruebas que se aplica cada año, la Agencia, para optimizar la eficiencia y calidad de los procesos, ha externalizado algunos servicios asociados a las operaciones de campo, a través de licitaciones públicas. No obstante, un equipo interno de profesionales tiene la responsabilidad de velar porque estos servicios externos cumplan estrictamente los criterios de calidad que exige la Agencia. Para ello, el equipo interno desarrolla un plan de supervisión para la ejecución de cada uno de los procesos.

Además, en el año 2017, con el propósito de asegurar la calidad de los procesos logísticos y de aplicación, la Agencia puso a disposición de la comunidad escolar una mesa de ayuda telefónica especial, implementada desde 2016, para responder consultas de los usuarios relacionadas con la aplicación.

Por último y de acuerdo con lo señalado anteriormente, con el propósito de desarrollar un proceso de aplicación Simce 2017 óptimo y de calidad, la Agencia desarrolló un riguroso plan de supervisión para las etapas que comprende el proceso logístico de empadronamiento, impresión y aplicación.

En el año 2017, se aplicaron pruebas Simce de manera censal en los niveles de 4.º básico, 8.º básico y II medio. Este capítulo aborda los procesos operativos para la aplicación en dichos niveles:

- Empadronamiento de los establecimientos y estudiantes.
- Impresión, mecanizado y distribución del material de aplicación.
- Aplicación de las pruebas.
- Entrega del material aplicado a procesamiento de datos.
- Atención a la comunidad escolar.

2.1 Empadronamiento de establecimientos y estudiantes

El proceso de empadronamiento 2017 corresponde a la validación de información de la base de datos del Sistema de Información General de Estudiantes (SIGE)¹¹ y la base de datos del Programa de Integración Escolar (PIE)¹² que contempla todos los establecimientos que rinden pruebas Simce.

La Ley General de Educación (Ley 20 370) establece que las pruebas Simce deben aplicarse en “todos los establecimientos educacionales de enseñanza regular del país” que tengan al menos un estudiante matriculado en el nivel escolar en que se aplica una determinada prueba, por tanto, únicamente quedan excluidas del proceso las escuelas de enseñanza no regular como: hospitalarias, vespertinas, carcelarias y de educación de adultos.

Para la definición de estos establecimientos se utilizan los datos que entrega la base de datos SIGE y PIE del Ministerio de Educación, los que se encuentran sistematizados en categorías de establecimientos según códigos. En la Tabla 1.2.1 se muestran las categorías de los establecimientos que fueron considerados para la aplicación de las pruebas Simce 2017.

Tabla 1.2.1

Categorías de establecimientos según códigos SIGE evaluados con las pruebas Simce 2017.

Código	Categoría
110	Educación Básica
310	Educación Media Humanístico-científica
410	Educación Media Técnico-profesional Comercial
510	Educación Media Técnico-profesional Industrial
610	Educación Media Técnico-profesional Técnica
710	Educación Media Técnico-profesional Agrícola
810	Educación Media Técnico-profesional Marítima
910	Educación Media Artística

El proceso de empadronamiento 2017 fue adjudicado mediante licitación pública¹³ y fue ejecutado durante el mes de junio.

¹¹ El SIGE es una base de datos del Ministerio de Educación que contiene la información del reporte de matrícula informado y actualizado mensualmente por los directores de los establecimientos educacionales del país.

¹² Corresponde a la base de datos con la información de aquellos estudiantes con Necesidades Educativas Especiales.

¹³ Licitación Pública “Servicio de Call Center para Proceso Simce 2017 de Ítem 1” ID N.º721703-6-LQ16.

En este proceso de validación se verifican y actualizan los datos del establecimiento, tales como: nombre, Rol de Base de Datos (RBD), dirección, teléfono y los datos de contacto del director. Además, se identifican las cantidades de estudiantes por curso y nivel de cada establecimiento, tanto regulares como especiales. En el caso de estudiantes con discapacidad sensorial¹⁴ (DS), se solicita nombre completo, RUN, tipo de discapacidad (DVT, DVP, DA)¹⁵ y las ayudas técnicas requeridas.

Todos estos datos son ingresados y actualizados en una plataforma de información llamada SIRE (Sistema de Ingreso y Registro de Empadronamiento). La información validada resulta clave para la coordinación y ejecución de los procesos de impresión, mecanizado, distribución, aplicación y captura de materiales, tales como: pruebas, cuestionarios y material complementario.

Respecto de los estudiantes, la población objetivo para las pruebas censales la constituyen los alumnos de los establecimientos antes señalados, incluyendo aquellos estudiantes en situación de discapacidad sensorial (DS) y con necesidades educativas permanentes (NEEP)¹⁶, para los que se construyen instrumentos o realizan procedimientos específicos. Para la identificación de los estudiantes que en 2017 rindieron la prueba Simce DS, la Agencia procede a registrar y validar a los alumnos de acuerdo con lo indicado en la base de datos PIE. Posteriormente, cuando la Agencia abre la Plataforma web NEEP es posible identificar los alumnos faltantes, dado que los directores acreditan la situación de cada estudiante con los certificados correspondientes.

Durante el proceso de empadronamiento 2017, para la aplicación de los niveles censales de 4.º básico, 8.º básico y II medio, se estableció una población total de 8346 establecimientos.

La construcción de la base de datos de pruebas censales, posterior al empadronamiento, se realiza en tres niveles: establecimientos, cursos y estudiantes. Esta base entrega la información necesaria para iniciar los procesos siguientes, tales como: impresión, mecanizado, distribución y aplicación.

La información obtenida en el proceso es fundamental para validar los datos de matrícula de estudiantes y estimar la población total de establecimientos y cursos a evaluar.

Todo el proceso de empadronamiento es supervisado por los equipos de la Agencia de Calidad incluyendo controles de calidad de los elementos críticos para el proceso: verificación de que el equipo que realiza la tarea cumpla con el perfil requerido, escucha de una muestra de los llamados realizados a los establecimientos y aplicación de procesos de validación de los registros recogidos.

¹⁴ Los estudiantes con discapacidad sensorial corresponden a aquellos que presentan discapacidad visual total o parcial y/o discapacidad auditiva, de acuerdo con lo establecido en el Decreto Supremo N.º 170 del Ministerio de Educación (Mineduc, 2009a).

¹⁵ Discapacidad Visual Total, Discapacidad Visual Parcial y Discapacidad Auditiva, respectivamente.

¹⁶ Necesidades educativas especiales.

2.2 Impresión, mecanizado y distribución del material de aplicación

La impresión, mecanizado y distribución del material son fases que, al igual que las anteriores, son clave para la obtención de resultados válidos en las evaluaciones. Un error en la impresión o mecanizado puede afectar gravemente el registro y la posterior captura de la información. Del mismo modo, una falla de impresión (ambigüedad en un color, pérdida de legibilidad de un texto o imagen, error en la compaginación, etc.) podría terminar invalidando las interpretaciones sobre los resultados de los estudiantes (Martin, Rust, & Adams (Eds.), 1999).

Asimismo, si se produce una mecanización con errores podría verse afectada la aplicación de uno o varios cursos, e incluso de un establecimiento completo, sin ser posible obtener resultados de esos estudiantes. Por todas estas razones, los procesos de impresión, mecanizado y distribución se realizan bajo supervisión por parte de los equipos de la Agencia, contemplando controles de calidad en terreno.

El proceso de impresión, mecanizado y distribución del material corresponde a un servicio externalizado adjudicado mediante licitación pública¹⁷ a dos proveedores. Este se ejecutó entre julio y octubre del año 2017.

El proceso de impresión incluye la impresión de cuadernillos, hojas de respuesta de las pruebas, cuestionarios y materiales de capacitación (pruebas de selección de examinadores, Manual de aplicación y Formularios de control, entre otros que se describen más adelante).

El mecanizado corresponde al proceso de ordenamiento y embalaje del material de aplicación en cajas por curso y sobres rotulados para su posterior distribución en los establecimientos. Cada sobre es identificado con una etiqueta codificada (código de barras) que contiene los datos de identificación del establecimiento, entre otros datos importantes.

Para el correcto mecanizado, se consideran todos los materiales necesarios para la aplicación (incluyendo instrumentos, formularios de registro y herramientas para las pruebas DS, como atril, regleta, punzón y otros). Estos materiales se organizan en cajas de acuerdo con el curso de cada establecimiento que aplicará (caja-curso), y finalmente se mecanizan en cajas contenedoras (que agrupan cajas-curso por cada día de aplicación, por establecimiento).

Para resguardar el proceso de aplicación, se mecanizan cajas de contingencia, que corresponden a cajas de idéntico contenido al de las cajas-curso, sin embargo, no se encuentran asociadas a un establecimiento ni curso en particular, por lo que se usan para reemplazo de cajas-curso en caso de errores, fallas de impresión de material u otro inconveniente.

La distribución consiste en el traslado del material mecanizado, para el que la Agencia instruye un esquema de tránsito del material que pone énfasis en el uso de códigos para identificar la trazabilidad durante todo el proceso de aplicación. Es por lo anterior, que el material mecanizado se traslada a los puntos de recepción informados previamente por cada proveedor de aplicación. Cada punto corresponde a un Centro de Acopio (CA) por región: bodegas habilitadas exclusivamente para el servicio de recepción,

¹⁷ Licitación Pública "Servicio de impresión, mecanizado y distribución", ID 721703-46-LR16.

revisión, almacenamiento, manipulación, custodia y despacho del material. La distribución comienza con el envío del material de capacitación, continúa con el material complementario y finaliza con el envío del material de aplicación. En el año 2017, se distribuyeron un aproximado de 9500 sobres complementarios y 9500 cajas contenedoras por nivel escolar evaluado.

Para asegurar un correcto desarrollo y calidad del proceso de impresión, mecanizado y distribución, la Agencia establece instancias y controles de calidad enfocados en revisar la infraestructura necesaria para estos procesos, incluyendo controles de calidad en terreno.

2.3 Aplicación de las pruebas

El proceso de aplicación de las pruebas Simce 2017 en territorio nacional se realiza mediante la externalización de servicios, adjudicados por licitación pública¹⁸.

Las pruebas Simce establecidas en el Plan de Evaluaciones para 2017 se aplicaron en las siguientes fechas:

- 8.º básico: 11 y 12 de octubre
- II medio: 17 y 18 de octubre
- 4.º básico: 7 y 8 de noviembre

2.3.1 Población evaluada

En términos de la población evaluada, las pruebas Simce 2017 en los niveles censales de 4.º básico, 8.º básico y II medio fueron aplicadas en promedio a un 99,2 % de los establecimientos y al 93,7 % de los estudiantes empadronados (663 615 estudiantes), con diferencias de cobertura según nivel evaluado, como se detalla en la Tabla 1.2.2. Los establecimientos en los cuales no fue posible aplicar las pruebas corresponden a aquellos que se encontraban cerrados o sin matrícula (donde la información no estaba actualizada en la base SIGE); no estaban en funcionamiento en las fechas de aplicación por motivos de fuerza mayor, o bien, no fue posible trasladar a los estudiantes por razones climáticas o de otro tipo.

¹⁸ Licitación pública "Aplicación de las Pruebas Simce año 2017 - 2018" ID 721703-1-LR17, para toda la extensión del territorio, desde la Región de Arica y Parinacota hasta la Región del Biobío, y mediante trato directo desde la Región de la Araucanía hasta Magallanes.

Tabla 1.2.2

Establecimientos y estudiantes evaluados en la aplicación censal de las pruebas Simce 2017.

Nivel escolar	N.º establecimientos evaluados	% cobertura establecimientos	N.º estudiantes matriculados	N.º estudiantes evaluados (a)	% estudiantes evaluados
4.º básico	7 344	98,7 %	244 123	230 604	94,5 %
8.º básico	5 975	99,4 %	236 475	222 567	94,1 %
II medio	2 912	99,4 %	227 976	210 444	92,3 %

(a) Corresponden a aquellos estudiantes que estuvieron presentes en la aplicación de al menos una de las pruebas Simce aplicadas en ese nivel escolar.

Las pruebas Simce DS de 4.º básico fueron rendidas por el 88,7 % de los estudiantes a nivel nacional que presentan discapacidad sensorial, lo que corresponde a 205 estudiantes, distribuidos en 149 establecimientos. En tanto en 8.º básico fueron rendidas por el 97,5 % de los estudiantes a nivel nacional en situación de discapacidad sensorial, lo que corresponde a 195 estudiantes, distribuidos en 146 establecimientos.

2.3.2 Fases del proceso de aplicación

El proceso de aplicación de las pruebas Simce incluye cinco fases operativas, las que son relevantes para la calidad de resultados obtenidos por los establecimientos educacionales en el territorio nacional. Estas son: habilitar infraestructura y capacidad operativa; reclutar, capacitar y seleccionar a examinadores y supervisores; realización de la visita previa a los establecimientos; aplicación en aula; y cuadratura y recolección de material hacia la captura.

A continuación, se describe brevemente cada una de estas fases:

i. Habilitar infraestructura y capacidad operativa

Corresponde a la habilitación de infraestructura regional para recibir, revisar y manipular el material enviado desde la imprenta, junto con el personal necesario para realizar las actividades operativas en terreno.

En esta fase los proveedores disponen una red nacional de centros que responden a condiciones de seguridad para el correcto resguardo de la confidencialidad y trazabilidad del material. Además, los centros de capacitación cuentan con medidas que aseguren la correcta capacitación de los postulantes a examinadores y supervisores.

La infraestructura regional que se habilita se compone de:

- Centros de Acopio (CA), donde se establece al menos uno por región.
- Centros de Operaciones (CO), donde la cantidad de centros se establece de acuerdo con la estimación de los proveedores y sus convocados para la capacitación.

Estos centros corresponden a lugares físicos donde se recibe, manipula, custodia y despacha el material para el desarrollo del proceso.

Tanto los centros de acopio como los de operaciones estuvieron ubicados en las capitales regionales o provinciales, dispuestos en zonas estratégicas y de fácil acceso.

Para contar con la habilitación de los respectivos centros, la Agencia dispone de una red de supervisores a nivel nacional, quienes revisan, por medio de visitas en terreno, que las características y requerimientos mínimos de infraestructura y ubicación en cada región estén de acuerdo con las características exigidas. Además, se dispone de la capacidad operativa regional, que tiene la labor de realizar todas las actividades operativas en terreno y en los respectivos centros, a modo de cumplir cabalmente los lineamientos entregados por la Agencia.

ii. Reclutar, capacitar y seleccionar a examinadores y supervisores

a. Proceso de convocatoria, capacitación y selección

Para el desarrollo de esta fase, la Agencia elabora un plan de capacitación que permite estandarizar la formación de los examinadores y supervisores. Este debe ser aplicado rigurosamente por los proveedores. El proceso de formación considera las competencias que deben tener todos los actores relevantes del proceso (relatores, examinadores y supervisores); las formas de evaluación, incluyendo aspectos relevantes a supervisar; y un plan de formación que incluye la malla curricular de cada postulante (examinador, examinador DS y supervisores).

En términos genéricos, las etapas que constituyen el proceso de capacitación (formación) de examinadores y supervisores del Simce 2017 son las siguientes:

- **Formación de relatores:** inicialmente, todos los relatores (encargados de capacitar a los examinadores y supervisores) pasan por un proceso de filtro curricular, es decir, se verifica la idoneidad del postulante con el perfil requerido. Posteriormente, los profesionales de la Agencia realizan una capacitación que considera una serie de actividades prácticas que se debe replicar en futuras relatorías. Finalmente, al término de la capacitación, se realizan evaluaciones de conocimiento y por competencias, donde se seleccionan aquellos participantes más idóneos para el cargo.
- **Convocatoria:** esta etapa corresponde al llamado realizado por los proveedores de aplicación, a través de medios de comunicación para convocar a postulantes para los roles de examinadores y supervisores del proceso Simce 2017. Previo a este llamado, los proveedores desarrollan un plan de medios (aprobado por la Agencia) en el que se establece el mensaje y medios de comunicación para publicar masivamente y dar inicio al reclutamiento. Posteriormente y luego de la convocatoria, se preseleccionan los candidatos de acuerdo con el perfil exigido por la Agencia.
- **Capacitación de examinadores:** los examinadores y supervisores preseleccionados son capacitados. Para ello, la Agencia proporciona a los proveedores manuales, presentaciones, actividades a realizar, mecanismos de evaluación y formularios, con la finalidad de que las capacitaciones que se realicen sean estandarizadas.

Para la capacitación DS, se desarrolla un capítulo exclusivo dentro del manual de aplicación donde se detallan los procedimientos necesarios para ejecutar eficientemente las pruebas. Este

documento incluye toda la información necesaria para cada examinador, independiente del tipo de establecimiento donde podría desempeñarse: especial¹⁹ o regular.

- **Selección:** se seleccionan para participar en el proceso a aquellos postulantes que cumplen con el perfil requerido, asisten a la capacitación, aprueban la Prueba de Evaluación de Competencias del Examinador (TEC), y obtienen el 90 % de respuestas correctas en las Pruebas de Aplicación Regular y Discapacidad Sensorial (PAR y PADS respectivamente).

La Prueba TEC evalúa competencias y habilidades conductuales de los postulantes en relación con el perfil. En particular, en esta evaluación se miden las siguientes competencias:

- orientación prosocial,
- estabilidad emocional,
- dominancia social,
- ajuste a las normas,
- manejo de impresión.

Las Pruebas de Aplicación Regular y DS (PAR y PADS), por su parte, tienen el propósito de medir el nivel de conocimiento y dominio de los candidatos respecto de los conceptos y orientaciones que la Agencia define para las aplicaciones Simce.

Además, para la selección de examinadores de las pruebas Simce DS, se considera que rindan una evaluación adicional: la Evaluación de Competencias Técnicas en DS, que determina el grado de dominio que tienen los postulantes sobre la atención educativa requerida por los estudiantes en situación de discapacidad sensorial, como las ayudas técnicas y los sistemas de comunicación.

Con el fin de controlar y evitar variaciones que pudiesen derivar en fallas sistemáticas o aleatorias y que eventualmente pudiesen afectar la aplicación, la Agencia supervisa una muestra representativa de sesiones de capacitación.

Cabe mencionar que durante el proceso se capacitó a un contingente superior de examinadores y supervisores al necesario para las aplicaciones, con el objeto de prevenir eventuales bajas en cantidad de personal preparado, por lo que cada proveedor de aplicación debió contar con al menos un 10 % de examinadores de holgura respecto de la cantidad establecida por nivel, incluyendo DS.

¹⁹ Las Escuelas Especiales son establecimientos educacionales especializados que educan a estudiantes con discapacidad que requieren apoyos permanentes e intensivos en el área de la comunicación, del funcionamiento adaptativo, de su autonomía e independencia personal, y adecuaciones curriculares de acceso o a los objetivos de aprendizaje según las características y necesidades de los estudiantes frente al currículo, para así asegurar su proceso educativo y de transición a la vida adulta.

b. Material de capacitación

Durante la capacitación se utiliza el Manual de Aplicación. En 2017 este incorpora toda la información para la aplicación de establecimientos en educación regular y de establecimientos con educación para estudiantes DS en un solo documento.

En este manual están consignadas las características de la aplicación, según la prueba a aplicar, las responsabilidades y funciones del examinador; los procedimientos que deben seguirse antes, durante y después de la aplicación; y las consideraciones especiales requeridas en algunas de las pruebas.

Los documentos que se utilizan en los talleres prácticos corresponden al “material complementario”, esto es, los formularios que el examinador deberá completar en las distintas fases del proceso. Para la aplicación regular, se utilizan dos formularios:

- **Formulario Lista de Curso:** constituye el documento guía de la aplicación por curso. Incluye el detalle de los códigos y series de las pruebas asignadas a cada estudiante en cada curso, la identificación del establecimiento y de los estudiantes, incluyendo nombre, RUN, fecha de nacimiento y sexo. Este formulario contempla el registro de hasta 48 estudiantes en cada curso y permite que los examinadores completen la información en la eventualidad de que se deba reasignar o incorporar a estudiantes nuevos no registrados en la base de datos original.
- **Formulario de Aplicación,** en el que, identificando los datos del establecimiento y del curso, se registra toda la información relacionada con la gestión de la aplicación: cantidad de material recibido y entregado, pruebas anuladas, uso de material de contingencia, entre otros. Este formulario permite, además, tanto al examinador como al director del establecimiento, registrar cualquier situación anómala que ocurra durante la aplicación, para lo que cuenta con una sección denominada Acta de Control de Eventos.

Finalmente, para la aplicación Simce DS se utiliza el Formulario DS, que corresponde a una fusión de los formularios utilizados en la aplicación regular, es decir, incluye las mismas secciones de ambos documentos, pero con adaptaciones adecuadas a los procesos de control y registro que se requieren en esta aplicación específica.

iii. Realización de la visita previa a los establecimientos

La visita previa a los establecimientos constituye el primer contacto entre el equipo de aplicación (examinadores y supervisores) y los establecimientos. En esta instancia se realizan los preparativos para la aplicación y se verifican y validan los datos de los estudiantes.

En este hito, los examinadores y supervisores se reúnen tanto con la dirección del establecimiento como con cada uno de los cursos en los que se aplicarán las pruebas.

En la reunión con la dirección del establecimiento, el examinador:

- Informa sobre los procedimientos de la aplicación y señala el cronograma de aplicación correspondiente a cada día.
- Entrega el modelo de comunicación informativa de la aplicación para los padres y apoderados.
- Solicita la colaboración de los directivos durante el día de la aplicación, para que el proceso se desarrolle en condiciones óptimas. Dada la importancia de la colaboración de los directores de los establecimientos, en 2017 la Agencia envió con antelación un documento en formato digital e impreso denominado Programa de Evaluaciones Educativas para Directores, tanto de Educación Básica como de Educación Media. En este documento se informa detalladamente el proceso y actividades en las se requiere la colaboración de los directores.
- Verifica que la información impresa en el Formulario Lista de Curso de cada estudiante (nombre, RUN, sexo), incluyendo si existen alumnos en situación de discapacidad sensorial, coincida con la información contenida en los libros de clases. En caso de detectar discrepancias, el examinador corrige o completa la información.

Durante la visita a los cursos, el examinador:

- Se presenta a los estudiantes y los motiva para que asistan el día de la aplicación y rindan las evaluaciones seriamente.
- Informa sobre los materiales que recibirán el día de la aplicación (cuadernillos, lápiz grafito y goma de borrar).
- Indica la prohibición de uso de artículos electrónicos durante la aplicación, tales como celulares, calculadoras, etc., así como de cualquier otro material de apoyo, como diccionarios, libros, cuadernos, etc.
- Verifica que la sala de clases que se ocupará el día de la aplicación posea una adecuada luminosidad, que se encuentre en buen estado, que cuente con el número suficiente de mesas y sillas, que no esté próxima a ruidos externos molestos, y que no posea material didáctico a la vista. En caso de que la sala no cumpla con algunos de estos requisitos, el examinador solicita al director del establecimiento el cambio de sala para el día de la aplicación.

iv. Aplicación en aula

El desarrollo de la etapa de aplicación en aula comprende la realización de las siguientes actividades por parte del examinador:

- Retirar el material de aplicación (caja-curso-día) desde el centro de operaciones o centro de acopio correspondiente al establecimiento.
- Llegar al establecimiento a la hora acordada (siempre antes de las 8:00 am), llevar peto, la credencial de examinador y la cédula de identidad.
- Saludar y presentarse ante el director o la autoridad competente (subdirector, jefe de la Unidad Técnico Pedagógica (UTP) o sostenedor) y reunirse con el profesor colaborador asignado por el director para informarle la programación del día y coordinar la organización de los estudiantes.
- Revisar el material recibido en la sala de clases.

- Preparar la sala de clases.
- Ordenar las pruebas respecto de los estudiantes.
- Hacer ingresar a los estudiantes a la sala.
- Registrar la asistencia.
- Verificar las series por estudiantes.
- Leer las instrucciones en voz alta a los estudiantes.
- Aplicar los instrumentos correspondientes a cada día de aplicación.
- Firmar y solicitar al director que firme el Acta de Control de Eventos, contenida en el Formulario de Aplicación.
- Recuperar todo el material de la aplicación, cuidando de dar cumplimiento a los tiempos mínimos y máximos estipulados para cada prueba y nivel.
- Realizar la revisión y cuadratura de todo el material aplicado.
- Realizar la cuadratura y cierre de la caja-curso-día.
- Devolver el material aplicado al centro de operaciones o centro de acopio respectivo.

Al igual que en la etapa de capacitación, la aplicación en aula es observada por la Agencia con el objeto de resguardar el cumplimiento de los estándares del proceso y los procedimientos de aplicación. En 2017, la supervisión fue realizada por un equipo de profesionales de la Agencia, a una muestra de 272 cursos a nivel nacional, donde a través de las pautas de supervisión, se identificó el nivel de adherencia de los examinadores a los lineamientos entregados.

v. Cuadratura y recolección de material hacia captura

Junto con las fases antes descritas, que involucran directamente a examinadores, supervisores de aplicación y demás estamentos de supervisión, se encuentran los procesos de cuadratura y recolección del material. En ellos se debe velar por recuperar la totalidad de los instrumentos aplicados y no aplicados durante el proceso Simce.

Luego de la aplicación por nivel, los examinadores y supervisores proceden a revisar y cuadrar cada uno de los documentos, cerrar la caja-curso correspondiente y devolverla a su respectivo centro de operaciones; donde se ordena y prepara el material para su retorno a los centros de captura.

Para verificar el correcto desarrollo y calidad del proceso de aplicación, los equipos de la Agencia establecen controles de calidad que involucran revisión y verificación en terreno de la infraestructura y revisión del proceso de traslado de materiales en los Centros de Acopio (CA), de los Centros de Operaciones (CO) y Centros de Capacitación (CC).

2.4 Entrega del material a procesamiento de datos

El proveedor de aplicación es el responsable, custodio y encargado de realizar el despacho del material de aplicación a los centros de captura para su correspondiente procesamiento. Al momento del despacho, las caja-curso deben contener todo el material de aplicación asociado a un curso (pruebas, cuestionarios y formularios de control), y a su vez estas cajas deben ser despachadas dentro de sus respectivas cajas contenedoras. Para verificar este proceso y asegurar un correcto traspaso de materiales de aplicación, la Agencia supervisa estos procedimientos a través de controles y comprobantes.

Con al menos 24 horas de antelación a una entrega en el centro de captura, los proveedores de aplicación informan a la Agencia y al mismo centro de captura, vía correo electrónico, el detalle del material a entregar.

Como una forma de asegurar un correcto traspaso de materiales de aplicación en los centros de captura, los proveedores de aplicación cuentan con personal propio en las dependencias de captura, con el objetivo de validar la entrega efectiva al proveedor de captura.

2.5 Atención a la comunidad escolar

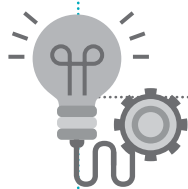
Con la finalidad de facilitar el acceso a la información y la generación de respuestas rápidas a consultas relacionadas con el proceso de aplicación de las pruebas Simce 2017, la Agencia pone a disposición de la comunidad escolar el servicio de Mesa de Ayuda Telefónica²⁰ para recibir, gestionar y responder todas las solicitudes de la comunidad escolar, como también mejorar el funcionamiento de la aplicación de las pruebas Simce.

La plataforma permite registrar, almacenar y extraer una base de datos con todas las consultas, reclamos y otros requerimientos ingresados. Además, se genera un manual de atención de usuarios, para que el proveedor de este servicio pueda utilizarlo como herramienta de apoyo, dado que revela el funcionamiento de la aplicación y los procedimientos para la atención de requerimientos.

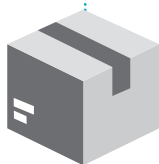
Para el proceso de formación de los agentes de la mesa de ayuda, se realizaron dos jornadas de capacitación. Estas consistieron en una capacitación teórica donde se expusieron todos los conceptos establecidos en el Manual de Atención de Usuarios con sus respectivas especificaciones de acuerdo con cada aplicación, y otra capacitación práctica, en la que se ejecutaron pruebas de la plataforma de información desarrollada para este servicio.

La totalidad de casos ingresados a la plataforma de información permite generar una base de datos de acontecimientos que reporta todas las situaciones acontecidas y manifestadas a la Agencia durante el proceso de aplicación. En el periodo en que se extendió este proceso ingresaron 904 casos, los que en su mayoría estuvieron relacionados con solicitudes de información y situaciones de carácter operativo, como consultas sobre las fechas de aplicación de las pruebas, entre otras.

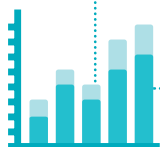
²⁰ Licitación Pública ID N.º 721703-6-LQ16. "Servicio de Call Center para Proceso Simce 2017 de Ítem 2". Servicio que se llevó a cabo entre los meses de septiembre y diciembre de 2017.



Capítulo I:
Diseño y construcción de las
pruebas Simce 2017



Capítulo II:
Logística y aplicación de las
pruebas Simce 2017



Capítulo III:
Procesamiento y análisis
de datos 2017



Capítulo IV:
Comunicación de resultados
educativos 2017

Capítulo III: Procesamiento y análisis de datos 2017

La fase de procesamiento y análisis de datos de las pruebas Simce 2017, al igual que las anteriores, se realiza bajo estrictos controles de calidad para asegurar que los resultados obtenidos reflejen correctamente los constructos medidos. El procesamiento supone la creación de una base de datos que, posterior a la validación, consolida la información de las respuestas de los estudiantes a las pruebas y la información de los cuestionarios para estudiantes, docentes, directivos y padres y apoderados. El análisis de los datos, por su parte, involucra tanto los análisis psicométricos de las pruebas, como los análisis agregados, tales como promedios nacionales, regionales, por género, etc.

En este capítulo se describen las acciones involucradas en la etapa de procesamiento y análisis de datos de las aplicaciones Simce, a saber:

- Captura, validación y consolidación de la base de datos
- Corrección de preguntas de respuesta abierta
- Análisis psicométricos de las pruebas
- Análisis de datos agregados

3.1 Captura, validación y consolidación de la base de datos de pruebas y cuestionarios

La primera acción antes del proceso de captura, validación y consolidación de los datos consiste en revisar el Plan de Gestión de Datos, documento que contiene el detalle de cada uno de los procesos informáticos que se ejecutan para la entrega de resultados de las pruebas Simce que aplican cada año, e incluye los criterios y etapas que deberán desarrollar los proveedores externos del servicio. Esta acción se realiza antes de la aplicación de las pruebas, en el 2017 involucró aspectos de mejora respecto del plan establecido en la medición anterior. Algunos de ellos fueron la actualización y mejoras en la optimización de rutinas de validación de datos, con las que se reduce el tiempo de ejecución de procesos masivos de revisión, mejorando así los tiempos de respuesta del servidor en el procesamiento de rutinas. También se mejora la asignación de los errores en la base de datos, para que procedan a ser validados a través del SVD (Sistema de Validación de Datos).

Una vez revisado el Plan de gestión de datos, se generan los insumos para el proceso de captura y validación de los datos. Como insumo para el procesamiento, se elabora un diccionario de datos, que define los criterios a partir de los cuales se sistematizará la información de las respuestas y las normas, codificaciones y los rangos de valores permitidos para cada uno de los campos que serán utilizados en el procesamiento. La información para construir el diccionario de datos se obtiene de los distintos documentos en formato PDF que serán utilizados durante la aplicación, como las listas de curso, los formularios de aplicación, los cuestionarios y las hojas de respuesta.

Tomando como base los documentos antes mencionados, se construyen las estructuras de las tablas de bases de datos, además se comienza con la actualización y mejoras de las rutinas que ayudan en el procesamiento de los datos, identificando los datos de los establecimientos, cursos y estudiantes.

La identificación de los establecimientos se realiza con el número de Rol de Base de Datos (RBD)²¹ y para la identificación del nivel escolar se utiliza un código de dos dígitos: el primer dígito identifica el grado y el segundo dígito es un cero (0) si corresponde a Educación Básica o dos (2) si es Educación Media. Así, en la aplicación 2017, se definieron tres códigos para la variable de identificación de los niveles, los que se muestran en la Tabla 1.3.1.

Tabla 1.3.1

Códigos de la variable de identificación por nivel escolar de la aplicación 2017.

Nivel escolar	Código
4.º básico	40
8.º básico	80
II medio	22

Para la identificación de los cursos se define un código de cinco dígitos, con códigos únicos, que se enumeran de forma continua para cada nivel.

La identificación de los estudiantes se realiza a través de un número de serie único al cual se asocian todas las pruebas y cuestionarios respondidos por los estudiantes, y luego se cataloga a los estudiantes en función de la variable de participación, cuya información se obtiene de las listas de curso. En la Tabla 1.3.2 se muestran las categorías de la variable de participación de estudiantes de la aplicación 2017.

Tabla 1.3.2

Categorías de la variable de participación de estudiantes en la aplicación 2017.

Categoría	Descripción
Presente	Estudiante asiste el día de la aplicación.
Ausente	Estudiante no asiste el día de la aplicación.
Nulo	Alumno atrasado, llega después de que se leen las instrucciones. Alumno abandona por impedimento temporal (vómitos, dolor de cabeza, dolor abdominal, etc.). Error en las pruebas y/o cuestionarios (mala compaginación, impresión, legibilidad, etc.). Alumno con mala conducta y/o sorprendido copiando. Alumno demora más de 15 minutos en regresar a la sala de clases al salir con permiso al servicio higiénico.

²¹ El Ministerio de Educación es el encargado de asignar un RBD a los establecimientos en Chile, haciendo un reconocimiento oficial del establecimiento.

La captura, validación y consolidación de la base de datos de pruebas y cuestionarios de la aplicación Simce 2017 se llevó a cabo entre octubre de 2017 y enero de 2018.

El proceso de captura se compone de diversas etapas, la primera es la recepción de las cajas en las bodegas del centro de captura con el material aplicado y no aplicado; la segunda etapa consiste en llevar el control de este material a través de un inventario o arqueo; la siguiente etapa consiste en separar las cajas y ordenar el material por tipo de documento y nivel al cual corresponde, y llevarlo a la línea de producción de este; luego, con el uso de un *software*, se procede a la lectura de la información de los distintos documentos aplicados, su conversión a imágenes con caracteres digitales y, finalmente, a archivos de datos electrónicos. Para convertir la imagen con caracteres digitales a un archivo de datos electrónico, en 2017 se utilizaron tres procesos: Reconocimiento de Caracteres Inteligente (ICR), Reconocimiento Óptico de Caracteres (OCR) y Lectura de Marcas Ópticas (OMR).

Una vez que se obtienen los archivos de datos electrónicos, se traspasan los datos extraídos de cada imagen a formato CSV (valores separados por comas), se realiza un control de calidad y posteriormente se entrega una muestra de datos de 300 cursos por nivel a la Agencia. A partir de la base de datos antes señalada, se ajustan las rutinas de programación y los paquetes de carga, que corresponden a los datos obtenidos de la captura, para detectar posibles inconsistencias, y se realizan controles de calidad de la captura, en función de lo establecido en el Plan de Gestión de Datos vigente. Para asegurar la calidad de estos procesos, cada año se actualiza el *software* de validación y corrección de los datos (SVD26), de acuerdo con las nuevas estructuras que presentan los formularios de control, las pruebas y los cuestionarios aplicados.

Utilizando un set de datos de 300 cursos por nivel escolar aplicado, el que es entregado por el proveedor de captura, se realizan pruebas de carga en los sistemas preparados por la Unidad de Gestión de Datos. En esta etapa se revisa el cumplimiento del formato establecido en el diccionario de datos y se verifica que las estructuras y contenidos cumplan con las especificaciones entregadas en reunión de inicio del proceso. La finalidad de este proceso es retroalimentar al proveedor en función de las inconsistencias detectadas, para que este realice las correcciones pertinentes previo a la entrega de las bases de datos definitivas.

En la Tabla 1.3.3 se muestra el detalle de los tipos de documentos capturados en el proceso de aplicación 2017, en cada nivel escolar y prueba aplicada.

Tabla 1.3.3

Tipos de documentos capturados en 2017 por nivel escolar y prueba aplicada.

Datos	4.º básico	4.º básico DS	8.º básico	8.º básico DS	II medio
Lista Curso (LC)	✓	✓	✓	✓	✓
Formulario Aplicación (FA)	✓	✓	✓	✓	✓
Preguntas Cerradas (PC)	✓	✓	✓	✓	✓
Preguntas Abiertas (PA)	✓	✓	✓	✓	✓
Cuestionario de Padres (CP)	✓	✓	✓	✓	✓
Cuestionario de Estudiantes (CE)		✓	✓	✓	✓
Cuestionario Docentes Lenguaje (DL)	✓				✓
Cuestionario Docentes Matemática (DM)			✓		✓
Cuestionario Docentes Cs. Naturales (DN)			✓		
Cuestionario Docentes		✓			

Con esta información la Agencia realiza los procesos de completitud, validación y consolidación de datos.

El proceso de completitud consiste en verificar la recepción, captura y procesamiento del total de material recibido por el proveedor de captura, que posteriormente es entregado a la Agencia. El material faltante se consigna en un campo de la base de datos para poder identificar si corresponde a formularios de control, pruebas o cuestionarios de contexto, y que posteriormente permita la generación de tres reportes: (a) reporte de completitud global (identificando a nivel de región, comuna y establecimiento el material esperado frente a lo recibido); (b) reporte de completitud detallado (contiene el detalle de lo faltante por tipo de documento); y (c) reporte de completitud de gestión (comunica el estado de la base de datos a nivel macro, estableciendo rangos de porcentaje del material faltante). Estos reportes facilitarán la gestión en la recuperación de material por parte de la Agencia, permitiendo prever posibles problemas en la consolidación de bases de datos o entrega de resultados.

La validación de los datos, por su parte, se realiza mediante el SVD y distingue la validación de pruebas, formularios de control y la validación de cuestionarios.

El proceso de validación de pruebas consiste en la revisión de la consistencia e integridad de los datos de captura enviados por el proveedor y la corrección de la aplicación de acuerdo con las reglas de validación. En este proceso, mediante una automatización de rutinas de programación, se marcan errores en los registros, los que se clasifican según formulario y prioridad, y se agrupan mediante un código. Luego, mediante el SVD un equipo de validadores, contratados y capacitados especialmente para este propósito a través de la licitación antes nombrada, es responsable de corregir una serie de errores. La asignación de errores es de manera aleatoria, para favorecer la imparcialidad del proceso.

Existen dos tipos de errores: se entiende por “error detectado modificable”, aquellas inconsistencias producidas por el llenado de datos incorrectos en los formularios por parte del examinador o una incorrecta captura del dato, y se entiende por “error detectado no modificable” las inconsistencias producidas por el estudiante, como, por ejemplo, que marque más de una alternativa ante una pregunta cerrada o que omita una respuesta; estas últimas quedan permanentemente con el error de alerta, sin embargo, se revisan para asegurar que no existe un problema de captura.

En el proceso de validación de datos de pruebas 2017 se corrigió el 100 % de los errores detectados modificables y se revisaron los errores no modificables.

La validación de cuestionarios, por su parte, consiste en verificar la consistencia de captura, revisando frecuencias de respuestas y muestras de documentos, además se comprueba que la serie del cuestionario corresponda al estudiante, curso y establecimiento respectivos. Para validar los cuestionarios se revisan los datos capturados del Formulario de Aplicación, el que contiene las cantidades de documentos utilizados, los que son comparados con los datos de los cursos que efectivamente rindieron las pruebas y con los cuestionarios que retornaron de la aplicación. A partir de esta comparación de información se validan los datos, se analiza la cuadratura de estos y se corrobora que no existan duplicados de series. Una vez validados los datos, se revisa caso a caso cada set de cuestionarios, para no incluir en el análisis posterior a aquellos que no correspondan. Por ejemplo, en el Cuestionario de Estudiantes, se requiere distinguir a aquellos que estuvieron presentes, tanto al momento de rendir las pruebas como de contestar el cuestionario, de aquellos que únicamente rindieron las pruebas. También se debe velar porque cada curso tenga asociado un Cuestionario Docente, es decir, que en la base de datos de cada curso exista un cuestionario docente asociado.

Una vez ratificados todos los datos, se genera un reporte de frecuencias de análisis en el que se distingue cantidad y porcentaje de ocurrencia para cada campo capturado, lo que permite obtener una visión general de la omisión de respuestas y de las múltiples marcas, e identificar valores fuera de rango. En función de este reporte se definen los campos que es necesario corregir o incluso reprocesar en captura si fuera necesario.

La última acción del proceso de elaboración y validación de la base de datos consiste en la consolidación, cuadratura y paridad de las bases. La consolidación consiste en traspasar los datos validados a una base de datos estructurada para llevar a cabo el proceso de análisis de datos. Esta base consolidada contiene la codificación de las respuestas de las preguntas abiertas, recibidas desde el proceso de corrección de preguntas abiertas y preguntas cerradas; además, se cargan los datos obtenidos de los cuestionarios, los que permiten realizar posteriormente los análisis agregados. Realizada la consolidación de datos, la cuadratura consiste en un control de calidad de los datos consolidados en donde se revisa el total de la asistencia en sus dos estados (Presente y Nulo) y se contabiliza el total de pruebas y cuestionarios aplicados, en donde a nivel de curso estos totales deben ser los mismos. Finalmente, luego de que se cuadre los cursos con sus respectivas pruebas y cuestionarios, se realiza el pareo de los datos, que consiste en verificar que cada alumno presente debe tener su prueba y cuestionarios, y cada prueba y cuestionario debe tener su alumno presente.

Los productos consolidados que se obtienen a partir de los procesos descritos en este apartado son dispuestos para la Unidad de Análisis Estadístico.

3.2 Corrección de preguntas de respuesta abierta

Como se señaló en el primer capítulo de esta sección, las preguntas de respuesta abierta se utilizan fundamentalmente para medir aprendizajes que no son susceptibles de ser evaluados con formato cerrado o cuyo constructo se mide en forma más precisa en formato abierto.

Este tipo de preguntas requiere un mayor tiempo de corrección que el de las de opción múltiple y un proceso que supone contar con equipos de correctores alineados con los criterios de evaluación y consistentes entre sí, para mantener la validez de la corrección (Brancato, Macchia, Murgia, Signore, Simeoni, Blanke, & Hoffmeyer-Zlotnik, 2006).

La corrección de preguntas abiertas se realiza mediante rúbricas aplicadas por correctores, por lo que se requiere establecer mecanismos que aseguren la confiabilidad de la corrección de las respuestas. Entre estos mecanismos se incluyen: una capacitación de correctores previa al inicio del proceso para lograr alineación con los criterios de evaluación, y un monitoreo constante de los porcentajes de consistencia entre correctores²², en función de estándares previamente establecidos.

El proceso Simce 2017 requirió la corrección de preguntas de respuesta abierta en todas las pruebas censales aplicadas. Al igual que en años anteriores, las respuestas de los estudiantes a las preguntas abiertas fueron dispuestas para su corrección en formato de imágenes digitalizadas (obtenidas en el proceso de captura, según se reseñó en este capítulo anteriormente), y para su corrección se utilizó el software de corrección de creación propia de la Agencia, Corrector.

El proceso de corrección 2017 se realizó conformando equipos expertos de las distintas disciplinas evaluadas. Administrativamente, este proceso fue licitado²³ y llevado a cabo por un proveedor externo con supervisión directa de profesionales de la Agencia.

En la Tabla 1.3.4 se muestra la cantidad total de correctores que participaron en el proceso 2017 y la cantidad de grupos de corrección conformados para cada asignatura y nivel escolar de las evaluaciones censales que contaron con preguntas de respuestas abiertas.

²² La consistencia entre correctores se refiere a la proporción de respuestas idénticas que son juzgadas del mismo modo por dos o más correctores distintos.

²³ Licitación pública "Servicio corrección de respuestas abiertas incluidas en las pruebas Simce años 2017 y 2018" ID 721703-11-LR17.

Tabla 1.3.4*Distribución de correctores de preguntas abiertas pruebas censales Simce 2017.*

Nivel Escolar	Prueba	N.º correctores (aproximado)	N.º grupos de correctores
4.º Básico	Lectura	48	4
	Matemática	12	1
8.º Básico	Lectura	24	2
	Matemática	12	1
	Ciencias Naturales	36	3
II Medio	Lectura	48	4
	Matemática	12	1
	Historia, Geografía y Cs. Sociales	12	1

La corrección de preguntas abiertas de pruebas se realiza en función de los manuales de corrección validados en la aplicación experimental previa, que en el caso de las pruebas censales 2017 correspondió a las aplicadas experimentalmente en 2016, tal como se menciona en el Anexo A.

El proceso se inicia con la selección y capacitación de correctores y supervisores, a cargo de la institución convocada mediante la licitación pública ya nombrada. Los equipos de trabajo conformados realizan una marcha blanca durante la cual se ejecuta una doble corrección del 10 % de las respuestas obtenidas de cada pregunta. Si durante la marcha blanca no se alcanza el mínimo de 80 % de consistencia, se repite la capacitación y, de ser necesario, se reemplazan correctores o supervisores, y se realiza una nueva marcha blanca. Una vez ajustados los equipos de trabajo, se corrige el 90 % restante de las respuestas, con doble corrección del 10 % y triple corrección del 10 % de las respuestas doblemente corregidas.

Al finalizar el proceso, los códigos asignados durante la corrección se traducen a puntajes que son equiparables con la escala de puntajes utilizada en la corrección de las preguntas de opción múltiple. En las pruebas censales 2017 se corrigió un total de 33 preguntas abiertas, distribuidas por asignatura y nivel escolar, según se muestra en la Tabla 1.3.5.

Tabla 1.3.5*Cantidad de preguntas abiertas pruebas censales Simce 2017.*

Nivel escolar	Prueba	N.º de preguntas abiertas
4.º básico	Lectura	8
	Matemática	4
8.º básico	Lectura	5
	Matemática	4
	Ciencias Naturales	4
II medio	Lectura	4
	Matemática	2
	Historia, Geografía y Cs. Sociales	2

3.3 Análisis psicométricos de las pruebas

Los análisis psicométricos tienen por objetivo entregar indicadores de rendimiento individual de los estudiantes en cada una de las pruebas aplicadas, tanto a nivel de puntaje, como de clasificación en Estándares de Aprendizaje, de haberlos.

Estos análisis se realizan sobre la base de lineamientos internacionales (AERA, APA y NCME, 2014) que aseguran la estandarización de los procesos, y están sujetos a estrictos controles de calidad, incluyendo un procesamiento en paralelo, lo que permite detectar inconsistencias y verificarlas, de haberlas, así como ajustarlas de manera inmediata.

En las aplicaciones Simce los análisis psicométricos se realizan utilizando como insumo los archivos denominados Planilla Maestra y Base de Datos de Estudiantes, los que contienen la información de los ítems aplicados (identificación, objetivo de evaluación, respuesta correcta o clave, posición, etc.) y de las respuestas de los estudiantes, respectivamente. Las respuestas a las preguntas cerradas tienen dos puntuaciones posibles: 0 (incorrecta) y 1 (correcta), y las respuestas a preguntas abiertas pueden tener una puntuación igual a las cerradas (respuesta dicotómica) o bien una de tres niveles o politómica: 0 (incorrecta), 1 (parcialmente correcta) y 2 (correcta), dependiendo de lo que establezca la pauta de corrección respectiva.

Los modelos que orientan los análisis psicométricos de las pruebas Simce son la Teoría Clásica y la Teoría de Respuesta al Ítem (IRT).

En 2017, la Teoría Clásica se utilizó para el análisis de las pruebas Simce DS, además del análisis de pruebas experimentales, detallado en el Anexo A y para su ejecución se utilizaron los softwares SAS 9.1 (versión 2006) y TIAPLUS: build 307 (versión 2010). El modelo IRT, por su parte, se utilizó para realizar los análisis de las pruebas censales de Lectura, Matemática, Historia, Geografía y Ciencias Sociales, y Ciencias Naturales. En particular, se utilizaron: el Modelo Logístico de Tres Parámetros (3PL) para las preguntas cerradas; el Modelo Logístico de dos Parámetros (2PL), para los ítems de respuesta abierta dicotómica; y el Modelo de Crédito Parcial Generalizado (MCPG), para los ítems de respuesta abierta politómica.

Realizar los análisis según el modelo IRT, hace necesario revisar cada año el cumplimiento de los tres supuestos bajo los cuales estos modelos resultan eficientes y robustos: (a) unidimensionalidad, esto es que las pruebas evalúen una dimensión de rendimiento predominante; (b) independencia local, es decir, que para los estudiantes con un mismo nivel de rendimiento no existan otros factores que expliquen sus respuestas correctas a los ítems; y (c) monotonicidad creciente, es decir, que la probabilidad de responder correctamente a un ítem aumente progresivamente en aquellos estudiantes con mayor nivel de rendimiento.

En la aplicación 2017, para la revisión del supuesto de unidimensionalidad se aplicó un Análisis Factorial Exploratorio (AFE) (De Vellis, 2012) mediante el *software* MPLUS 2010. Este *software* analiza la información a través de la magnitud de los valores propios resultantes y del porcentaje de varianza que estos entregan en el modelo, y luego permite visualizar los resultados en un *scree plot*²⁴. Como

²⁴ Método gráfico que muestra los valores propios de la matriz de correlación en orden descendente, esto es, se comienza

resultado, en todas las pruebas 2017 se evidenció la existencia de un factor dominante, representado por un quiebre significativo entre la magnitud de los factores primero y segundo²⁵.

Para la revisión del supuesto de independencia local, en 2017 se realizó una revisión cualitativa tendiente a verificar que no existieran ítems en esta situación, o que la respuesta correcta o incorrecta a un ítem no influyera en la respuesta correcta o incorrecta a otro. Adicionalmente, en aquellos conjuntos de ítems que compartían un mismo estímulo (imagen, texto, etc.), conocidos como *testlet*, se detectaron los ítems con valores inusualmente altos en el parámetro de discriminación, y se recalibró la prueba eliminando cada uno de manera separada. Si la estimación de los parámetros de discriminación restantes (sin el ítem eliminado) no poseía valores altos, se estimó que existía dependencia local y, por tanto, se eliminó el ítem.

Por último, para la revisión del supuesto de monotonicidad creciente, se revisaron las curvas características y las curvas de información y se verificó, en las curvas empíricas de cada ítem, que la correspondiente a la clave aumentara progresivamente su probabilidad en todos los ítems considerados para el análisis y su puntuación²⁶.

Respecto de los parámetros que se calculan en función de estos modelos, cabe señalar que estos dependen de si la pregunta es cerrada o abierta. Como se muestra en la Tabla 1.3.6, para las preguntas cerradas de la aplicación 2017 se calcularon seis indicadores según Teoría Clásica²⁷ y cinco según IRT. Para las abiertas, como se muestra en la Tabla 1.3.7, se calcularon dos indicadores de Teoría Clásica y cuatro de IRT.

Tabla 1.3.6

Indicadores 2017 para preguntas de respuesta cerrada según Teoría Clásica e IRT.

Teoría Clásica	IRT
• Porcentaje de respuesta para cada alternativa. • Correlación ítem-total de cada alternativa. • Índice de dificultad del ítem (porcentaje de aciertos). • Capacidad discriminativa del ítem (correlación ítem-total corregida). • Curvas empíricas de respuesta al ítem. • Porcentaje de omisión por ítems.	• Discriminación. • Dificultad. • Azar. • RP67. • Indicador de comportamiento diferencial del ítem (por género, ruralidad y GSE).

con el valor propio de mayor magnitud, y se finaliza con el menor.

²⁵ Los *scree plot* de la aplicación 2017 se encuentran disponibles en www.agenciaeducacion.cl índice Simce/bases de datos nacionales 2017, a través de petición formal en el formulario OIRS, presente en la misma web.

²⁶ Los parámetros de los ítems y curvas características y de información de las pruebas se encuentran disponibles en www.agenciaeducacion.cl, índice Simce/bases de datos nacionales 2017, a través de petición formal con el formulario OIRS, presente en la misma web.

²⁷ Adicionalmente se realizó una revisión de la correlación ítem-total de cada alternativa, con el objetivo de detectar posibles errores en la asignación de claves y remediar la situación antes de la etapa de calibración.

Tabla 1.3.7

Indicadores 2017 para preguntas de respuesta abierta según Teoría Clásica e IRT.

Teoría Clásica	IRT
- Distribución de respuestas por código de puntuación. - Capacidad discriminativa del ítem (correlación ítem-total, corregida).	- Discriminación. - Dificultad. - RP67. - Indicador de comportamiento diferencial del ítem (por género, ruralidad y GSE).

A continuación, se describen los análisis psicométricos realizados para puntuar a los estudiantes y clasificarlos según estándares en función de sus respuestas a las pruebas censales.

3.3.1 Análisis psicométricos de las pruebas Simce censales

Para el análisis psicométrico de las pruebas Simce censales se utiliza como metodología base un modelo IRT. Para la puntuación de los estudiantes se emplea la estimación Expected A Posteriori (EAP), que, de acuerdo con lo planteado por Bock y Mislevy (1982), presenta diversas ventajas respecto de otros estimadores, como los basados en Máxima Verosimilitud (ML) o Máxima Verosimilitud a Posteriori (MAP). Entre estas ventajas destacan la facilidad de cálculo (dado que no es un proceso iterativo y no es necesario obtener la derivada de la función de verosimilitud) y la estabilidad para pruebas con extensiones diversas. Adicionalmente, según lo señalado por ACER (2013), la estimación EAP permite distinguir, con mayor precisión que otras, el promedio de agrupamiento de estudiantes.

Bajo el modelo 3PL, la probabilidad de responder correctamente está dada por:

$$P_i = c_i + (1 - c_i) \frac{\exp[Da_i(\theta - b_i)]}{1 + \exp[Da_i(\theta - b_i)]}$$

Donde:

- a_i, b_i y c_i = parámetros del ítem i
- D = constante de escalamiento

En función de lo anterior, la probabilidad de respuesta X_i a un ítem está dada por:

$$\begin{aligned}
 P(X_i | \theta, a_i, b_i, c_i) &= P(X_i = 1 | \theta, a_i, b_i, c_i) \cdot P(X_i = 0 | \theta, a_i, b_i, c_i) \\
 &= (P_i)^{X_i} \cdot (1 - P_i)^{(1 - X_i)} \\
 &= (P_i)^{X_i} \cdot (Q_i)^{(1 - X_i)}
 \end{aligned}$$

Considerando las respuestas de N examinados a los K ítems, la función de verosimilitud se calcula en función de lo siguiente:

$$L(x_{11}, x_{12}, \dots, x_{ji}, \dots, x_{NK} | \Theta, a_i, b_i, c_i) = \prod \prod (P_{ji})^{x_{ji}} \cdot (Q_{ji})^{(1-x_{ji})}$$

De este modo, para obtener las estimaciones de las habilidades y de los parámetros es necesario resolver las siguientes ecuaciones:

$$\frac{\partial \ln L}{\partial \Theta_j} = 0, j = 1, \dots, N$$

$$\frac{\partial \ln L}{\partial a_i} = 0, i = 1, \dots, K$$

$$\frac{\partial \ln L}{\partial b_i} = 0, i = 1, \dots, K$$

$$\frac{\partial \ln L}{\partial c_i} = 0, i = 1, \dots, K$$

El análisis psicométrico de las pruebas Simce censales 2017 se hizo en cuatro etapas.

En la primera etapa se realizó una estimación de parámetros de los ítems que componen cada prueba, sin considerar ningún anclaje con la medición anterior (en caso de que existiera). En términos de equiparación entre formas de la prueba, se optó por una calibración concurrente mediante ítems comunes entre ellas. Para las preguntas cerradas se utilizó una calibración bajo el modelo logístico 3PL, con constante de escalamiento 1,7, dado que para este valor el modelo se aproxima a la curva normal acumulada (Martínez, 2005; Camilli, 1994). Los ítems de respuesta abierta dicotómica, por su parte, se analizaron bajo el modelo 2PL, porque el efecto azar no necesita ser estimado.

Por su parte, las preguntas abiertas de respuesta politómica se analizaron bajo el Modelo de Crédito Parcial Generalizado (MCPG), que es similar al modelo 2PL, pero permite tener más categorías. Todo lo anterior se realizó con el programa PARSCALE 4 (2003). Desde el punto de vista técnico, la calibración se realizó con el método ML, a través del algoritmo *Expectation Maximization* (EM), especificando un máximo de 100 iteraciones, con 41 puntos de cuadratura y un criterio de convergencia de 0,005. Cada ítem, independientemente de su formato, se especificó en un bloque distinto (comando BLOCK). Los criterios que se debieron satisfacer para determinar la solución fueron los siguientes: (a) la función converge en menos de 100 iteraciones y (b) no existen errores estándar de dificultad mayores a 0,210.

En la segunda etapa, las preguntas o ítems de *equating* se compararon en términos de los parámetros estimados mediante la calibración libre (primera etapa) y de los parámetros obtenidos en el análisis de las pruebas anteriores. Este análisis de variación de los parámetros (conocido como *drift*, por su denominación en inglés) permitió determinar cómo los parámetros de los ítems se afectan con el paso del tiempo. Para detectar la posible presencia de un efecto z , se utilizó la metodología de distancias perpendiculares (Crocker y Algina, 1986), según la cual un ítem es removido del conjunto de equiparación si la distancia estandarizada a la recta, en valor absoluto, es mayor a 2 para los parámetros de dificultad o de discriminación. Los ítems removidos o desanclados son incorporados al análisis de la prueba como ítems nuevos.

Una vez obtenido el conjunto de ítems de equiparación que no presentaba *drift* entre las dos mediciones comparadas, se inició la tercera etapa, que consistió en la equiparación de las pruebas definitivas con la última medición de la serie, con el propósito de conseguir estimaciones en una métrica comparable entre ambas mediciones. En esta etapa, se recalibró la prueba, importando como parámetros fijos los parámetros de los ítems de anclaje de las pruebas definitivas y los restantes ítems se ajustaron a la métrica de estos.

Para obtener la calibración óptima de las pruebas definitivas con parámetros anclados a la medición anterior, se requirió que el conjunto de preguntas satisficiera criterios aún más exigentes que los utilizados en la calibración libre. Estos criterios fueron: (a) convergencia de la función en menos de 100 iteraciones; (b) errores estándar de la dificultad, menores a 0,210; (c) azar menor o igual a 0,35, en los ítems de cuatro alternativas y menor o igual a 0,45, en los de tres alternativas; (d) discriminación mayor o igual a 0,6; y (e) parámetros de dificultad entre -2,4 y 2,4.

Por último, luego de obtener estimaciones estables en los parámetros de las pruebas anclados a la última medición de la serie, se procedió a ejecutar la cuarta y última etapa, consistente en el cálculo de la puntuación de cada estudiante. Esta se realizó a partir del patrón de respuestas individual, fijando los parámetros de todas las preguntas o ítems ajustados y utilizando la estimación EAP. Para cada uno de los niveles y asignaturas evaluados, la habilidad estimada y su respectivo error de estimación fueron transformados a la escala Simce anclada²⁸, que posee una media de 250 puntos y una desviación estándar de 50 puntos. Adicionalmente, para la puntuación de los ejes de Lectura, Matemática, Ciencias Naturales e Historia, Geografía y Ciencias Sociales, se utilizó la solución final de la calibración de la prueba completa respectiva.

Una vez obtenidos los puntajes por eje, se usó una escala con una media de 5 puntos y una desviación estándar de 2, la que fue truncada en los extremos dejando un mínimo de 0 puntos y un máximo de 10 puntos para los puntajes individuales. La elección de esta escala permitió obtener diferencias significativas de aproximadamente 1 punto, lo que simplificó las comparaciones.

La particularidad del proceso 2017 se evidencia en la entrada en vigencia de las Bases Curriculares de 8.º básico, con ello el marco de referencia de la prueba cambió y, por tanto, para mantener la comparabilidad de la serie de resultados, es necesario ajustes en la metodología de análisis. Con este objetivo, se trabajó en una asesoría con *Educational Testing Service* (ETS) denominada "Estudios puente en evaluaciones Simce frente a cambios en las bases curriculares de 8.º básico y II medio", donde se abordó la metodología más adecuada para el propósito buscado y la forma de implementarlo.

Como información adicional a la estimación de parámetros obtenidos en la calibración de ítems, se calcula el índice denominado RP67, que corresponde a la habilidad necesaria para obtener 67 % de probabilidad de acertar la respuesta correcta de una pregunta o un ítem. Este indicador es utilizado para asociar los ítems a los Estándares de Aprendizaje y corresponde a una estimación de la dificultad reescalada a 0,67 (2/3) de probabilidad de acierto del ítem o alternativa, en el caso de ítems politómicos²⁹. El objetivo de calcular este valor es proveer información necesaria para establecer los futuros Estándares de Aprendizaje.

²⁸ La escala Simce anclada para cada nivel escolar se definió en su año de origen, cuando se comenzó a analizar en función de IRT. Para II medio el año de origen fue 1998; para 4.º básico, 1999; para 8.º básico, 2000; para 2.º básico, 2012, y para 6.º básico, 2013.

²⁹ En el caso de las preguntas o ítems de respuesta cerrada, esta técnica supone una complejidad añadida por el uso del modelo 3PL, en el que, a diferencia de los modelos 1PL (o Modelo Rasch) y 2PL, el parámetro de dificultad no representa directamente el valor en la escala de habilidad para el cual un estudiante cualquiera tiene 50 % de probabilidad de responder correctamente. En los modelos de 1PL y 2PL, el cálculo de la dificultad reescalada se obtiene mediante una interpolación directa, mientras en el caso del modelo logístico 3PL es preciso utilizar una ecuación.

La estimación del valor RP67 para ítems dicotómicos se obtiene mediante la siguiente ecuación, la cual es derivada a partir de la probabilidad de responder correctamente un ítem en el modelo 3PL:

$$\Theta_{3PL} = \left(\frac{1}{Da} \right) \cdot \ln \left(\frac{RP - c}{1 - RP} \right) + b$$

El procedimiento anterior es válido solo para preguntas cerradas y para preguntas abiertas con dos alternativas de puntuación. En el caso de los ítems de respuesta abierta puntuados de manera politómica, se calcula la dificultad reescalada utilizando diferentes ecuaciones, dependiendo de la cantidad de opciones, como lo establece el modelamiento de alternativas con el MCPG.

Finalmente, a modo de confirmación, se verifica la presencia de Funcionamiento Diferencial de los Ítems (DIF por sus siglas en inglés), según corresponda al grado evaluado (Educación Básica o Educación Media), aun cuando en el análisis experimental de las preguntas se haya descartado la presencia de sesgo en las variables género, ruralidad³⁰ y grupo socioeconómico (GSE). Para calcular DIF se considera como grupo de referencia aquel que, en teoría, podría estar aventajado (en género, los hombres; en ruralidad, los establecimientos urbanos y en GSE, el estrato más alto) y como indicador de DIF se contempla una diferencia del parámetro, en valor absoluto, igual o mayor a 0,4³¹.

3.3.2 Estudio puente 8.º básico

Un estudio puente es un conjunto de métodos que permiten establecer la conexión o relación estadística entre las escalas provenientes de dos sistemas de medida distintos, pero que se supone miden un constructo que se considera común, en este caso las pruebas de 8.º básico 2015, referidas al Marco Curricular ajustado en 2009, y las pruebas de 8.º básico 2017, referidas a las Bases Curriculares vigentes.

En la práctica, en el proceso 2017 una muestra del 5 % de los estudiantes (10 000 aproximadamente) rindieron las pruebas 2015 y se procedió de la siguiente manera:

1. Se calibró conjuntamente la muestra de estudiantes 2017 con la población que rindió la prueba en 2015. Se utilizaron ponderadores para que cada grupo tuviera el mismo peso (10 000). Se emplearon los criterios tradicionales de calibración.
2. Se calibró de manera libre la prueba 2017, siguiendo los criterios tradicionales.
3. Se ajustó la distribución de la población 2015 y la muestra 2017, y lo reportado en 2015 (media y desviación estándar).
4. Con la métrica obtenida en el punto anterior para la muestra 2017, se ajustó la distribución (media y desviación estándar) de la población 2017.
5. Se reportaron los resultados.

Con esto se realizó un puente entre la prueba antigua y la nueva a través de la aplicación de las pruebas en una población común.

³⁰ En el medio no se realiza análisis DIF por ruralidad, debido a que la población de establecimientos rurales en ese nivel escolar es casi nula.

³¹ Dado que el programa Parscale realiza una estandarización del parámetro de dificultad del grupo de referencia, este valor representa la diferencia en porcentajes de desviaciones estándar.

3.3.3 Clasificación por Estándares de Aprendizaje

En las asignaturas y niveles escolares en los que existen Estándares de Aprendizaje, los resultados se reportan en función de estos. Es decir, junto con la puntuación, se clasifica a los estudiantes según niveles que describen lo que los estudiantes saben y pueden hacer en las evaluaciones Simce, y constituyen un referente de distintos grados de cumplimiento de los objetivos de aprendizaje establecidos en el currículo nacional. Estos estándares comprenden tres niveles de aprendizaje (Adecuado, Elemental e Insuficiente) que son inclusivos, es decir, un estudiante que alcanza el nivel Adecuado no solo demuestra que cumple con los requisitos establecidos para dicho nivel, sino que además cumple con los del nivel Elemental.

La clasificación de los resultados 2017 en función de Estándares de Aprendizaje se realizó en 4.º básico y II medio, ya que al existir un cambio curricular en 8.º básico, los estándares de este nivel quedaron desactualizados y no pueden ser utilizados.

Para la clasificación en estándares se utilizan puntajes de corte (véanse los utilizados en la aplicación 2017, en la Tabla 1.3.8).

Tabla 1.3.8

Puntajes de corte para la clasificación 2017 según Estándares de Aprendizaje.

Nivel escolar	Prueba	Insuficiente	Elemental	Adecuado
4.º básico	Lectura	$X < 241$ puntos	$241 \leq X < 284$ puntos	$X \geq 284$ puntos
	Matemática	$X < 245$ puntos	$245 \leq X < 295$ puntos	$X \geq 295$ puntos
II medio	Lectura	$X < 250$ puntos	$250 \leq X < 295$ puntos	$X \geq 295$ puntos
	Matemática	$X < 252$ puntos	$252 \leq X < 319$ puntos	$X \geq 319$ puntos

Nota: X = puntaje obtenido por el estudiante.

3.3.4 Análisis psicométricos de las pruebas para Discapacidad Sensorial

Para los efectos de los análisis psicométricos, la prueba de Discapacidad Sensorial (Simce DS) 2017, requirió algunos tratamientos especiales que se describen a continuación.

En las pruebas Simce DS, para transformar las respuestas de los estudiantes en puntuaciones de rendimiento individual, garantizando la calidad psicométrica de los resultados obtenidos, se debió realizar inicialmente algunos análisis exploratorios para verificar tanto la frecuencia como la distribución de cada alternativa de respuesta y de las respuestas omitidas, además de obtener una primera aproximación del comportamiento de los datos, en cuanto al ajuste al modelo que se pretende utilizar. Para el caso de las preguntas cerradas se calcularon los siguientes indicadores:

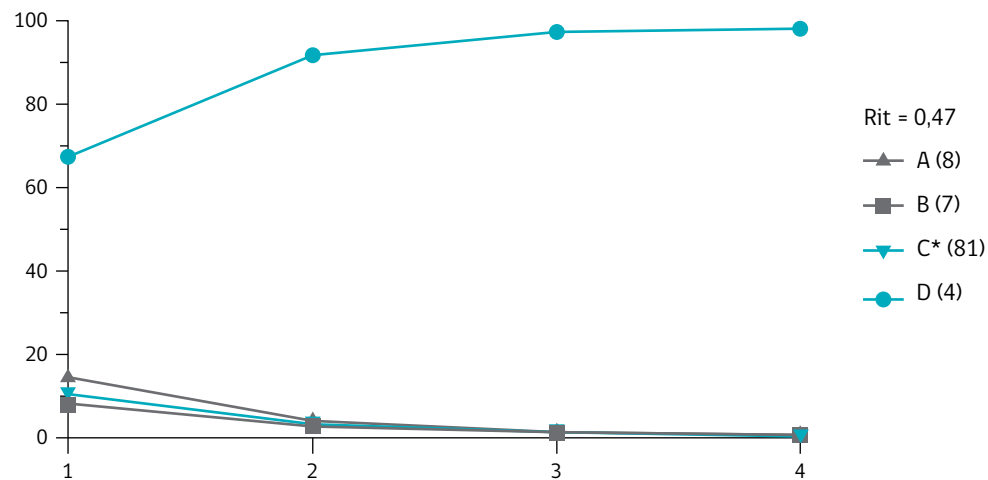
- Porcentaje de respuesta para cada alternativa.
- Correlación ítem-total de cada alternativa.
- Índice de dificultad del ítem (porcentaje de aciertos).
- Curvas empíricas de respuesta al ítem.
- Proporción de respuestas omitidas para cada ítem.

Luego, con el fin de optimizar el tiempo de procesamiento, se utilizaron las curvas empíricas generadas por el programa Tiaplus 2014, para las subpoblaciones (estudiantes con discapacidad visual parcial y discapacidad auditiva³²), cuya gráfica utiliza una curva para cada alternativa de respuesta señalando la que es correcta, como se muestra en la Figura 1.3.1.

Posteriormente, las puntuaciones de las pruebas Simce DS fueron calculadas en función del modelo de Teoría Clásica (por cuanto el reducido tamaño de las subpoblaciones de estudiantes que rindieron esta prueba no permite implementar IRT), en donde el puntaje observado del test se define como una combinación del puntaje verdadero más el componente de error. Esto implica que, a menor error, mayor es la relación entre el puntaje observado y el puntaje verdadero. Según este modelo, el puntaje verdadero de un sujeto está dado por $V=X+e$ (Lord y Novick, 2008), es decir, el puntaje verdadero (V) se estima a través de una combinación lineal de los puntajes observados en la medición (X) más un componente de error (e).

Figura 1.3.1

Ejemplo de curva empírica del ítem.



Una vez obtenida la estimación de habilidad individual, esta se transformó, utilizando una escala Simce de discapacidad sensorial³³, en la que se trabaja con los porcentajes de respuestas correctas (PRC) y se genera un rango de valores entre 0 y 100, donde cada puntuación se obtiene del siguiente modo:

$$PRC = \sum_{i=1}^N \frac{\text{respuesta } i \cdot 100}{N^{\circ} \text{ total preguntas}}$$

³² Para los estudiantes con discapacidad visual total, no se generaron estas curvas porque en 2017 solo siete estudiantes presentaron esa condición y esa cantidad no permite separar sus habilidades en subgrupos.

³³ Esta escala es diferente a las usadas en las pruebas Simce de aplicación regular, por cuanto en estas pruebas DS no solo los ítems han sufrido acomodaciones específicas, sino que además la población está constituida por un reducido número de estudiantes.

3.4 Análisis de datos agregados

Este tipo de análisis permite obtener resultados de distintas agregaciones que son de utilidad, tanto para informar la política pública como para comunicar los resultados a los actores del sistema educativo (docentes, equipos directivos, padres, etc.)³⁴. Entre estas agregaciones se cuentan nacional, regional, comunal, etc.

Estos análisis tienen dos objetivos centrales: (a) conocer la realidad actual del desempeño de los estudiantes del país y (b) monitorear su evolución a través del tiempo, comparándolos con resultados de evaluaciones anteriores. El resultado promedio de una agregación se obtiene promediando el puntaje de todos los estudiantes que la componen. Para lo primero, se realizan comparaciones de rendimiento en un determinado año de aplicación y, para lo segundo, comparaciones entre años, como se describe a continuación.

- Diferencia significativa (en el año): compara el rendimiento promedio obtenido en una agregación con el rendimiento promedio de otra agregación en la misma medición. Por ejemplo, se comparan los puntajes promedio de 4.º básico Matemática Simce 2017 de una región con el puntaje promedio nacional Simce 2017.
- Variación significativa (entre años): compara el rendimiento promedio obtenido en una agregación con el rendimiento promedio de la misma agregación de la medición anterior³⁵. Por ejemplo, se comparan los puntajes promedio Simce 2016 y Simce 2017 en la evaluación de 4.º básico Matemática.

En las pruebas que cuentan con Estándares de Aprendizaje, es posible, además, determinar la proporción de estudiantes de una agregación que se ubica en cada nivel: Adecuado, Elemental e Insuficiente.

Adicionalmente, los análisis agregados se realizan cruzando las variables de rendimiento (obtenidas de las pruebas) con otras variables, obtenidas de los cuestionarios, como la clasificación por GSE. Esta variable permite comparar el desempeño académico de una comunidad educativa respecto de otra, controlando las características socioeconómicas de cada una y obteniendo como resultado comparaciones más justas³⁶.

A continuación, se describen los análisis agregados realizados en las pruebas censales Simce.

³⁴ El detalle de este proceso se encuentra en el capítulo IV sobre comunicación de resultados.

³⁵ Es importante considerar que "medición anterior" refiere a la aplicación anterior de la misma prueba y nivel, que no es necesariamente el año anterior, puesto que cada evaluación y nivel tiene una ocurrencia específica.

³⁶ Para la construcción de los grupos de establecimientos similares según GSE se utiliza una técnica estadística de análisis de conglomerados o *clusters*. La clasificación de las escuelas, según esta metodología, se hace por nivel evaluado, incluso cuando los establecimientos imparten todos los grados evaluados. Para conocer más detalles del cálculo de GSE revisar www.agenciaeducacion.cl.

3.4.1 Análisis agregados de las pruebas censales Simce

A partir de la información obtenida de los estudiantes que rindieron las pruebas regulares Simce 2017, se realizaron los siguientes análisis:

- Porcentaje de estudiantes evaluados, a nivel nacional y por región: número total de estudiantes que rindió la prueba, sobre el total de estudiantes perteneciente al nivel evaluado del sistema.
- Porcentaje de establecimientos evaluados, a nivel nacional y por región: número total de establecimientos donde se aplicó cada prueba, sobre el total de establecimientos del sistema que imparte el nivel escolar evaluado.
- Porcentaje de asistencia anual a nivel nacional: porcentaje promedio de asistencia que tuvieron todos los estudiantes del nivel escolar evaluado, durante el año escolar en curso.
- Porcentaje de asistencia Simce a nivel nacional: porcentaje de estudiantes que rinden la prueba sobre el total de estudiantes matriculados en el nivel escolar evaluado.
- Distribución de estudiantes evaluados: número de estudiantes evaluados por agregación. Se calcula para las agregaciones: nacional, regional, comunal, provincial, por departamento provincial, género, dependencia administrativa y GSE.
- Distribución de establecimientos evaluados: número de establecimientos, por agregación, que contaron con estudiantes evaluados. Se calcula para las agregaciones: nacional, regional, comunal, provincial, departamento provincial, género, dependencia administrativa y GSE.

Junto con lo anterior, se calculan el puntaje promedio y la distribución según Estándares de Aprendizaje, para cada asignatura en cada nivel escolar evaluado³⁷, según las agregaciones de interés. Adicionalmente, tanto para la comparación de puntajes promedio como para la estimación de la proporción de estudiantes en cada nivel de los Estándares de Aprendizaje, se calcula la significancia estadística de las diferencias y variaciones entre agregaciones.

En la Tabla 1.3.9 se muestra una síntesis de los análisis agregados que se realizan, tanto a nivel de puntaje, como a nivel de estándares.

Tabla 1.3.9

Agrupación de análisis agregados de las pruebas Simce 2017.

Análisis	Descripción
Puntaje	Puntaje promedio de una agregación.
	Diferencia y significancia entre puntajes promedio de dos agregaciones en el mismo año.
	Variación y significancia entre puntajes promedio de una agregación, obtenida en dos evaluaciones consecutivas.
Estándares de aprendizaje	Distribución en Estándares de Aprendizaje de una agregación.
	Diferencia y significancia de la distribución en Estándares de Aprendizaje de dos agregaciones en el mismo año.
	Variación y significancia entre la distribución en Estándares de Aprendizaje de una agregación obtenida en dos evaluaciones consecutivas.

³⁷ Es importante destacar que Simce reporta resultados por cada una de las asignaturas evaluadas en un determinado nivel escolar, porque cada asignatura tiene una escala propia, por lo que resulta técnicamente imposible generar análisis combinándolas.

Todos los análisis antes señalados se rigen por filtros y restricciones. Los filtros tienen por objeto definir la población estadísticamente comparable. Por ejemplo, al comparar los resultados Simce 2016 y 2017 en 4.º básico Matemática, según GSE y dependencia administrativa, se considera solo a los establecimientos que en ambas evaluaciones mantuvieron su GSE y dependencia. Las restricciones, por su parte, determinan la publicación de datos y por lo tanto son aplicadas de forma posterior al cálculo de los análisis agregados para resguardar la confidencialidad y robustez de estos. Por ejemplo, no se publican resultados según Estándares de Aprendizaje para aquellos establecimientos que hubieran tenido menos de diez estudiantes evaluados en una determinada asignatura y nivel escolar.

3.4.2 Análisis agregados de las pruebas para Discapacidad Sensorial

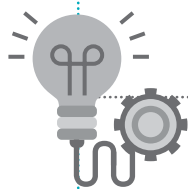
Al igual que para los análisis psicométricos, existen diferencias entre los análisis agregados que se realizan para las pruebas censales y los que se ejecutan sobre las pruebas Simce DS.

Para la prueba Simce DS 2017, los análisis agregados reportaron el porcentaje promedio de respuestas correctas a nivel nacional para cada tipo de discapacidad: visual total, visual parcial y auditiva, en cada una de las pruebas: Lectura y Matemática. Este dato se entregó agregado también por tipo de establecimiento (regular/especial) y dependencia administrativa, por género de los estudiantes y por macrozona³⁸.

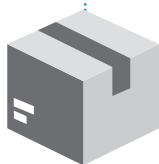
3.4.3 Análisis de tendencias

Como se mencionó al comienzo de este apartado, junto con conocer la realidad actual del desempeño de los estudiantes del país, los análisis agregados buscan entregar información que permita monitorear su evolución a través del tiempo, comparándolos con resultados de evaluaciones anteriores. En función de lo indicado recientemente, se realizan análisis que muestran las tendencias de resultados a nivel nacional cada año, para los distintos niveles escolares evaluados, en las diversas asignaturas medidas. Estos análisis permiten observar la evolución de puntajes, a partir de una medición de referencia, que generalmente es una década.

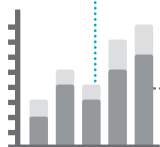
³⁸ Para los análisis antes descritos se implementaron dos metodologías no paramétricas: la prueba U de Mann-Whitney, para las diferencias entre grupos independientes; y la prueba de Rangos Asignados de Wilcoxon, para la comparación de zona geográfica.



Capítulo I:
Diseño y construcción de las
pruebas Simce 2017



Capítulo II:
Logística y aplicación de las
pruebas Simce 2017



Capítulo III:
Procesamiento y análisis
de datos 2017



Capítulo IV:
Comunicación de resultados
educativos 2017

Capítulo IV: Comunicación de resultados educativos 2017

En atención a que la Ley que establece el Sistema Nacional de Aseguramiento de la Calidad de la Educación señala que la Agencia de Calidad debe “proporcionar información en materias de su competencia a la comunidad en general y promover su correcto uso” (Ley N.º 20.529, art. 10, letra e), la comunicación de los resultados Simce constituye uno de sus mandatos centrales.

Al respecto, la Agencia ha establecido los siguientes objetivos estratégicos³⁹:

- Proveer información útil y orientaciones para el trabajo pedagógico e institucional en los establecimientos, con foco en el aprendizaje de los estudiantes.
- Entregar orientaciones técnicas, pedagógicas y de gestión a todos los establecimientos, para instalar capacidades en las comunidades escolares orientadas al mejoramiento de los procesos institucionales y los resultados educativos.
- Entregar información para la toma de decisiones y la articulación de los distintos actores del sistema escolar, así como impulsar el trabajo colaborativo y en red.
- Contribuir al sistema a partir de su carácter técnico, basado en evidencia, independiente, transfiriendo la información generada a partir de las evaluaciones y estudios que realiza la Agencia, para influir en la toma de decisiones en materia de política educativa, identificando fortalezas y debilidades de los procesos y resultados educacionales.

Para cumplir en parte con estos objetivos, la Agencia realiza una serie de publicaciones asociadas a las aplicaciones Simce que buscan informar y orientar sobre la evaluación y el uso de sus resultados a los distintos actores del sistema educativo, entre los que se cuentan: docentes, directivos, sostenedores de establecimientos, padres y apoderados, líderes en política educativa e investigadores.

Antes de la aplicación 2017 de las pruebas Simce, la Agencia publicó en la página web institucional los Programas de Evaluaciones Educativas para Docentes y los Programas de Evaluaciones Educativas para Directivos, con el propósito de entregar información relevante sobre la evaluación. También se publicaron programas de evaluaciones educativas para directivos de escuelas especiales, de establecimientos que aplicaron pruebas experimentales y de establecimientos con cursos que rinden las pruebas y poseen menos de 10 estudiantes.

Los documentos Programas de Evaluaciones Educativas para Docentes Educación Básica y Educación Media, tuvieron como propósito informar a los profesores el cronograma de aplicación; los niveles, asignaturas y aprendizajes curriculares evaluados; y las características generales de las pruebas y de los cuestionarios.

Los Programas de Evaluaciones Educativas para Directivos Educación Básica y Educación Media, por su parte, tuvieron como propósito informar a los directores sobre el proceso de aplicación y favorecer su colaboración en el desarrollo regular de este. Estos documentos incluyeron el cronograma de aplicación de las pruebas y cuestionarios; información específica sobre los estudiantes que rendirían

³⁹ Disponibles en <https://www.agenciaeducacion.cl/nosotros/objetivos-estrategicos/>.

las evaluaciones (considerando a aquellos que podrían no rendirlas por presentar algún impedimento) y procedimientos para los cuales se requeriría la colaboración de los directivos del establecimiento.

Posterior a la evaluación Simce, los docentes, directivos y sostenedores reciben información sobre los resultados de las pruebas y de los cuestionarios, así como orientaciones para el uso pedagógico de la información, de manera de impulsar la planificación de acciones concretas que aporten al proceso de mejora continua de la calidad de la educación impartida. A los padres y apoderados, por su parte, se les entrega información sintetizada y organizada de manera sencilla, para facilitar la comprensión por parte de todos y, de ese modo, favorecer su interés por involucrarse de manera directa en el sistema. Finalmente, los líderes en política educativa y los investigadores, aunque tienen acceso a la información general que se publica a nivel nacional, también pueden acceder, si lo solicitan formalmente, a información específica (que incluye bases de datos) sobre los resultados de una o más pruebas Simce y sobre los Indicadores de Desarrollo Personal y Social (IDPS). Lo anterior, con el objeto de que puedan desarrollar sus proyectos contando con información de la fuente directa.

En este capítulo se describen los principales documentos comunicacionales que publicó la Agencia para comunicar los resultados educativos en el contexto de las pruebas censales Simce 2017.

4.1 Informes de resultados

Después de la aplicación Simce 2017, la Agencia publicó informes con los resultados de aprendizaje para Educación Básica y Educación Media. Ello, con el propósito de propiciar la realización de diagnósticos más completos e integrales de los aprendizajes de los estudiantes.

Los informes de resultados contemplaron tres versiones inicialmente: una para docentes y directivos; una para padres y apoderados, y una para sostenedores. Estos informes se evalúan cada año por las audiencias respectivas, con el objeto de mejorar la calidad de los productos comunicacionales y fomentar el uso adecuado de la información. En estos procesos de evaluación se determina, por ejemplo, la necesidad de incorporar nueva información, de cambiar el enfoque o de profundizar en ciertos contenidos.

4.1.1 Informe de Resultados Educativos para Docentes y Directivos

El Informe de Resultados para Docentes y Directivos tiene por objetivo comunicar los resultados educativos (de aprendizaje y de los IDPS) del conjunto de estudiantes del establecimiento, en las distintas asignaturas y niveles escolares evaluados, y entregar orientaciones para una correcta interpretación y uso de estos⁴⁰.

En este documento se entregan los siguientes resultados educativos:

⁴⁰ De manera complementaria, se publicaron dos documentos (uno para Educación Básica y otro para Educación Media) con algunos ejemplos de preguntas respondidas por los estudiantes en las pruebas Simce, con el propósito de contribuir al análisis que realiza el equipo docente de sus evaluaciones de aula y de que sean utilizados para conocer diferentes maneras de evaluar una determinada habilidad o conocimiento.

Resultados de Indicadores de Desarrollo Personal y Social (IDPS)

- Puntaje en cada indicador y sus dimensiones; distribución de las respuestas en cada nivel de los indicadores.
- Resultados según grupo socioeconómico (GSE) y género.

Resultados de aprendizaje Simce:

- Puntajes promedio en cada prueba, distribución según Niveles de Aprendizaje en 4.º básico y resultados según ejes.
- Tendencias de puntaje promedio del establecimiento a través del tiempo.
- Resultados según grupo socioeconómico (GSE) y género.
- Puntajes promedio por curso y puntaje según eje, solo en el caso del informe privado que descargan los directores con su Clave Única.

En el año 2017, al igual que en 2016, se incluyó un capítulo de orientaciones para el uso de resultados educativos, con recomendaciones para equipos directivos y docentes, relacionadas con la mejora de la gestión institucional y la gestión pedagógica. También se presentó una propuesta de plan de trabajo para que los equipos directivos, en conjunto con los docentes, puedan analizar los resultados educativos y relacionarlos con datos internos de gestión, con el fin de contextualizarlos en la realidad del establecimiento.

4.1.2 Informe de Resultados Educativos para Padres y Apoderados

El Informe de Resultados para Padres y Apoderados resumió los resultados educativos a las familias de los estudiantes, con el propósito de involucrar a los padres y apoderados en el proceso educativo y convocarlos a participar activamente en él. Este informe incluye resultados de cada IDPS (escala de 0 a 100 puntos) comparados con otros establecimientos de similares características socioeconómicas; resultados de aprendizaje Simce comparados con establecimientos de similares características socioeconómicas; y distribución de los estudiantes en niveles de desempeño, en las asignaturas evaluadas que cuentan con Estándares de Aprendizaje. Asimismo, entrega la clasificación de la Categoría de Desempeño que le corresponde al establecimiento en el ciclo anterior⁴¹, y la descripción de lo que considera la herramienta con que se obtiene dicha clasificación.

Junto con lo anterior, se entregó a los padres y apoderados información general sobre las funciones de la Agencia y recomendaciones para participar activamente en el proceso de aprendizaje de sus pupilos.

⁴¹ Refiere a la clasificación notificada en diciembre de 2017 a los sostenedores según las evaluaciones del año anterior al informe.

4.1.3 Informe de Resultados Educativos para Sostenedores

En tanto, el Informe de Resultados para Sostenedores incluyó los principales resultados educativos, tanto a nivel general de su administración (Capítulo 1) como el detalle de cada uno de los establecimientos de tal administración (Capítulo 2). Además, se incorporaron orientaciones para el correcto análisis y uso de la información.

4.1.4 Informe de Resultados Educativos para Servicios Locales de Educación Pública

Posteriormente, con la implementación de la Ley N.º 21040 que crea el Sistema de la Educación Pública, se elaboró un informe para los primeros Servicios Locales de Educación que comenzaron a funcionar en nuestro país: Barrancas, Costa Araucanía, Puerto Cordillera y Huasco.

Este informe respondió a las necesidades vigentes al momento en que los Servicios entraron en funcionamiento, y su propósito era servir como diagnóstico para que los equipos administrativos, técnicos y pedagógicos tuvieran una mirada integral del territorio y focalizaran sus esfuerzos en aquellos aspectos que requerían ser atendidos de manera más urgente. Dado este contexto, este documento fue enviado en dos oportunidades a los jefes de los Servicios Locales: en marzo, con resultados hasta el año 2016 y en junio, cuando se agregaron los resultados 2017.

El informe entrega los siguientes resultados educativos:

Características territoriales del Servicio Local de Educación Pública

- Dependencia administrativa y financiera.
- Matrícula.
- Asistencia escolar.
- Indicadores de GSE y de Índice de Vulnerabilidad Escolar.
- Resultados de la Evaluación Docente.

Resultados Educativos agrupados por Servicio Local y por comunas

- Categoría de Desempeño del último año.
- Puntaje en cada indicador y distribución de las respuestas en cada nivel de los indicadores.
- Resultados Simce: puntajes promedio en cada prueba; tendencias de puntaje promedio del establecimiento a través del tiempo; distribución según Niveles de Aprendizaje, GSE y género.

Orientaciones para el análisis y uso de resultados educativos del Servicio Local

- Propósitos a los que aporta el uso de datos.
- Condiciones que inciden en el proceso de uso y transformación de datos en conocimiento.
- Orientaciones generales para trabajar con los datos en los establecimientos del Servicio Local.

4.2 Otros instrumentos comunicacionales

Los resultados nacionales se comunican y difunden a la opinión pública a través de una conferencia de prensa, instancia en la que se presentan los principales resultados de aprendizaje y de los Indicadores de Desarrollo Personal y Social. Al mismo tiempo que se informan los resultados de cada establecimiento del país, a través de los documentos indicados en el apartado anterior, la Agencia pone a disposición de la comunidad otros dispositivos e insumos que se describen a continuación:

4.2.1 Chatbot

Para avanzar en la personalización de la información y orientación, durante el 2018 y para la entrega de resultados de las evaluaciones 2017, se incorporó en la web una herramienta de innovación digital, el Chatbot, un asistente virtual conversacional de inteligencia artificial, cuyo objetivo de creación es facilitar la búsqueda de resultados educativos por tipo de establecimiento, nivel y comuna, al entregar un enlace directo para la descarga de los resultados educativos y así facilitar la experiencia de búsqueda.

4.2.2 Localiza tu colegio

Al igual que el Chatbot, para la entrega de resultados de las evaluaciones 2017, a inicios del 2018, se creó la herramienta Localiza que permite buscar establecimientos por territorio y cercanía, logrando una mirada integral de resultados educativos georeferenciados. Su innovación radica en que su formato digital tiene bajo impacto ambiental y crea un mayor bienestar para los padres/apoderados y toda la sociedad, al contar con la información de los establecimientos cercanos geográficamente al momento de realizar la búsqueda de un nuevo establecimiento o para conocer cómo se encuentra el establecimiento en comparación con el resto de la comuna.

4.2.3 Resúmenes ejecutivos

Los resúmenes ejecutivos tienen el propósito de entregar una síntesis de los principales resultados de las evaluaciones Simce a la audiencia general, a través de los medios de comunicación masivos, con el objeto de informar y transparentar los procesos, además de propiciar el análisis y debate en torno a la educación del país.

Para la aplicación 2017 se publicaron tres resúmenes ejecutivos:

- **Minuta comunicacional:** Este documento contiene una síntesis de los principales resultados de aprendizaje y de los Indicadores de desarrollo personal y social 2017, además de algunas conclusiones generales obtenidas a partir de dichos resultados. Este resumen ejecutivo se utilizó en las conferencias de prensa realizadas en los meses de mayo (conferencia de prensa Resultados Nacionales) y septiembre de 2018 (2.º básico Comprensión de Lectura). Los videos de estas conferencias fueron transmitidos vía *streaming* y publicados en el sitio web de la Agencia.

- **Resúmenes ejecutivos para las conferencias de prensa regionales:** Estos documentos, uno por región, contienen una síntesis de los principales resultados Simce obtenidos por los estudiantes de la región. El público objetivo de estos documentos informativos son los secretarios regionales ministeriales y los jefes de los Departamentos Provinciales de Educación, a quienes se les envía la información por correo directo. Esta información luego fue publicada en el sitio web de la Agencia.
- **Presentación de resultados educativos para directores:** Junto con la entrega del Informe de Resultados para Docentes y Directivos, se entregó este documento que resume los principales resultados de cada establecimiento (puntaje promedio, distribución de estudiantes en cada nivel de aprendizaje, tendencias de resultados, puntaje promedio por curso evaluado y comparación con establecimientos similares, género y GSE) con el objeto de facilitar a los directores la presentación de los resultados a su equipo directivo, a los docentes y a la comunidad educativa en general.

4.2.4 Material de apoyo pedagógico: Guía Metodológica de Uso de Datos

Para contribuir a un mayor uso de los resultados educativos orientados a la toma de decisiones de mejora escolar, desde la Agencia de Calidad de la Educación se elaboró y presentó a las comunidades educativas la Guía Metodológica para el Uso de Datos, herramienta que propone una aproximación práctica a las tareas esenciales que se generan a partir de la evaluación.

Esta guía busca orientar a las escuelas sobre el uso de los datos educativos, tanto los que entrega la Agencia como los que se generan dentro del mismo establecimiento, con el objetivo de que sean transformados en conocimiento útil para la toma de decisiones que impacten positivamente en el mejoramiento educativo.

4.2.5 Bases de datos

En relación con las bases de datos, cabe señalar que la Agencia provee de dos tipos de bases de datos: una pública y una privada (esta última para investigadores). La base de datos pública cuenta con los resultados por establecimiento, comuna, provincia, región, GSE y dependencia administrativa. La versión 2017 de esta base se puso a disposición de la comunidad en agosto de 2018 en el sitio web de la Agencia. Por otro lado, el acceso a la base de datos con información específica de los resultados educativos (sin individualización de estudiantes, profesores ni apoderados), se entregó a los investigadores que la requirieron, previa solicitud formal a la Agencia; en esta se debían especificar los fines del estudio.

La difusión de todos los documentos descritos en este capítulo, referidos a la aplicación 2017, se realizó a través de medios impresos y plataformas digitales. Entre estas últimas se cuentan: el sitio web de la Agencia; las redes sociales oficiales (Twitter y Facebook de la Agencia); correos electrónicos masivos y personales, y plataforma *streaming* en el caso de las videoconferencias. Esta socialización mediante diversos medios de comunicación y difusión, busca asegurar un acceso efectivo y expedito a la información por parte de los distintos actores del sistema educativo.

4.3 Síntesis de productos comunicacionales Simce 2017

A continuación, se señalan los productos comunicacionales de la aplicación Simce 2017.

Tabla 1.4.1

Documentos comunicacionales de la aplicación Simce 2017.

Tipo de documento	Título de la publicación	Soporte	Fecha de la publicación
Programas de Evaluaciones Educativas	Programas de Evaluaciones Educativas Educación Básica para Docentes	Digital	Marzo 2017
	Programas de Evaluaciones Educativas Educación Media para Docentes	Digital	Marzo 2017
	Programas de Evaluaciones Educativas Educación Básica para Directores	Digital	Septiembre 2017
	Programa de Evaluaciones Educativas Educación Media para Directores	Digital	Septiembre 2017
	Programa de Evaluaciones Educativas para Directores Pruebas Experimentales Educación Básica y Media	Digital	Septiembre 2017
	Programa de Evaluaciones Educativas para Directores Escuelas Especiales Educación Básica	Digital	Septiembre 2017
	Programa de Evaluaciones Educativas para Directores Prueba Extendida Educación Básica	Digital	Septiembre 2017
Informes de Resultados Educativos	Informe Servicio Local de Educación Pública	Digital	Marzo 2018
	Informe de Resultados Educativos para Docentes y Directivos Educación Básica	Digital e impreso	Mayo 2018
	Informe de Resultados Educativos para Docentes y Directivos Educación Media	Digital e impreso	Mayo 2018
	Informe de Resultados Educativos para Padres y Apoderados 4° Básico	Digital e impreso	Mayo 2018
	Informe de Resultados Educativos para Padres y Apoderados 8° Básico	Digital e impreso	Mayo 2018
	Informe de Resultados Educativos para Padres y Apoderados II Medio	Digital e impreso	Mayo 2018
	Informe de Resultados Educativos para Sostenedores Educación Básica	Digital e impreso	Mayo 2018
	Informe de Resultados Educativos para Sostenedores Educación Media	Digital e impreso	Mayo 2018
	Informe Servicio Local de Educación Pública	Digital	Junio 2018
Otros documentos comunicacionales	PPT Directores (4°, 8° Básico y II Medio)	Digital	Mayo 2018
	PPT Resultados Regionales	Digital	Mayo 2018



**Estudios Nacionales
Muestrales**

Introducción

Los Estudios Nacionales Muestrales evalúan el aprendizaje de los estudiantes del país en áreas curriculares relevantes para monitorear los logros del sistema educativo e informar a la política pública. Lo anterior, permite al Estado rendir cuentas sobre la efectividad de sus políticas y programas. Estos estudios se componen de evaluaciones que se aplican a una muestra representativa de establecimientos educacionales.

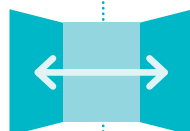
La información obtenida a través de los resultados de estos estudios se pone a disposición de la comunidad para su conocimiento y uso. Para ello, se informan resultados a nivel nacional y de grupos de interés, lo que permite focalizar recursos, identificar fortalezas y aspectos a mejorar a nivel nacional y en los diferentes grupos de interés.

De acuerdo con el Plan de Evaluaciones Nacionales e Internacionales 2016-2020, en el año 2017 se realizaron los Estudios Nacionales de carácter muestral en las asignaturas de Lectura, Formación Ciudadana e Inglés, aplicados a 2.º básico, 8.º básico y III medio, respectivamente.

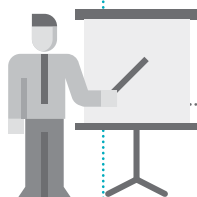
Estudios nacionales muestrales aplicados en 2017

Estudio	Nivel
Lectura	2.º básico
Formación Ciudadana	8.º básico
Inglés	III medio

El ciclo evaluativo de los Estudios Nacionales Muestrales 2017 constó de cuatro etapas: diseño y construcción de los instrumentos; logística y aplicación de las pruebas; procesamiento y análisis de los datos; y comunicación de los resultados educativos. Además, se incluye la instancia de diseño y selección de la muestra.



Capítulo I:
Visión panorámica de los Estudios
Nacionales Muestrales 2017



Capítulo II:
Descripción de los Estudios
Nacionales Muestrales 2017

Capítulo I: Visión panorámica de los Estudios Nacionales Muestrales 2017

a. Diseño y construcción de las pruebas

El diseño y la construcción de los instrumentos que se aplican en los Estudios Nacionales Muestrales se lleva a cabo mediante un proceso que dura aproximadamente un año y medio. En este proceso, al igual que con las evaluaciones Simce, se resguarda el cumplimiento de estándares internacionales, como los definidos en el documento *Standards for Educational and Psychological Testing* (AERA, APA y NCME, 2014). El diseño y construcción de las pruebas se inicia con la definición del Marco de Referencia y las especificaciones técnicas, continúa con la elaboración y validación de las preguntas, y culmina con el armado y diagramación de las formas de prueba definitivas que se aplican a los estudiantes.

b. Logística y aplicación de las pruebas

El proceso de logística y aplicación de los instrumentos de los Estudios Nacionales en territorio nacional consta de las mismas etapas que el de la aplicación Simce: empadronamiento; impresión, mecanizado y distribución del material de aplicación; aplicación de las pruebas; entrega del material para procesamiento. Sin embargo, para los Estudios Nacionales se incluye una etapa previa de selección de la muestra a la que se le aplicarán los instrumentos. Todas las etapas cumplen con los mismos criterios de calidad que en la aplicación Simce.

En términos de la población evaluada, los Estudios Nacionales fueron rendidos por una muestra de estudiantes en los niveles correspondientes. La cobertura general se presenta en la Tabla 2.1.1.

Tabla 2.1.1

Establecimientos y estudiantes evaluados en la aplicación de los Estudios Nacionales 2017.

Estudio nacional	N.º establecimientos en la muestra	% aplicación establecimientos de la muestra	N.º estudiantes en la muestra	N.º estudiantes evaluados	% estudiantes evaluados
Lectura 2.º básico	260	100 %	8 263	7 349	88,9 %
Formación Ciudadana 8.º básico	242	100 %	9 800	8 701	88,8 %
Inglés III medio	137	100 %	9 165	7 340	80,1 %

Al igual que con las evaluaciones Simce, la aplicación de los Estudios Nacionales se realiza mediante la externalización de servicios, adjudicados por licitación pública⁴². En 2017, estos estudios se aplicaron en las siguientes fechas:

- 2.º básico Lectura: 24 de octubre
- 8.º básico Formación Ciudadana: 9 de noviembre
- III medio Inglés: 13 al 15 de noviembre

c. Procesamiento y análisis de datos

En lo que se refiere al procesamiento y análisis de los datos de los Estudios Nacionales, este incluye las mismas etapas que en el de las evaluaciones Simce: Captura, validación y consolidación de datos; Corrección de preguntas abiertas; Análisis psicométrico de las pruebas; Análisis de datos agregados. Estas etapas se realizan bajo los mismos estándares que las evaluaciones Simce.

En la Tabla 2.1.2 se muestra el detalle de los tipos de documentos capturados en el proceso de aplicación 2017, en cada nivel escolar y prueba aplicada.

Tabla 2.1.2

Tipos de documentos capturados en 2017 por nivel escolar y prueba muestral aplicada.

Datos	2º básico	8º básico FC*	III medio
Lista Curso (LC)	✓	✓	✓
Formulario Aplicación (FA)	✓	✓	✓
Preguntas Cerradas (PC)	✓	✓	✓
Preguntas Abiertas (PA)	✓	✓	
Cuestionario de Padres (CP)	✓	✓	✓
Cuestionario de Estudiantes (CE)		✓	✓
Cuestionario Docentes	✓	✓	✓
Cuestionario Director		✓	

FC: Formación ciudadana, incluye DS.

⁴² Licitación pública "Aplicación de las Pruebas Simce año 2017 - 2018" ID 721703-1-LR17, para toda la extensión del territorio, desde la Región de Arica y Parinacota hasta la Región del Biobío, y mediante trato directo desde la Región de la Araucanía hasta Magallanes

En la Tabla 2.1.3 se muestra la cantidad de preguntas abiertas contenidas en las pruebas muestrales 2017, mientras que en la Tabla 2.1.4 se muestra la cantidad total de correctores que participó en el proceso 2017 y la cantidad de grupos de corrección conformados para cada asignatura y nivel escolar de las evaluaciones muestrales que contaron con este tipo de preguntas.

Tabla 2.1.3

Cantidad de preguntas abiertas pruebas muestrales 2017.

Nivel escolar	Prueba	N.º de preguntas abiertas
2.º básico	Lectura muestral	5
8.º básico	Formación Ciudadana muestral	11

Tabla 2.1.4

Distribución de correctores de preguntas abiertas pruebas muestrales 2017.

Nivel escolar	Prueba	N.º correctores (aproximado)	N.º grupos de correctores
2.º básico	Lectura muestral	36	3
8.º básico	Formación Ciudadana muestral	48	4

En lo que respecta a los análisis psicométricos de los Estudios Nacionales, estos se realizan sobre la base de lineamientos internacionales (AERA, APA y NCME, 2014) que aseguran la estandarización de los procesos y están sujetos a estrictos controles de calidad, incluyendo un procesamiento en paralelo, lo que permite detectar inconsistencias y verificarlas, de haberlas, así como ajustarlas de manera inmediata.

Los análisis psicométricos se realizan utilizando como insumo los archivos denominados Planilla Maestra y Base de Datos de Estudiantes, los que contienen la información de los ítems aplicados (identificación, objetivo de evaluación, respuesta correcta o clave, posición, etc.) y de las respuestas de los estudiantes, respectivamente. Las respuestas a las preguntas cerradas tienen dos puntuaciones posibles: 0 (incorrecta) y 1 (correcta), y las respuestas a preguntas abiertas pueden tener una puntuación igual a las cerradas (respuesta dicotómica) o bien una de tres niveles o politómica: 0 (incorrecta), 1 (parcialmente correcta) y 2 (correcta), dependiendo de lo que establezca la pauta de corrección respectiva.

Respecto del modelo que orienta los análisis psicométricos de las pruebas muestrales, en 2017 para realizar los análisis de las pruebas muestrales de lectura se usó el Modelo Logístico de Tres Parámetros (3PL) para las preguntas cerradas; el Modelo Logístico de dos Parámetros (2PL), para los ítems de respuesta abierta dicotómica; y el Modelo de Crédito Parcial Generalizado (MCPG), para los ítems de respuesta abierta politómica. En el caso de Formación Ciudadana, el análisis se realizó bajo el Modelo Logístico de Dos Parámetros (2PL) para las preguntas cerradas y para los ítems de respuesta abierta dicotómica. En lo que respecta a la prueba de Inglés, esta fue analizada por los proveedores del instrumento (*The British Council*), por lo que la Agencia no llevó a cabo ese proceso.

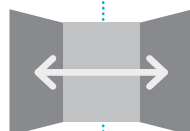
d. Comunicación de resultados educativos

En lo que se refiere a la comunicación de los resultados educativos, en la Tabla 2.1.5 se muestran los productos comunicacionales de la aplicación de los Estudios Nacionales 2017.

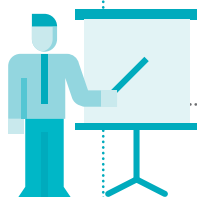
Tabla 2.1.5

Documentos comunicacionales de la aplicación de Estudios Nacionales 2017.

Tipo de documento	Título de la publicación	Soporte	Fecha de la publicación
Programas de Evaluaciones Educativas	Programa Estudio Nacional Formación Ciudadana 8.º básico 2017	Digital	Agosto 2017
	Programa Estudio Nacional Inglés III Medio 2017	Digital	Agosto 2017
	Programa Estudio Nacional Comprensión de Lectura 2.º básico 2017	Digital	Agosto 2017
Informes de Resultados Educativos	Informe de Resultados Nacionales Comprensión de Lectura 2.º básico 2017	Digital	Agosto 2018
	Informe de Resultados Nacionales Inglés III Medio 2017	Digital	Enero 2019
	Presentación de Resultados Nacionales Estudio Nacional de Formación Ciudadana 8.º básico 2017	Digital	Segundo semestre 2018



Capítulo I:
Visión panorámica de los Estudios
Nacionales Muestrales 2017



Capítulo II:
Descripción de los Estudios
Nacionales Muestrales 2017

Capítulo II: Descripción de los Estudios Nacionales Muestrales 2017

A continuación, se describe cada uno de los Estudios Nacionales Muestrales aplicados en 2017.

1.1 Lectura 2.º básico

Descripción general

El Estudio Nacional de Lectura de 2.º básico corresponde a la continuación de la prueba censal de lectura que se aplicó en nuestro país entre 2012 y 2015. La transformación de la prueba de censal a muestral responde a la recomendación recibida desde el Equipo de Tarea para la revisión del Simce 2014 y el Ministerio de Educación.

Tomando en cuenta dicha recomendación, el Estudio Nacional de Lectura de 2.º básico es una evaluación que se aplica a una muestra representativa de estudiantes del país, cada dos años.

Aprendizajes evaluados

La prueba aplicada para el Estudio Nacional de Lectura 2.º básico 2017 evaluó los Objetivos de Aprendizaje del Eje de Lectura, estipulados para 1.º y 2.º básico en las Bases Curriculares 2012 (Decreto N.º 439, Mineduc 2011). Además, se distinguen los mismos tres ejes de habilidades utilizados en las evaluaciones Simce: Localizar, Interpretar y relacionar, y Reflexionar (ver descripción en Evaluaciones Simce, Capítulo I), los que se evalúan a partir de textos literarios y no literarios.

Tabla 2.2.1

Matriz⁴³ de evaluación del Estudio Nacional de Lectura, 2.º básico, ejes de habilidades.

Habilidad	Cantidad de ítems	%
Localizar	16	38 %
Interpretar y relacionar	23	55 %
Reflexionar	3	7 %
Total	42	100 %

⁴³ Corresponde a la matriz final, es decir, la resultante luego del análisis psicométrico.

Formato de preguntas

Esta evaluación, al igual que las pruebas Simce, contiene preguntas con formato de selección múltiple y de respuesta abierta; sin embargo, incluye también otros dos tipos de formatos de preguntas: de selección de respuesta, donde los estudiantes deben encerrar en un círculo la o las respuestas correctas; y de ordenación, donde los estudiantes deben enumerar acciones o acontecimientos.

Las preguntas de selección múltiple de las pruebas de 2.º básico se construyen únicamente con tres opciones de respuesta (una clave y dos distractores), con el objeto de disminuir la carga verbal para estos estudiantes que se encuentran en etapas iniciales de lectura. Estas se responden en el mismo cuadernillo de preguntas, cuyo formato y extensión son adecuados para la edad de los estudiantes que participan en esta aplicación.

Cantidad de formas

El Estudio Nacional de Lectura 2.º básico cuenta con 3 formas equivalentes, las que cuentan con un núcleo de preguntas comunes que se usan para fines de equiparación.

Particularidades

La prueba de 2.º básico se aplica a estudiantes de ocho años, por lo que posee características particulares que la diferencian de otras pruebas de lectura aplicadas a estudiantes de mayor edad; por ejemplo, de las pruebas Simce de Lenguaje y Comunicación: Lectura. Estas diferencias son:

- Las hojas del cuadernillo, al igual que la letra utilizada, tienen un tamaño mayor al usado comúnmente en otras evaluaciones.
- El cuadernillo incluye un total de 25 preguntas basadas en los textos presentados.
- No tiene límite de tiempo para que respondan los estudiantes.
- Todas las preguntas se responden directamente en el cuadernillo.

Muestra

En lo que respecta a la muestra a la que se le aplicó el estudio, se generó el respectivo muestreo estratificado por región y dependencia administrativa del establecimiento (municipal/particular subvencionado/particular pagado), lo que permite reportar resultados a un nivel de agregación mínimo de establecimientos. Así, se constituyó una muestra nacional de al menos 7.500 estudiantes, para obtener una muestra de al menos 500 estudiantes en cada región.

La asignación del tamaño de la muestra en 2017 se realizó en dos fases: en la primera se usó una afijación simple por región, y en la segunda, una asignación proporcional según dependencia administrativa dentro de cada región. El factor de expansión para el análisis de datos que se utilizó fue el inverso de la probabilidad de selección de un estudiante en cada estrato.

El Estudio Nacional de Lectura 2.º básico se aplicó a una muestra de 260 establecimientos en los que se evaluó a un total de 7 349 estudiantes.

Análisis de resultados

En el caso del Estudio Nacional de Lectura 2.º básico, el análisis se realizó siguiendo los mismos criterios de una prueba censal indicados en la sección Evaluaciones Simce, Capítulo III de este informe. Esto, considerando que los resultados de este estudio debían estar en la misma métrica de la prueba Simce de 2.º básico aplicada hasta 2015.

Para reportar resultados por Estándares de Aprendizaje se utilizaron los puntajes de corte establecidos por el Ministerio de Educación. La Tabla 2.2.2 detalla los correspondientes a la prueba muestral de Lectura 2.º básico.

Tabla 2.2.2

Puntajes de corte para la clasificación 2017 según Estándares de Aprendizaje. Nivel 2.º básico.

Nivel escolar	Prueba	Insuficiente	Elemental	Adecuado
2.º básico	Estudio Nacional de Lectura	$X < 215$ puntos	$215 \leq X < 265$ puntos	$X \geq 265$ puntos

Nota: X = puntaje obtenido por el estudiante.

Resultados educativos

En el caso del Estudio Nacional de Lectura 2.º básico, los análisis agregados reportaron los puntajes promedio y distribución en los Estándares de Aprendizaje para diversas agregaciones, similar a lo reportado en las pruebas censales 2017, puesto que se utilizaron expansores para la muestra con el fin de reportar a toda la población. Asimismo, se consideró el error muestral en el cálculo de la significancia, tanto para las diferencias como para las variaciones.

1.2 Formación Ciudadana 8.º básico

Descripción general

El Plan Nacional de Evaluaciones 2016-2020 incorpora el Estudio de Formación Ciudadana el año 2017. Este complementa la información entregada por las pruebas censales entregando resultados nacionales respecto de esta área de formación.

Este estudio muestral tiene el objeto de medir los conocimientos, habilidades y actitudes de los estudiantes de 8.º básico en temáticas de Educación Cívica y Formación Ciudadana. Contar con este tipo de información resulta clave para la toma de decisiones de política pública orientadas a promover el compromiso democrático, el conocimiento cívico y la participación de los estudiantes, al tiempo que posibilita una adecuada retroalimentación de los procesos y prácticas docentes.

En específico, el estudio evalúa los conocimientos y actitudes con respecto a: conocimiento cívico, disposición a participar y actitudes, y valores democráticos. Esto, mediante la aplicación de una prueba de conocimientos a los estudiantes y cuestionarios a distintos actores de la comunidad educativa (directores, padres, profesores y estudiantes).

Aprendizajes evaluados

La evaluación se organiza en torno a cuatro grandes dominios de contenidos o temas, los que a su vez se dividen en subtemas que pretenden especificar los tópicos que se quiere relevar.

Los temas describen áreas de los ámbitos cívico, político y ciudadano, sobre los cuales los estudiantes deberían adquirir ciertos conocimientos y habilidades, y hacia los cuales deberían desarrollar también ciertas actitudes y comportamientos. Cada uno de los temas y subtemas incorpora tanto aspectos cognitivos –lo que los estudiantes saben sobre cierto contenido–, como aspectos afectivos/comportamentales –percepciones o actitudes que poseen sobre algún tema– (Schulz, Ainley, Fraillon, Losito y Agrusti, 2016).

1.2.1 Dominios de contenidos: temas y subtemas

Tomando como referencia el modelo de ICCS (Schulz, et al, 2016) y de Cox (2010), se definieron cinco temas relevantes para la evaluación. Estos buscan abarcar los ejes de contenido y los objetivos de aprendizaje actitudinales del currículo nacional vigente y, además, relevan los tópicos que a nivel internacional son considerados como centrales en los procesos de enseñanza y aprendizaje de la ciudadanía (Torney-Purta, 2002; Cohen, 2006; Ainley, Schulz, & Friedman, 2009; Cox, 2010).

Para esto, se consideraron todas las asignaturas que se imparten en los niveles de 1.º a 8.º básico, recogiendo los Objetivos de Aprendizaje que tienen un carácter transversal, como los vinculados al desarrollo de actitudes y valores centrales para la vida en sociedad.

Los temas presentes en la evaluación muestral de Formación Ciudadana 2017 los que se presentan a continuación:

a. Principios fundamentales de una sociedad democrática

Se enfoca en las bases éticas compartidas por todos los participantes de las sociedades cívicas actuales, apuntando a su defensa, promoción y protección (Schulz et al., 2016).

Como subtemas se han definido: la libertad, la equidad, la justicia y el bien común, y se han asociado algunos conceptos clave, tales como, el pluralismo, la solidaridad, la empatía y los derechos humanos, entre otros.

b. Participación política y compromiso cívico

Refiere a las diferentes formas, mecanismos y medios, a través de los cuales los individuos se involucran y participan con el objetivo de influir en la vida cívica y la toma de decisiones.

Los subtemas definidos son: derechos políticos, sociales y culturales; deberes ciudadanos, respeto por la ley y la institucionalidad; formas de participación e involucramiento cívico y político; y compromiso activo con la comunidad. por consiguiente, los conceptos centrales de esta temática son, entre otros: participación, toma de decisiones, cooperación, resolución de conflictos, responsabilidad, trabajo en equipo.

c. Institucionalidad dentro de una sociedad democrática

Considera los mecanismos, sistemas y organizaciones, formales e informales, públicas y privadas, que son el sustento de la vida en democracia, debido principalmente a que se encuentran en la base del contrato cívico entre los ciudadanos y sus sociedades (Hahn, 2010; Schulz et al., 2016).

Dentro de este tema se releva la organización política del Estado y sus instituciones, algo central para la gobernanza cívica y la legislación en pos del interés común, así como las bases legales y constitucionales que le dan forma.

d. Identidad, multiculturalismo y diversidad

Enfatiza en las bases culturales, simbólicas y valóricas de la identidad personal, así como del sentido de pertenencia a nivel comunitario, nacional e internacional (cosmopolitismo), relevando la importancia de la diversidad étnica, racial y cultural orientada a una convivencia democrática tolerante, respetuosa y pacífica.

Los subtemas son: identidad personal y autoimagen cívica; identidad nacional y grupal; y multiculturalismo, estereotipos y prejuicios.

e. Desafíos de la sociedad democrática actual

En esta temática se abordan algunos de los problemas y desafíos políticos, económicos y sociales más importantes a los que se enfrentan las democracias actuales, en el marco del proceso de globalización y liberalización de la economía mundial.

Los subtemas que se incorporan, considerando su relevancia para el contexto nacional, son: desigualdad social y económica, equidad de género, riesgos para la democracia y medio ambiente.

La Tabla 2.2.3 muestra los temas y subtemas de los dominios de contenido del Estudio Nacional de Formación Ciudadana 2017.

Tabla 2.2.3

Dominio de contenidos. Temas de la evaluación muestral Formación Ciudadana 2017.

Temas	Subtemas	Conceptos
a. Principios fundamentales de una sociedad democrática	Libertad	Democracia, cohesión social, justicia social, empatía, solidaridad, derechos humanos.
	Equidad	
	Justicia	
	Bien común	
b. Participación política y compromiso cívico	Derechos políticos, sociales y culturales	Participación, voto, toma de decisiones, medios de comunicación y redes sociales, cooperación, igualdad de derechos, responsabilidad, compromiso democrático, resolución de conflictos, trabajo en equipo.
	Deberes ciudadanos, respeto por la ley, las normas y la institucionalidad	
	Formas de participación e involucramiento cívico y político	
	Compromiso activo con la comunidad	
c. Institucionalidad dentro de una sociedad democrática	Instituciones, organización política y deberes/responsabilidades del Estado (poderes, cargos, autoridades, FF.AA, etc.)	Constitución, leyes/ legislación, gobierno y gobernanza, soberanía, representatividad, burocracia, institucionalidad y estructuras económicas.
	Instituciones civiles (partidos políticos, sindicatos, organizaciones culturales y religiosas, ONG)	
	Empresas, instituciones privadas y aspectos económicos	
	Legislación y normativa legal vigente	
d. Identidad, multiculturalismo y diversidad	Identidad personal y autoimagen cívica	Nacionalismo, patriotismo, cultura, identidad y cosmopolitismo, diversidad, tolerancia, respeto, multiculturalismo, migración, diferencia, ciudadanía global.
	Identidad nacional y grupal	
	Multiculturalismo, estereotipos y prejuicios	
e. Desafíos de la sociedad democrática actual	Desigualdad social y económica	Distribución del ingreso, flexibilidad y derechos laborales, sexismo y machismo, violencia de género, autoritarismo, desafección, desconfianza, calentamiento global, reciclaje, contaminación ambiental, guerra y paz, refugiados.
	Equidad de género	
	Riesgos para la democracia y conflictos internacionales	
	Medio ambiente	

1.2.2 Dominios y procesos cognitivos

A través de esta evaluación se espera obtener información respecto de las habilidades de los estudiantes para procesar los contenidos de los cinco temas antes presentados. Estas transitan desde una aplicación simple y directa del conocimiento sobre una situación concreta, hasta la selección, asimilación y utilización del conocimiento y comprensión de múltiples conceptos, con el objetivo de alcanzar conclusiones sobre situaciones complejas, abstractas y poco familiares (Schulz et al., 2016).

a. Conocimiento

El dominio del conocimiento se refiere a la formación cívica y ciudadana aprendida, y que los estudiantes utilizan cuando se comprometen en tareas cognitivas más complejas que les ayudan a dar sentido a su mundo cívico. Los procesos cognitivos asociados a este dominio son:

1. **Definir:** identificar afirmaciones que definen conceptos y contenidos cívicos y de ciudadanía.
2. **Describir:** identificar afirmaciones que describen las características clave de conceptos y contenidos cívicos y de ciudadanía.
3. **Ilustrar con ejemplos:** identificar ejemplos que respalden o clarifiquen afirmaciones sobre conceptos y contenidos cívicos y de ciudadanía (Schulz et al., 2016).

b. Razonamiento, análisis y aplicación

Se refiere al modo en que los estudiantes utilizan la información cívica y ciudadana para llegar a conclusiones más amplias que el contenido de un único concepto. El razonamiento comprende desde la aplicación directa del conocimiento y la comprensión para sacar conclusiones sobre situaciones conocidas concretas, hasta la selección y asimilación del conocimiento y comprensión de múltiples conceptos. Los procesos cognitivos asociados a este dominio son:

1. **Interpretar información:** identificar declaraciones sobre contenidos presentados en textos, gráficos o tabulados que dan sentido a la información a la luz de un concepto cívico y de ciudadanía.
2. **Relacionar:** utilizar aspectos definitorios clave de un concepto cívico y de ciudadanía para explicar o reconocer cómo un ejemplo ilustra un concepto.
3. **Justificar:** usar evidencia y conceptos cívicos y de ciudadanía para construir o reconocer un argumento razonado para apoyar un punto de vista.
4. **Integrar:** identificar las conexiones entre los diferentes conceptos a través de los temas y entre los dominios de contenido cívico y de ciudadanía.
5. **Generalizar:** identificar los principios conceptuales cívicos y de ciudadanía que se manifiestan como ejemplos específicos y explicar cómo estos pueden aplicarse en otros contextos cívicos y de ciudadanía.
6. **Evaluar:** identificar juicios sobre las ventajas y desventajas de puntos de vista alternativos o aproximaciones a conceptos y acciones cívicas y de ciudadanía.
7. **Predecir:** identificar los resultados probables de determinadas estrategias o acciones de políticas cívicas y de ciudadanía.
8. **Proponer soluciones:** identificar los cursos de acción o pensamiento que se pueden utilizar para solucionar problemas cívicos y de ciudadanía expresados como conflicto, tensión, o ideas no resueltas o disputadas.

1.2.3 Dominios afectivo-comportamentales

Además de los conocimientos y habilidades asociadas a procesos cognitivos reseñados anteriormente, los estudiantes, profesores, padres y apoderados han desarrollado creencias, percepciones, disposiciones, intenciones de comportamientos y comportamientos, todo lo cual es posible de conceptualizar bajo dos dominios afectivo-comportamentales: actitudes y compromiso (Schulz et al., 2016). La información relacionada con ambos dominios se obtiene mediante la aplicación de cuestionarios que permiten conocer: en qué medida los diferentes actores de la comunidad educativa están de acuerdo o en desacuerdo con ciertas afirmaciones; cuán a menudo realizan ciertas actividades; y cuáles son sus percepciones sobre diferentes temáticas de la vida cotidiana.

a. Actitudes

El dominio de las actitudes refiere a los juicios o evaluaciones relacionadas con ideas, personas, objetos, eventos, situaciones o relaciones. Este dominio, a diferencia de lo realizado por ICCS en 2009, considera tanto las actitudes y valores más permanentes y profundamente enraizados en los individuos, como también aquellas actitudes asociadas a temas más específicos y contingentes, que son temporalmente situadas y susceptibles a ser transformadas. Estas últimas, son las más relevantes en el contexto escolar si se considera al estudiante en proceso de formación valórica y actitudinal.

Es importante destacar que, si bien son los estudiantes el foco principal de esta evaluación, la batería de instrumentos intentará recoger también las actitudes (creencias, percepciones, juicios y disposiciones) tanto de padres como de profesores. La idea, como se planteó anteriormente, es despejar las eventuales relaciones que pueden existir entre las actitudes de padres y profesores, y las actitudes de los estudiantes en torno a temas sensibles y fundamentales para la vida en democracia, tales como la confianza en las instituciones, la diversidad (homosexualidad, inmigración, etc.), la desigualdad y la igualdad de derechos, entre otros.

Como se adelantó, las actitudes pueden ser también clasificadas según el dominio de contenido (tema) al que refieran. La categorización de estas actitudes y la relación con cada uno de los temas se presenta en la Tabla 2.2.4.

Tabla 2.2.4*Clasificación de actitudes. Temas de la evaluación muestral Formación Ciudadana 2017.*

Temas	Actitudes
a. Principios fundamentales de una sociedad democrática	• Percepciones sobre la democracia. • Actitudes hacia la igualdad de derechos y la libertad. • Actitudes relacionadas con la empatía y el bien común.
b. Participación política y compromiso cívico	• Percepciones y creencias sobre los roles ciudadanos (buena ciudadanía).
c. Institucionalidad dentro de una sociedad democrática	• Actitudes hacia el respeto y la desobediencia de las leyes. • Confianza en las instituciones. • Actitudes hacia las prácticas de transparencia, probidad y corrupción.
d. Identidad, multiculturalismo y diversidad	• Percepciones y creencias sobre la inmigración. • Percepciones y creencias sobre la discriminación. • Actitudes hacia la homosexualidad. • Creencias sobre el país de residencia, la comunidad y la escuela. • Actitudes hacia la diversidad.
e. Desafíos de la sociedad democrática actual	• Percepciones y creencias sobre la desigualdad. • Percepciones y creencias sobre la violencia (física y psicológica, machismo, bullying, etc.). • Creencias y actitudes sobre el medio ambiente y su conservación

Fuente: Elaboración propia a partir de Schulz et al., 2016.

b. Compromiso

El segundo dominio afectivo-comportamental corresponde al compromiso, y refiere concretamente a la participación y a las disposiciones hacia un activo involucramiento en la sociedad (interés, autoeficacia, participación, etc.), tanto de estudiantes como de padres y profesores. Junto a esto, este dominio también considera las expectativas de participación futura de los estudiantes.

Esto último requiere ítems que pregunten a los estudiantes sobre sus intenciones hacia la acción cívica en un futuro cercano o cuando sean adultos, así como también ítems que midan el grado de interés de los estudiantes y si se sienten competentes para comprometerse en algún tipo de acción cívica y/o política (Schulz et al., 2016).

Considerando la distinción entre formas de participación e involucramiento latente y manifiesto, se asume que las disposiciones básicas hacia el compromiso cívico y ciudadano (el interés en temas y asuntos políticos, la autoeficacia, etc.) así como las intenciones de comportamientos, entendidas como la preparación subyacente para realizar alguna acción, son de crucial importancia al estudiar el compromiso cívico y político en los adolescentes. La especificación de cada una de las disposiciones,

intenciones de comportamiento y formas de participación se exponen en la Tabla 2.2.5.

Tabla 2.2.5

Tipología del dominio afectivo-comportamental Compromiso.

Temas	Actitudes
Disposiciones	• Interés en asuntos políticos y sociales. • Sentido de autoeficacia ciudadana.
Participación cívica	• Participación cívica a través de las redes sociales. • Participación en organizaciones o grupos de carácter cívico-ciudadano. • Participación de los estudiantes en actividades del establecimiento.
Intenciones de comportamiento de los estudiantes	• Expectativas de participación cívica, legal o ilegal, en apoyo a alguna causa o como forma de protesta. • Expectativas de participación política en la vida adulta. • Expectativas de participación futura en actividades del establecimiento.

Fuente: Elaboración a partir de Schulz et al., 2016.

Formato de preguntas

Esta evaluación, al igual que las pruebas Simce, contiene preguntas con formato de selección múltiple y de respuesta abierta.

Cantidad de formas

El Estudio Nacional de Formación Ciudadana 2017 incluye la aplicación de dos formas equivalentes de la prueba de conocimientos que se aplica a los estudiantes, una forma de cuestionario aplicada a los padres y apoderados de los estudiantes, una forma de cuestionario aplicada a los directores de los establecimientos, una aplicada a los docentes de la asignatura y una forma de cuestionario actitudinal aplicado a los estudiantes.

Particularidades

El estudio de Formación Ciudadana permite complementar la información que se obtiene de la participación en el Estudio Internacional ICCS aplicado el 2017.

Muestra

Para este estudio se definió una muestra representativa a nivel nacional, regional y por género. Se evaluaron 242 establecimientos y a 8 701 niñas y niños. Para ello, se generó el respectivo muestreo estratificado por región y dependencia administrativa del establecimiento (municipal/particular subvencionado/particular pagado), lo que permite reportar resultados a un nivel de agregación mínimo de establecimientos. Así, se constituyó una muestra nacional de al menos 7 500 estudiantes, para obtener una muestra de al menos 500 estudiantes en cada región.

La asignación del tamaño de la muestra en 2017 se realizó en dos fases: en la primera se usó una afijación simple por región y en la segunda, una asignación proporcional según dependencia administrativa dentro de cada región. El factor de expansión para el análisis de datos que se utilizó fue el inverso de la probabilidad de selección de un estudiante en cada estrato.

Análisis de resultados

La batería de instrumentos para evaluar Formación Ciudadana consideró una serie de cuestionarios y una prueba de conocimiento. Esta última fue analizada utilizando un modelo de dos parámetros (2PL), pues presentó el mejor ajuste a los datos. Se utilizaron los criterios de ajuste al modelo ya indicados. Asimismo, se emplearon expansores con el fin de que la muestra represente a toda la población de 8.º básico. Se decidió aplicar una media de 500 y una desviación de 100, misma escala del Estudio ICCS 2016 y equiparar los puntajes de la muestra a los puntajes obtenidos en dicho estudio. Para calcular el puntaje de cada estudiante, se aplicó una equiparación entre las puntuaciones 2PL de la muestra y las puntuaciones de ICCS 2016, para los casos de puntajes que no se pudo equiparar, se les aplicó una transformación lineal.

Resultados educativos

En el caso de Formación Ciudadana 8.º básico, los análisis agregados reportaron los puntajes promedio para diversas agregaciones. Se utilizaron expansores para la muestra con el fin de reportar a toda la población. Asimismo, se consideró el error muestral en el cálculo de la significancia para las diferencias.

Por otro lado, con la batería de cuestionarios específicos de esta prueba se generaron diversos indicadores que dieron cuenta de la percepción en temas de Formación Ciudadana de los jóvenes de 8.º básico, padres o apoderados y docentes.

1.3 Idioma Extranjero Inglés III medio

Descripción general

Con el fin de concretar el Estudio Nacional de Inglés 2017, se realizó una licitación pública⁴⁴ para la adquisición de la prueba estandarizada de inglés como segunda lengua, en las áreas de comprensión lectora y comprensión auditiva, y el procesamiento de sus resultados. En lo que respecta a los procesos operativos asociados a la aplicación (impresión, distribución y aplicación), estos fueron realizados mediante los mismos procedimientos que las pruebas Simce y otros estudios nacionales, y se rigen bajo sus mismos estándares de seguridad y control, con el objetivo de garantizar la confidencialidad de la información del instrumento.

⁴⁴ Licitación Pública "Adquisición prueba muestral Simce Inglés 2017 y procesamiento de sus resultados" ID N.º721703-47- LR16. Esta licitación fue adjudicada por el British Council.

El Estudio Nacional de Idioma Extranjero Inglés evalúa las habilidades de comprensión lectora y comprensión auditiva en inglés en una muestra representativa de los estudiantes que cursan III medio. Para ello, se utilizó la prueba Aptis for Teens desarrollada por el British Council.

Aptis for Teens es una herramienta de evaluación del inglés como segunda lengua o como lengua extranjera, dirigida específicamente a personas de 13 a 17 años, ya que evalúa las habilidades usando temas y escenarios familiares para dicha población. Las preguntas son diseñadas considerando actividades de la vida diaria de los adolescentes, por ejemplo, redes sociales, tareas, y eventos escolares y deportivos.

La evaluación original incluye cinco dominios: Núcleo (gramática y vocabulario), Comprensión Lectora, Comprensión Auditiva, Producción Escrita y Producción Oral, los que pueden aplicarse seleccionando cualquier combinación de ellos. El Estudio Nacional aplicado el 2017 solo aplicó los componentes de Comprensión Lectora y Comprensión Auditiva.

La evaluación contiene preguntas que se asocian a los niveles A1, A2 y B1 del Marco Común Europeo de Referencia para las Lenguas (Common European Framework of Reference for Languages, CEFR). Para efecto del estudio aplicado en Chile, se considera como nivel de "certificación" el nivel A2 de dicho marco.

Aprendizajes evaluados

El Estudio Nacional de Inglés III medio 2017 evaluó las habilidades de comprensión lectora y de comprensión auditiva correspondientes al CEFR presentes en los Objetivos Fundamentales y Contenidos Mínimos Obligatorios del Ajuste Curricular 2009 (Decreto N.º 254, Mineduc 2009c) de la asignatura Idioma Extranjero Inglés. La evaluación consta de dos pruebas: Comprensión Lectora y Comprensión Auditiva que evalúan conocimientos y habilidades posibles de ser medidos en una prueba de lápiz y papel.

a. Sección de Comprensión Lectora

En la prueba de Comprensión Lectora se evalúa el nivel de habilidad para comprender diferentes tipos de textos (textos informativos, diálogos, entre otros) de diferente extensión que están escritos en inglés. Para ello, se incluyen varias preguntas referidas a cada texto. Cada estudiante debe responder de forma escrita un total de 25 preguntas que corresponden a diferentes niveles de exigencia de acuerdo con los niveles del CEFR incorporados en la evaluación, esto en un tiempo de 30 minutos.

b. Sección de Comprensión Auditiva

En la prueba de Comprensión Auditiva se evalúa el nivel de habilidad para comprender anuncios y conversaciones orales breves entre hablantes nativos angloparlantes. Estas producciones orales corresponden a situaciones reales y variadas en las que se utiliza un lenguaje auténtico y cotidiano. Cada grabación se escucha dos veces y, después de ello, los estudiantes deben responder una o dos preguntas referidas a ellas, las que se presentan por escrito, en inglés, en un cuadernillo de preguntas. La prueba tiene 25 preguntas que corresponden a diferentes niveles de exigencia de acuerdo con los niveles del CEFR incorporados en la evaluación, el tiempo asignado para responder la prueba es de 55 minutos.

c. Descripción de los niveles CEFR reportados en las pruebas

En la Tabla 2.2.6 se describen las habilidades de comprensión lectora y comprensión auditiva que corresponden a los niveles de logro reportados en el Estudio Nacional de Inglés.

Tabla 2.2.6

Habilidades de comprensión lectora y comprensión auditiva. Estudio Nacional de Inglés 2017⁴⁵.

Niveles	Habilidades comprensión lectora	Habilidades comprensión auditiva
Básico e intermedio (A2 y B1)	- Comprende textos siempre que estén relacionados con temas cotidianos que poseen estructuras gramaticales simples y de mediana complejidad. - Reconoce las ideas principales de un texto siempre que trate temas cotidianos.	- Comprende las ideas principales de un discurso siempre que esté articulado con claridad, trate temas cotidianos, la pronunciación tenga un acento normal y esté acompañado de un apoyo extralingüístico. - Identifica palabras, frases y expresiones de uso frecuente. - Identifica el tema principal en diálogos y monólogos breves y simples, siempre que su pronunciación sea lenta y clara y se acompañen de un apoyo extralingüístico.
Principiantes (A1)	- Comprende textos siempre que sean muy breves y sencillos, en tanto lectura frase por frase o relectura, captura de nombres, palabras y frases básicas relacionadas con su persona, familia y entorno inmediato. - Comprende textos e identifica el tema principal e información explícita destacada, siempre que se acompañen de abundante apoyo visual.	- Reconoce palabras y expresiones básicas de uso habitual, siempre que sean relativas a su persona, familia y entorno inmediato. - Comprende diálogos y monólogos siempre que sean pronunciados de forma muy lenta y con pausas, y que su articulación sea cuidadosa.

Fuente: Marco Común Europeo de Referencia para las Lenguas (Instituto Cervantes, 2002). Adaptación.

d. Estudiantes en situación de discapacidad

Cabe mencionar que los estudiantes de III medio en situación de discapacidad sensorial, de los establecimientos que participaron de este estudio nacional, pudieron rendir las secciones de Comprensión Lectora y Comprensión Auditiva, con acomodaciones para discapacidad visual parcial y total. Los estudiantes en situación de discapacidad auditiva rindieron la sección de Comprensión Lectora regular, no así la sección de Comprensión Auditiva, previo acuerdo entre el apoderado y el director.

Formato de preguntas

Esta evaluación contiene preguntas con formato de selección múltiple (en ambas pruebas) y de términos pareados (en la prueba de Comprensión Lectora).

⁴⁵ Puede conocer más información de MCER en https://cvc.cervantes.es/ensenanza/biblioteca_ele/marco/cvc_mer.pdf.

Cantidad de formas

Tanto la prueba de Comprensión Lectora como la de Comprensión Auditivas cuentan con dos formas equivalentes; cuyas preguntas son todas distintas entre formas.

Particularidades

Esta prueba tiene la particularidad de ser diseñada y analizada por el British Council. La prueba está referida a los niveles del CEFR y los puntajes de corte se establecieron específicamente, a través de un ejercicio nacional de *standard setting*, para ser aplicados en Chile.

Además, para la aplicación de la sección de Comprensión Auditiva se requiere usar una radio y un CD proporcionado por el British Council.

Muestra

El estudio se aplicó a una muestra representativa a nivel nacional de 7 340 estudiantes correspondientes a 137 establecimientos educativos. Para ello, se generó el respectivo muestreo estratificado por región y dependencia administrativa del establecimiento (municipal/particular subvencionado/particular pagado), lo que permite reportar resultados a un nivel de agregación mínimo de establecimientos. Así, se constituyó una muestra nacional de al menos 7 500 estudiantes, para obtener una muestra de al menos 500 estudiantes en cada región.

La asignación del tamaño de la muestra en 2017 se realizó en dos fases: en la primera se usó una afijación simple por región y en la segunda, una asignación proporcional según dependencia administrativa dentro de cada región. El factor de expansión para el análisis de datos que se utilizó fue el inverso de la probabilidad de selección de un estudiante en cada estrato.

Análisis de resultados

En el caso del Estudio Nacional de Inglés III medio, el análisis de los resultados es responsabilidad del proveedor del instrumento, en 2017 The British Council, quien realiza el análisis psicométrico del instrumento y entregan a la Agencia un archivo con el listado de estudiantes con su puntuación y su clasificación.

Para determinar los puntajes de corte que delimitan los niveles, la Agencia realizó un ejercicio de *standard setting* (establecimiento de puntajes de corte) en el contexto nacional, de manera que la operacionalización de las categorías considerara las características específicas de la población evaluada.

En este contexto, se implementaron dos jornadas de establecimiento de puntajes de corte, una para las formas aplicadas a Comprensión Lectora y otra, para las utilizadas para evaluar Comprensión Auditiva.

El método seleccionado para recomendar los puntajes de corte fue el método Angoff Modificado, el cual fue aplicado por un panel de docentes y especialistas en la enseñanza del inglés como lengua extranjera. Cada jornada tuvo una duración de tres días en los que se incluye una introducción al procedimiento, una capacitación en el Nivel del Marco Común Europeo A2 y en el método, y tres rondas de trabajo para recomendar los puntajes de corte.

Durante cada jornada se recomendó el puntaje de corte y un intervalo para cada una de las dos formas. Para ello, se definió en conjunto los conocimientos y habilidades que posee un estudiante que alcanza mínimamente el nivel A2 del Marco Común Europeo en Comprensión Lectora o en Comprensión Auditiva, según corresponde; se realizó una ronda de práctica del método Angoff Modificado; y se llevaron a cabo tres rondas de estimaciones individuales, entre las cuales se presentó a los participantes información del porcentaje de estudiantes chilenos que respondió correctamente cada pregunta y el porcentaje de estudiantes que alcanzaría el Nivel A2 de usarse los puntajes de corte recomendados.

Como resultado de las jornadas se obtuvo una recomendación de puntaje de corte para diferenciar entre nivel A1 y A2 para cada forma de la prueba y un intervalo de confianza asociado a cada puntaje de corte.

La jornada de Comprensión Lectora se llevó a cabo en el Hotel San Francisco los días 13, 14 y 15 de junio de 2018. El panel estuvo conformado por un total de once especialistas: tres profesores de aula, tres representantes del Programa Inglés Abre Puertas (PIAP) del Mineduc, tres docentes universitarios y dos representantes del Centro de idiomas de la PUC.

Por su parte, la jornada de Comprensión Auditiva se llevó a cabo en el Hotel San Francisco los días 20, 21 y 22 de junio de 2018. El panel estuvo conformado por un total de ocho especialistas: tres profesores de aula, dos representantes del Programa Inglés Abre Puertas (PIAP) del Mineduc y tres docentes universitarios.

Los puntajes de corte recomendados para las cuatro formas analizadas (dos de Comprensión Lectora y dos de Comprensión Auditiva) fueron los mismos: 13 preguntas correctas de un total de 25 preguntas.

Resultados educativos

Para el Estudio Nacional de Inglés III medio, los análisis agregados reportaron promedios del puntaje total, Comprensión Lectora y Comprensión Auditiva, así como porcentaje de estudiantes en nivel básico e intermedio (A2 y B1). Todo ello para diversas agregaciones, similar a lo reportado en las pruebas censales 2017, puesto que se utilizaron expansores para la muestra con el fin de reportar a toda la población. Asimismo, se consideró el error muestral en el cálculo de la significancia tanto para las diferencias como para las variaciones.



**Referencias
bibliográficas**



Anexos

Referencias bibliográficas

Australian Council for Educational Research (ACER). (2013). *Evaluación de los procesos y los productos asociados a la elaboración de instrumentos, las operaciones de campo y la gestión de datos de las pruebas nacionales de Simce*. [Documento interno de trabajo]. Santiago: Chile.

AERA, APA y NCME. (2014). *Standards for Educational and Psychological Testing*. Washington: autores.

Ainley, J., Schulz, W., & Friedman, T. (2009). *ICCS 2009 Encyclopedia Approaches to civic and citizenship education*. IEA. Amsterdam.

Bock, R. y Mislevy, R. (1982). Adaptive EAP estimation of ability in a microcomputer environment, *Applied Psychological Measurement*, 6. Illinois, Chicago: University of Illinois.

Brancato, G., Macchia, S., Murgia, M., Signore, M., Simeoni, G., Blanke, K., & Hoffmeyer-Zlotnik, J. (2006). Handbook of recommended practices for questionnaire development and testing in the European statistical system. *European Statistical System*.

Camilli, Gregory. (1994). Teacher's Corner: Origin of the Scaling Constant $d = 1.7$ in Item Response Theory. *Journal of Educational and Behavioral Statistics*.

Case, B., Jorgensen M. y Sucker S. (2004). *Alignment in Educational Assessment: Assessment Report*. San Antonio, Texas: Pearson Education Inc. Camilli, Gregory. (1994). Teacher's Corner: Origin of the Scaling Constant $d = 1.7$ in Item Response Theory. *Journal of Educational and Behavioral Statistics*.

Cohen, J. (2006). Social, Emotional, Ethical, and Academic Education: Creating a climate for learning, Participation in Democracy and Well-Being. *Harvard Educational Review*, 76(2), 201-285.

Cox, C. (2010). *Oportunidades de aprendizaje escolar de la ciudadanía en América Latina: currículos comparados*. Bogotá: SREDECC.

Cox, C., Bascopé, M., Castillo, J., Miranda, D. y Bonhomme, M. (2014). *Educación ciudadana en América Latina: Prioridades de los currículos escolares*. Ginebra. UNESCO.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart and Winston, 6277 Sea Harbor Drive, Orlando, FL.

DeVellis, F. (2012). *Scale development: theory and applications*. Thousand Oaks, California: Sage Publications.

DeVellis, F. (2012). *Scale development: theory and applications*. Thousand Oaks, California: Sage Publications.

Downing, S. y Haladyna, T. (2006). *Handbook of Test Development*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Haladyna, T. M. y Rodríguez, M. C. (2013). *Developing and Validating Items*. Nueva York: Routledge.

- Hahn, C. L. (2010). Comparative civic education research: What we know and what we need to know. *Citizenship Teaching and Learning*, 6(1), 5–23.
- Instituto Cervantes (2012). *Marco Común Europeo de Referencia para las Lenguas: Enseñanza, Aprendizaje, Evaluación*. Madrid: MECD-Anaya.
- Kolen, M. y Brennan, R. (2004). *Test equating, Scaling and Linking: Methods and Practices*. Nueva York: Springer.
- Ley General de Educación (N.º 20 370). *Establece la Ley General de Educación*, Santiago de Chile.
- Lord, F. y Novick, M. (2008). *Statistical theories of mental test scores*. Reading, Massachusetts: Addison-Wesley.
- Martin, M., Rust, K., Adams, R. (eds.) (1999). *Technical Standards for IEA Studies*. Amsterdam, Países Bajos: International Association for the Evaluation of Educational Achievement (IEA).
- Martínez, M. R. (2005). *Psicometría: Teoría de los Test Psicológicos y Educativos*. Madrid: Editorial SINTESIS, S.A.
- Ministerio de Educación, (2009a). Decreto N.º 170: *Fija normas para determinar los alumnos con necesidades educativas especiales que serán beneficiarios de las subvenciones para educación especial*. Santiago de Chile.
- Ministerio de Educación. (2009b). Decreto N.º 254: *Objetivos Fundamentales y Contenidos Mínimos Obligatorios de la Educación Básica y Media. Actualización 2009*. Santiago de Chile.
- Ministerio de Educación. (2011). Decreto N.º 439/2012: *Bases Curriculares Educación Básica: Lenguaje y Comunicación, Matemática, Ciencias Naturales, Historia, Geografía y Ciencias Sociales, Idioma Extranjero Inglés*. Santiago de Chile.
- Ministerio de Educación. (2013). Decreto Supremo de Educación N.º 614/2013: *Bases Curriculares de 7º básico a 2º medio: Lengua y Literatura, Matemática, Ciencias Naturales, Historia, Geografía y Ciencias Sociales, Idioma Extranjero Inglés, Educación Física y Salud*. Santiago de Chile.
- Murphy, E. J. y Elliot, S. (1991). A Sustained Strategy for Preventing Potential Bias in the TMTE, *Teacher Certification Testing: Recent Perspectives*. Amherst, Massachusetts: NESInc.
- Ricker, K. (2006). Setting Cut-Scores: A Critical Review of the Angoff and Modified Angoff Methods. *Alberta Journal of Educational Research*, 52,1. Alberta, Canadá: University of Alberta.
- Schulz, W., Ainley, J., Fraillon, J., Losito, B. y Agrusti, G. (2016). IEA *International civic and citizen education study. Assessment framework*. IEA. Amsterdam
- Torney-Purta, J. (2002). The school's role in developing civic engagement: A study of adolescents in twenty-eight countries. *Applied developmental science*, 6(4), 203–212.

Anexo A: Pruebas experimentales Simce 2017

Las pruebas experimentales tienen un doble objetivo: en el caso de las preguntas cerradas, tienen como fin la obtención de datos psicométricos (desde la Teoría Clásica e IRT) que sirvan de insumo para seleccionar los ítems que integrarán las pruebas censales del año siguiente. En el caso de las preguntas abiertas, su corrección permite adicionalmente obtener información para validar las rúbricas y elaborar los manuales de corrección, proceso que se describe de manera general al final de este anexo.

Las pruebas experimentales son de carácter muestral y su aplicación se realiza en conjunto con las pruebas censales y demás instrumentos aplicados cada año y en los mismos niveles escolares.

Las pruebas experimentales que sirvieron de base para las pruebas 2017 (PE2017) se aplicaron en el año 2016 en todas las áreas y niveles evaluados en 2017.

En la aplicación de las PE2017 se consideró una muestra de estudiantes de los niveles evaluados, pertenecientes a establecimientos de las regiones de Valparaíso, Metropolitana y del Biobío⁴⁶. Para la determinación de esta muestra se contemplaron los mismos métodos descritos para la aplicación muestral (contenidos en el capítulo 2 de este informe), con la salvedad de que se excluyeron todos los establecimientos en los que se estuviera llevando a cabo un estudio paralelo en el mismo nivel escolar, para evitar una posible sobrecarga de evaluaciones.

Para la definición de la muestra de las PE2017 se establecieron estratos⁴⁷. En los niveles de Educación Básica estos estratos se conformaron según las siguientes variables: dependencia (municipal/particular subvencionado/particular pagado), ruralidad (urbano/rural) y rendimiento Simce. Esta última variable se dividió en tres categorías (cuyos puntos de corte están dados por el promedio $\pm 0,5$ desviaciones estándar, respecto del grado y área correspondiente), para asegurar la participación de establecimientos de estudiantes de distinto rendimiento. Las muestras fueron calculadas por prueba y nivel escolar.

Para las pruebas de II medio los estratos se conformaron con las siguientes variables: dependencia administrativa (municipal/particular subvencionado/particular pagado), modalidad educacional (científico-humanista/técnico-profesional/polivalente) y rendimiento Simce, dividido en tres categorías.

⁴⁶ No se seleccionó una muestra representativa de las distintas regiones, por cuanto la aplicación de la metodología IRT no lo requiere para analizar los datos.

⁴⁷ En caso de que en la formación de estratos se detectara una cantidad considerable de establecimientos nuevos (alrededor de 100 casos) que no contaran con resultados Simce de la medición anterior, se procedió a crear un estrato adicional que los agrupara. Además, se procuró que todos los estratos tuvieran una representación de al menos 1 %. En caso contrario, estos últimos fueron asociados al estrato más afín.

Respecto del tamaño de la muestra, cabe señalar que para la obtención de una estimación robusta de acuerdo al modelo utilizado (IRT de tres parámetros⁴⁸), se requiere contar con un mínimo de 1 500 respuestas efectivas para cada pregunta; sin embargo, para las PE2017 se estimó un mínimo de 2 500 estudiantes por pregunta, considerando la amplia gama de análisis requeridos y la tasa de omisión promedio en este tipo de evaluaciones (menor a 5%). A partir de este dato se determinó el número de establecimientos que compondrían la muestra, la cantidad de formas de prueba de cada evaluación y la tasa de estudiantes por establecimiento.

La cobertura efectiva que tuvo la aplicación de las PE2017 fue de 189 496 estudiantes, distribuidos por nivel escolar, como se muestra en la Tabla A1.

Tabla A1

Estudiantes según nivel escolar en las PE2017 y porcentaje de la muestra teórica que representan.

Nivel escolar	Nº de evaluados	% de la muestra teórica
4.º básico	33 377	88,4 %
8.º básico	41 971	86,7 %
II medio	36 569	80,5 %

La información sobre las preguntas que conformaron cada una de las PE2017 por asignatura y nivel escolar, se encuentra en las planillas maestras disponibles, previa petición formal, en www.agenciaeducacion.cl. El tipo de información que se encuentra en estas planillas corresponde a la identificación y a los datos métricos de los ítems, según se detalla en las tablas A2 y A3.

Respecto de las preguntas abiertas de pruebas experimentales, como se señaló al comienzo, el proceso de corrección permite validar las rúbricas y elaborar los manuales de corrección. Este proceso se inicia con la revisión de las preguntas a corregir, que consiste en un análisis cualitativo de las preguntas, a partir del cual se eliminan aquellas que presentan problemas, basándose fundamentalmente en el registro de consultas realizadas por los estudiantes durante la aplicación.

Una vez definidas las preguntas a corregir, se selecciona una muestra de respuestas con el objeto de ajustar las rúbricas preliminares⁴⁹ y los respectivos manuales de corrección. En la PE2017 la muestra constó de 500 respuestas de cada una de las preguntas abiertas seleccionadas. En el proceso de ajuste de rúbricas y manuales, participan profesionales expertos de las distintas asignaturas, tanto externos como pertenecientes a la Agencia.

⁴⁸ Existe variada información con respecto al mínimo necesario para estimaciones robustas. El valor más alto encontrado para un modelo logístico de tres parámetros es de entre 1 000 y 1 500 casos por ítem (Murphy y Elliot, 1991 y Ricker, 2006).

⁴⁹ Durante el proceso de elaboración de las preguntas abiertas se genera una rúbrica preliminar basada en información teórica sobre las respuestas esperadas. Estas rúbricas suelen incluir desde las respuestas más completas y correctas posibles, hasta las más erróneas e incompletas. Luego se establece una escala descriptiva (asociada a una puntuación) y, a partir de esta, se delimitan dos categorías (correcto e incorrecto) o tres categorías (correcto, incorrecto y parcialmente correcto).

El proceso se inicia con la corrección de 200 respuestas, utilizando la rúbrica preliminar. En función de este proceso, se ajustan las rúbricas y se corrigen las restantes 300 respuestas, con el objeto de validarla. El resultado de esta fase del proceso es un Manual de Corrección ajustado que incluye la rúbrica definitiva, que identifica las distintas categorías de respuesta en función de códigos y ejemplos reales de respuestas de los estudiantes comentados, que permitirán orientar de mejor forma los procesos de corrección de la prueba experimental. Adicionalmente, el manual incluye una serie de ejercicios prácticos para que los correctores puedan ejercitar con antelación.

En una segunda etapa, se selecciona y capacita a los correctores y supervisores que participarán en el proceso, con el objeto de asegurar que posean un dominio acabado del protocolo de corrección, de la rúbrica y de los demás procedimientos necesarios para resguardar la consistencia entre correctores.

En una tercera etapa, se realiza una marcha blanca, durante la cual se corrige 10 % del total de las respuestas obtenidas. En este periodo se incluyen, en caso de ser necesario, precisiones al Manual de Corrección.

Finalmente, se corrigen todas las respuestas obtenidas de la aplicación experimental, realizando doble corrección del 25 % de las respuestas, con el objeto de calcular la consistencia entre correctores. En función del grado de consistencia, se evalúa si el Manual posee la calidad adecuada para ser utilizado en una aplicación censal. En las PE2017, para la corrección de preguntas abiertas, se estableció un requerimiento de consistencia igual o superior a 80 %. Adicionalmente, los supervisores de cada grupo de corrección corrigieron 25 % de las respuestas doblemente corregidas por su grupo, para asegurar el correcto alineamiento de todos los integrantes del grupo respecto de los criterios contenidos en los Manuales de corrección.

Tabla A2

Código y descripción de los campos de identificación de los ítems de las PE2017.

	Código del campo	Descripción de campo
Identificación	ÁREA	Área a la que pertenece el ítem (todos los ítems).
	ID_COD	Código identificador del ítem (todos los ítems).
	ID_TIPO	Tipo de ítem (todos los ítems).
	POS_PD	Posición en la prueba definitiva de aplicación anterior.
	POS_EXP	Posición en la PE2017.
	EQ_PD_AÑOS_2017	Ítems de <i>equating</i> entre años de la prueba definitiva 2017.
	EQ_PD_FORMAS_2017	Ítems de <i>equating</i> entre formas de la prueba definitiva 2017.
	POS_PD_2017	Posición(es) en la prueba definitiva 2017.
	CLAVE	Opción de respuesta correcta.
	ALT_VAL	Cantidad de alternativas válidas.
	CE_MCNB	Nivel escolar al que corresponde el desempeño y/o tarea evaluada.
	CE_OA	Objetivo de aprendizaje (de acuerdo al currículo vigente).
	CE_OBJ_EV	Objetivo de evaluación de las especificaciones técnicas.
	CE_EJE_TEMÁTICO_PD_2016	Eje Temático de la prueba definitiva 2017 según currículo ^a .
	CE_DOMINIO_COGNITIVO_PD_2017	Dominio Cognitivo de la prueba definitiva 2017, según currículo ^b .
	CE_TÍTULO_ESTIMULO_COMÚN	Estímulo al que se asocia el ítem.

a En la prueba de Lectura corresponde a tipo de texto.

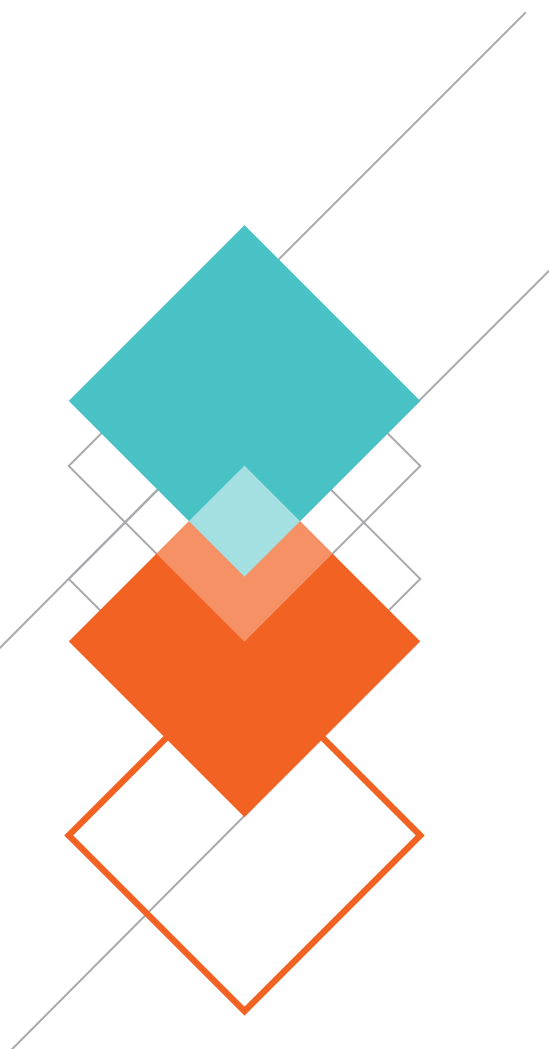
b En la prueba de Lectura corresponde a habilidades.

Tabla A3

Código y descripción de los campos de datos métricos de los ítems de las PE2017.

	Código del campo	Descripción de campo
DATOS MÉTRICOS PREGUNTAS OPCIÓN MÚLTIPLE	%_A_2017	Porcentaje de respuestas de la alternativa A.
	%_B_2017	Porcentaje de respuestas de la alternativa B.
	%_C_2017	Porcentaje de respuestas de la alternativa C.
	%_D_2017	Porcentaje de respuestas de la alternativa D.
	%_OMIT_2017	Porcentaje de respuestas omitidas.
	CORR_A_2017	Correlación opción-test considerando la alternativa A.
	CORR_B_2017	Correlación opción-test considerando la alternativa B.
	CORR_C_2017	Correlación opción-test considerando la alternativa C.
	CORR_D_2017	Correlación opción-test considerando la alternativa D.
	DISC_IRT_2017	Parámetro de discriminación IRT.
	DIFIC_IRT_2017	Parámetro de dificultad IRT.
	AZAR_IRT_2017	Parámetro de azar.
	RP67_2017	Parámetro de dificultad en escala Simce ^a .
	GENERO_2017	Indicador de funcionamiento diferencial del ítem (género favorecido).
	RURALIDAD_2017	Indicador de funcionamiento diferencial del ítem (zona favorecida).
	GSE_AltoMedio	Indicador de funcionamiento diferencial del ítem. Se indica el grupo favorecido por el ítem.
	GSE_AltoBajo	Indicador de funcionamiento diferencial del ítem. Se indica el grupo favorecido por el ítem.
	GSE_MedioBajo	Indicador de funcionamiento diferencial del ítem. Se indica el grupo favorecido por el ítem.
DATOS MÉTRICOS PREGUNTAS ABIERTAS	%_PARC	Porcentaje de respuesta para la categoría 'parcialmente correcta'.
	%_CORR	Porcentaje de respuesta para la categoría 'correcta'.
	%_OMITA	Porcentaje de respuestas omitidas.
	CORR_A	Correlación opción-test considerando categoría con puntuación máxima.
	DISC_A_IRT	Parámetro de discriminación IRT.
	DIFIC_PARC_IRT	Parámetro de dificultad IRT para la categoría 'parcialmente correcta'.
	DIFIC_CORR_IRT	Parámetro de dificultad IRT para la categoría 'correcta'.
	RP67_PARC	Parámetro de dificultad en escala Simce para la categoría 'parcialmente correcta' ^b .
	RP67_CORR	Parámetro de dificultad en escala Simce para la categoría 'correcta' ^c .
	GENERO_A	Indicador de funcionamiento diferencial del ítem (género favorecido).
	RURALIDAD_A	Indicador de funcionamiento diferencial del ítem (zona favorecida).
	GSE_AltoMedio	Indicador de funcionamiento diferencial del ítem. Se indica el grupo favorecido por el ítem.
	GSE_AltoBajo	Indicador de funcionamiento diferencial del ítem. Se indica el grupo favorecido por el ítem.
	GSE_MedioBajo	Indicador de funcionamiento diferencial del ítem. Se indica el grupo favorecido por el ítem.

a b c Ajustado a una probabilidad de respuesta de 67 %



@agenciaeduca
instagram.com/agenciaeducacion
facebook/Agenciaeducacion
contacto@agenciaeducacion.cl
www.agenciaeducacion.cl

