



Informe del comportamiento psicométrico de los ítems piloto de

Saber 7°

Aplicación 2022



Informe del comportamiento psicométrico de los ítems piloto de Saber 7°

Aplicación 2022



Presidente de la República
Gustavo Francisco Petro Urrego

Ministro de Educación
Alejandro Gaviria Uribe

Viceministro de Educación
Preescolar, Básica y Media
Hernando Bayona Rodríguez

Director General
Andrés Molano Flechas

Secretaria General
Luisa Fernanda Trujillo Bernal

Directora de Evaluación
Natalia González Gómez

Director de Producción y Operaciones
Oscar Orlando Ortega Mantilla

Director de Tecnología e información
Sergio Andrés Soler Rosas

Subdirectora de Diseño de Instrumentos
Lady Catheryne Lancheros Florian

Subdirectora (E) de Análisis y Divulgación
Natalia González Gómez

Subdirector de Estadísticas
Cristian Fabian Montaña Rincón

Subdirectora de Producción de Instrumentos
Nubia Rocío Sánchez Martínez

Subdirectora de Aplicación de Instrumentos
Yamile Ariza Luque

Subdirector de Información
William Alfredo Sandoval Sandoval

Subdirector de Abastecimiento y Servicios Generales
Hans Ronald Niño García

Subdirector Financiero y Contable
William Abel Otero Millán

Subdirector de Desarrollo de Aplicaciones
Armando Alfonso Leyton González

Subdirectora de Talento Humano
Lorena Catalina Ramírez Duque

Jefa Oficina Asesora de Comunicaciones y Mercadeo
Andrea Linares Gómez

Jefa Oficina Asesora de Gestión de Proyectos de Investigación
María Luisa Miranda Miranda

Jefe Oficina Asesora de Planeación
Luis Alberto Colorado Aldana

Jefa Oficina de Control Interno
Adriana Bello Cortés

Jefa Oficina Asesora Jurídica
Ana María Cristina de la Cuadra

Asesora Unidad Atención al Ciudadano
Alba Liliana Abril Daza

Elaboración del documento
Jenny Cárdenas Buitrago
Alejandro Corrales Espinosa
Andrea Catalina Gonzáles García
Sergio Armando López Castillo
Jonnathan David Rico Marin

Revisión
Juan Sebastián Argüello Chiquiza
Astrid Julieth Betancourt Pineda
Karen Rosana Córdoba Perozo
Lina Paola Delgadillo Rincón
Patricia Escudero Montañez
Luís Adrián Quintero Sarmiento

Procesamiento estadístico
Juan Sebastián Argüello Chiquiza
Astrid Julieth Betancourt Pineda
Lina Paola Delgadillo Rincón
Patricia Escudero Montañez
Mishell Marcela Ramos de la Hoz

Diseño, diagramación e ilustración
Mónica Liliana López Piratoba

Corrección de estilo
Paul Cifuentes Velásquez

Bogotá D.C., noviembre 2022

<http://www.icfes.gov.co/>

<https://www.facebook.com/icfescol>

<https://www.instagram.com/icfescol/>

<https://twitter.com/ICFEScol>

Todos los derechos de autor reservados ©.

Cite este documento así:

APA 7

Instituto Colombiano para la Evaluación de la Educación (Icfes). (2022). Informe del comportamiento psicométrico de los ítems piloto de Saber 7°

Icontec

INSTITUTO COLOMBIANO PARA LA EVALUACIÓN DE LA EDUCACIÓN. Informe del comportamiento psicométrico de los ítems piloto de Saber 7° Aplicación 2022. Bogotá: Icfes, 2022.

Para tener acceso a las funcionalidades interactivas de este documento, haga uso de las siguientes herramientas actualizadas al 2020, según su sistema operativo:

 Windows

 macOS



Adobe
Acrobat



Google
Chrome



Opera



Mozilla
Firefox



Microsoft
Edge



Adobe
Acrobat



Google
Chrome



Opera

Términos y condiciones de uso para las publicaciones y obras que son propiedad del Icfes

El Instituto Colombiano para la Evaluación de la Educación (Icfes) pone a disposición de la comunidad educativa, y del público en general, **de forma gratuita y libre de cualquier cargo**, un conjunto de publicaciones disponibles en su portal www.icfes.gov.co. Estos materiales y documentos están normados por la presente política, y se encuentran protegidos por derechos de propiedad intelectual a favor del Icfes y del Ministerio de Educación Nacional. Si tiene conocimiento de alguna utilización contraria a lo establecido en estas condiciones de uso, por favor infórmenos al correo prensaicfes@icfes.gov.co.

Queda prohibido el uso o publicación total o parcial de este material con fines de lucro. **Únicamente está autorizado su uso para fines académicos e investigativos**. Ninguna persona, natural o jurídica, nacional o internacional, podrá vender, distribuir, alquilar, reproducir, transformar¹, promocionar o realizar acción alguna con la cual se lucre directa o indirectamente con este material². Esta publicación cuenta con el registro ISBN (International Standard Book Number o Número Normalizado Internacional para Libros), que facilita la identificación no solo de cada título, sino, también, de la autoría, la edición, el editor y el país en donde se edita.

En todo caso, cuando se haga uso parcial o total de los contenidos de esta publicación, el usuario deberá

¹ La transformación es la modificación de la obra a través de la creación de adaptaciones, traducciones, compilaciones, actualizaciones, revisiones, y, en general, cualquier modificación que se pueda realizar, haciendo que la nueva obra resultante se constituya en una obra derivada protegida por el derecho de autor, con la única diferencia, respecto de las obras originales, que aquellas requieren, para su realización, de la autorización expresa del autor o propietario para adaptar, traducir, compilar, etc. En este caso, el Icfes prohíbe la transformación de esta publicación.

consignar o hacer referencia a los créditos institucionales del Icfes, respetando los derechos de cita. En otras palabras, se podrá hacer uso de esta publicación si dicho uso se contempla en los fines aquí previstos. Es posible, entonces, transcribir pasajes del texto si se cita siempre la fuente de autor. Por supuesto, estas citas no deberían ser excesivas ni frecuentes para que, así, no se considere una reproducción simulada y sustancial que redunde en perjuicio del Icfes.

Asimismo, los logotipos institucionales son marcas registradas de propiedad exclusiva tanto del Ministerio de Educación Nacional como del Icfes, respectivamente. Por tanto, cuando su uso pueda causar confusión, los terceros no podrán usar las marcas de propiedad del Ministerio de Educación Nacional y del Icfes con signos idénticos o similares respecto a cualquier producto o servicio prestado por esta entidad.

En todo caso, queda prohibido su uso sin previa autorización expresa por parte del Icfes. La infracción de estos derechos se perseguirá civil y penalmente (en caso de que sea necesario), de acuerdo con las leyes nacionales y tratados internacionales aplicables.

El Icfes realizará cambios o revisiones periódicas a los presentes términos de uso y los actualizará en esta publicación.

² Como titular de la autoría conjunta de los derechos patrimoniales junto al Icfes, el Ministerio de Educación Nacional podrá usar, disponer, reproducir, distribuir y comunicar, sin limitación alguna de tiempo, modo y lugar, la batería de pruebas o ítems, guías de orientación y uso de resultados y demás temas educativos que posteriormente se ofrezcan, donde sea pertinente el uso de estos productos sin fines comerciales.

Tabla de contenidos

Introducción 07

Características de la aplicación de las pruebas Saber 7° 08

Objetivo de la prueba Saber 7° 08
Planeación técnica de la definición del instrumento y aplicación de pruebas 09
Cuestionarios aplicados 11
Caracterización de los participantes 12

Análisis psicométrico de las pruebas 13

Metodología de análisis psicométrico para las competencias básicas 13
Matemáticas 16
Competencias Comunicativas en Lenguaje: Lectura 17
Ciencias Naturales y Educación Ambiental 18
Competencias Ciudadanas 19
Pensamiento Ciudadano 19
Acciones y Actitudes 20

Cuestionarios auxiliares 23

Características de los Cuestionarios Auxiliares 24
Características del cuestionario de habilidades socioemocionales (HSE) 24
Características del cuestionario de factores asociados (FA) 26
Características del cuestionario socioeconómico (SE) 27
Metodología de análisis psicométrico para los CA 28
Resultados 29
Cuestionario de habilidades socioemocionales 29
Cuestionario de factores asociados 31
Cuestionario socioeconómico 32
Referencias 34

Índice de figuras

Figura 1. Modelo usado en el diseño de las pruebas Saber	9
Figura 2. Porcentaje de participación de estudiantes de grado 7 por región	12
Figura 3. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Matemáticas	16
Figura 4. Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba de Matemáticas	16
Figura 5. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Competencias Comunicativas en Lenguaje: Lectura	17
Figura 6. Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba de Competencias Comunicativas en Lenguaje: Lectura	17
Figura 7. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Ciencias Naturales y Educación Ambiental	18
Figura 8. Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba de Ciencias Naturales y Educación Ambiental	18
Figura 9. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Pensamiento Ciudadano	19
Figura 10. Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba Pensamiento Ciudadano	19
Figura 11. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Acciones y Actitudes ²⁰	19
Figura 12. Porcentaje promedio de respuestas, por opción, para las formas piloteadas de la prueba de Acciones y Actitudes	20

Índice de tablas

Tabla 1. Porcentajes de cobertura de estudiantes, por región, para cada sesión de aplicación.	10	Tabla 13. Estructura del cuestionario de factores asociados por componente y evidencia	24
Tabla 2. Saber 7°: Estructura del cuadernillo estándar grupo 1	11	Tabla 14. Número de ítems por componente del cuestionario de factores asociados	25
Tabla 3. Saber 7°: Estructura del cuadernillo estándar grupo 2	11	Tabla 15. Número de preguntas del Cuestionario Socioeconómico para grado séptimo	25
Tabla 4. Saber 7°: Estructura del cuadernillo para habilidades socioemocionales y factores asociados	11	Tabla 16. Análisis de distribución de frecuencias de respuesta por forma del Cuestionario de Habilidades Socioemocionales – grado séptimo	27
Tabla 5. Valores de referencia para los Estadísticos y parámetros psicométricos	15	Tabla 17. Análisis de distribución de frecuencias de respuesta por evidencia del Cuestionario de Habilidades Socioemocionales – grado séptimo	27
Tabla 6. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Matemáticas	16	Tabla 18. Análisis confiabilidad por índices y subíndices del Cuestionario de Habilidades Socioemocionales – grado séptimo	28
Tabla 7. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Competencias Comunicativas en Lenguaje: Lectura	17	Tabla 19. Análisis de distribución de frecuencias de respuesta por forma del cuestionario de factores asociados	29
Tabla 8. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Ciencias Naturales y Educación Ambiental	18	Tabla 20. Análisis de distribución de frecuencias de respuesta por componentes y evidencias del cuestionario de factores asociados	29
Tabla 9. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Pensamiento Ciudadano	19	Tabla 21. Valores del Alpha de Cronbach para los componentes del cuestionario de factores asociados	30
Tabla 10. Valores del Alpha de Cronbach para afirmaciones y evidencias de la prueba de Acciones y Actitudes	21	Tabla 22. Análisis univariados por componente del Cuestionario socioeconómico para grado séptimo	31
Tabla 11. Valores del Alpha de Cronbach para afirmaciones y evidencias de la prueba de Acciones y Actitudes	23		
Tabla 12. Número de ítems por dimensión y forma del Cuestionario de Habilidades Socioemocionales	23		

Introducción

El presente informe se rinde en el contexto del contrato interadministrativo número CO1.PCCNTR.3443608 de 2022, suscrito entre el Ministerio de Educación Nacional (MEN) y el Instituto Colombiano para la Evaluación de la Educación (Icfes). El objetivo de este documento es presentar los resultados de los análisis psicométricos realizados al pilotaje de las pruebas Saber 7°, con el fin de determinar las características de la medición de de la aplicación realizada en el año 2022 con estudiantes de grado 8.

El documento está dividido en tres capítulos. El primero contiene información general sobre el contexto de las pruebas y su aplicación, información que se expone a profundidad en los documentos e informes construidos para dichos propósitos. Por lo anterior, en este capítulo solo se resaltan los elementos más relevantes para la contextualización de los resultados psicométricos. Además, también se describen las características de la muestra de estudiantes usada para los pilotajes de las diferentes formas de las pruebas de Saber 7°. En esta, se observa que la mayoría de estudiantes son de la región Caribe, del nivel socioeconómico 2 y están matriculados en establecimientos educativos del sector oficial que se encuentran en zonas urbanas, lo cual es un reflejo de las características de los estudiantes de grado 7, dado que se usó una muestra representativa a nivel nacional.

El segundo capítulo presenta los resultados de los análisis psicométricos para los pilotos de las pruebas de competencias básicas, de Saber 7°. Contiene una breve contextualización de las pruebas y de la metodología empleada para los análisis donde

se presentan los conceptos psicométricos que sustentan los análisis, además de la correspondiente interpretación de los resultados para las pruebas de Matemáticas, Lectura, Ciencias Naturales y Competencias Ciudadanas, la cual está compuesta por Pensamiento Ciudadano y Acciones y Actitudes. En esta, se podrá observar, por ejemplo, que la prueba que representó la mayor dificultad fue la de Matemáticas, mientras que la de menor dificultad fue la de Competencias Comunicativas en Lenguaje: Lectura, según el parámetro de dificultad del modelo logístico de dos parámetros usado para la calificación de los ítems de las pruebas.

Por último, en el tercer capítulo, contiene una breve contextualización del cuestionario socioemocional, del de factores asociados y del socioeconómico. Posteriormente, se exponen los resultados de los análisis psicométricos en los que se observa que, por ejemplo, los valores más altos de Alpha de Cronbach se encuentran para el índice de habilidades socioemocionales, en el caso del cuestionario de habilidades de este tipo, y el componente de involucramiento parental para el cuestionario de factores asociados.

Cada uno de los resultados presentados ilustra, de manera global, el comportamiento psicométrico de los instrumentos utilizados en las pruebas Saber 7°. Los datos producto de la aplicación aportan evidencias que sustentan la calidad de la medición realizada y, en últimas, la validez de los resultados obtenidos con los instrumentos.



CAPÍTULO 1.

Características de la aplicación de las pruebas Saber 7°

Objetivo de la prueba Saber 7°

En general, las pruebas Saber 3°, 5°, 7° y 9° buscan contribuir al mejoramiento de la calidad de la educación colombiana y, además, proveer información sobre el desarrollo de competencias esenciales para la vida personal, social y laboral (Icfes, 2022a). También, las pruebas Saber 7° son un punto de referencia importante para el seguimiento a las trayectorias escolares, ya que constituyen un punto de referencia bienal para la identificación y el análisis del desempeño de los estudiantes en el tránsito que realizan entre la básica primaria y la básica secundaria (Icfes, 2022b), por lo que resulta de gran importancia revisar la calidad de los instrumentos para asegurar la medición longitudinal propuesta.

Planeación técnica de la definición del instrumento y aplicación de pruebas

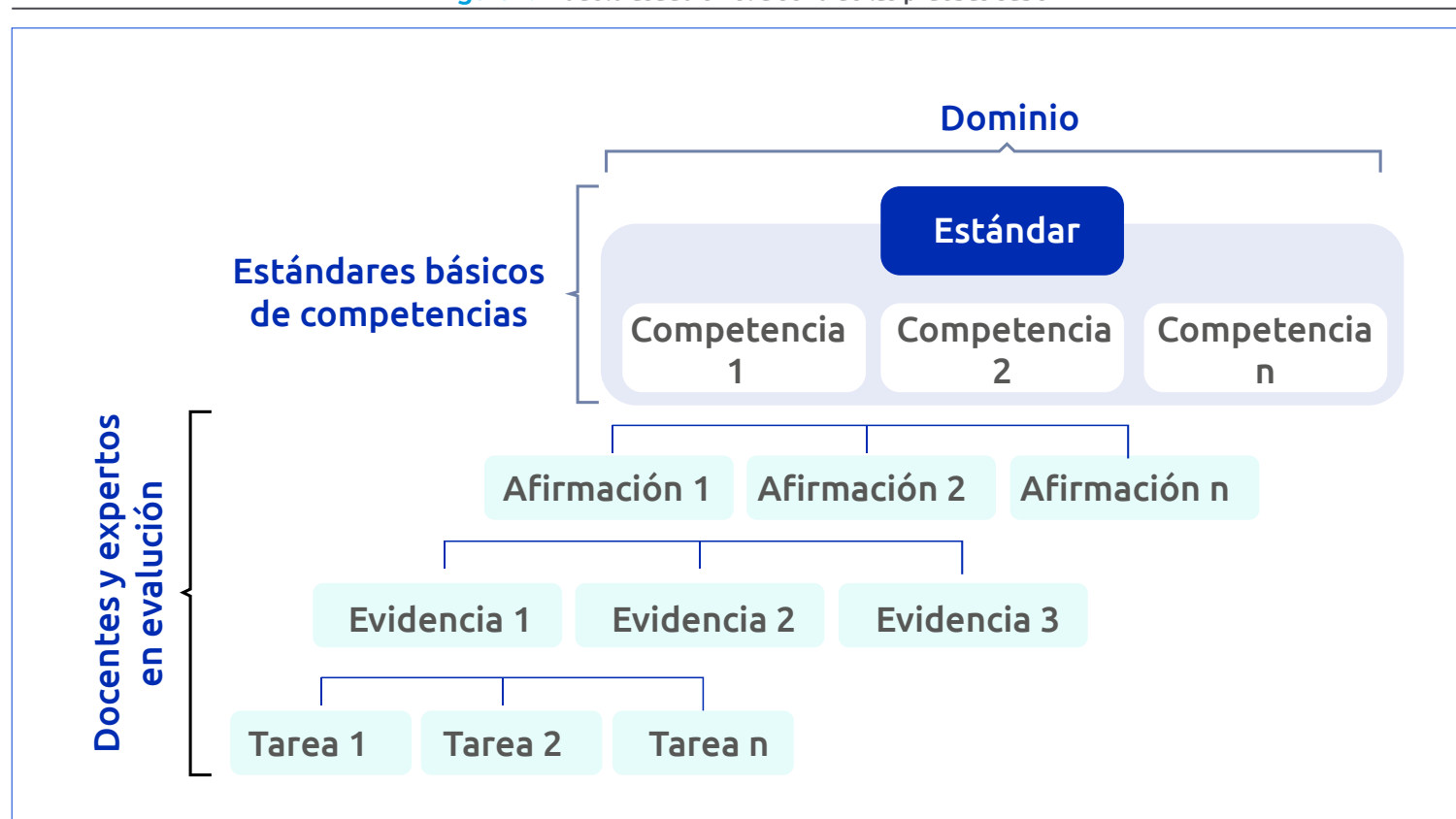
Marco de referencia para la elaboración de pruebas

La metodología usada para la construcción de los instrumentos aplicados en las pruebas Saber se conoce como diseño centrado en evidencias (DCE). El DCE es un marco para el diseño y construcción de diferentes tipos de evaluaciones que permite hacer explícitas las especificaciones de una evaluación, las cuales contienen los argumentos sustantivos que la sustentan (Mislevy, et. al, 2003). Esta metodología es usada en las evaluaciones que elabora el Icfes, ya que permite definir, de forma

esquemática, los elementos a evaluar dentro de las pruebas. También permite establecer conexiones entre las competencias básicas que deberían desarrollar los estudiantes y los elementos que las evalúan dentro de los instrumentos de medición.

La [Figura 1](#) muestra la estructura básica del modelo de DCE usado para la construcción de las pruebas Saber. Si el lector desea profundizar en este tema, puede consultar la guía introductoria al diseño centrado en evidencias elaborada por el Icfes (Icfes, 2018a).

Figura 1. Modelo usado en el diseño de las pruebas Saber



Fuente: "Guía introductoria al diseño centrado en evidencias". Icfes, 2018a, p. 8


Aplicación de las pruebas

Las pruebas Saber 7° fueron aplicadas en dos sesiones, a estudiantes de grado 8, durante el mes de abril de 2022. En la primera sesión de aplicación, se respondieron las preguntas de competencias básicas, así como el cuestionario socioeconómico. En la segunda sesión, se respondieron las preguntas de los cuestionarios de habilidades socioemocionales y factores asociados, estos últimos también diligenciados por docentes y directivos docentes de los grados y establecimientos participantes. Es importante tener en cuenta que los datos presentados en este informe corresponden, solamente, a los análisis realizados a las respuestas de los estudiantes.

Al igual que para la aplicación de 2021, el diseño de la muestra fue probabilístico estratificado y multietápico, en el que se consideraron las variables de grado, región, zona y sector, con lo cual se logró una representatividad para los resultados de las pruebas de competencias básicas, como de la información recolectada con los cuestionarios auxiliares para las variables de región, zona y sector, como estratos de la muestra.

Se proyectó que 1308 sedes participaran en la aplicación de las pruebas Saber 3°, 5°, 7° y 9°, pero, por razones como cambios en la matrícula y situaciones de orden público, entre otras, en 51 sedes se canceló la aplicación de las pruebas. A pesar de esto, el porcentaje de cobertura total de estudiantes, en grado 8, para la primera sesión, fue de 70,1%, mientras que, en la segunda sesión, fue de 67,9% (Icfes, 2022a). La Tabla 1 muestra los porcentajes de cobertura por región.

Tabla 1. Porcentajes de cobertura de estudiantes, por región, para cada sesión de aplicación.

Región	Primera sesión	Segunda sesión
 Caribe	71,4	69,2
 Centro Oriente	72,9	70,9
 Centro Sur	68,1	66,2
 Eje Cafetero – Antioquia	69,8	66,1
 Llano	65,0	63,9
 Pacífico	70,1	67,9
Total por grado	70,1	67,9

Los cuestionarios fueron aplicados en formato de lápiz y papel, con excepción de los que fueron diligenciados por personas con discapacidad, los cuales fueron aplicados a través de un computador, en las sedes que contaban con computadores y conectividad a internet.

Cuestionarios aplicados

Como se mencionó anteriormente, los cuestionarios aplicados para Saber 7° se correspondieron con las dos sesiones de aplicación. Para la primera sesión, se incluyeron, en el cuadernillo, las competencias básicas y el cuestionario socioeconómico; en la segunda sesión, se entregaron los cuadernillos con las preguntas para habilidades socioemocionales y factores asociados. Un elemento para tener en cuenta es que las pruebas de competencias básicas no se aplicaban, por completo, en un solo cuadernillo. Por esta razón, dos cuadernillos fueron entregados durante la primera sesión a dos grupos de estudiantes, los cuales contenían cuatro pruebas de competencias básicas y el cuestionario socioeconómico. Las [tablas 2 y 3](#) muestran la composición de cada uno.

Los cuadernillos para personas con discapacidad estaban organizados de la misma forma que se muestra en las [tablas 2 y 3](#), aunque los tiempos de aplicación aumentaban en 45 minutos. Por otra parte, la [Tabla 4](#) presenta la distribución de preguntas en los cuadernillos de la segunda sesión. Es importante tener en cuenta que los estudiantes respondieron 104 preguntas en esta sesión: 62 de habilidades socioemocionales y 42 de factores asociados.

Tabla 2. Saber 7°: Estructura del cuadernillo estándar grupo 1

	Número de preguntas Grado 7	Tiempo de aplicación
Matemáticas	36	4 horas y 20 minutos
Competencias Comunicativas en Lenguaje: Lectura	36	
Competencias Comunicativas en Lenguaje: Escritura	1	
Ciencias Naturales y Educación Ambiental	30	
Cuestionario socioeconómico	33	
Total de preguntas	136	

Tabla 3. Saber 7°: Estructura del cuadernillo estándar grupo 2

	Número de preguntas Grado 7	Tiempo de aplicación
Matemáticas	36	4 horas y 20 minutos
Competencias Comunicativas en Lenguaje: Lectura	36	
Competencias Ciudadanas: Acciones y Actitudes	60	
Competencias Ciudadanas: Pensamiento Ciudadano	30	
Cuestionario socioeconómico	33	
Total de preguntas	195	

Tabla 4. Saber 7°: Estructura del cuadernillo para habilidades socioemocionales y factores asociados

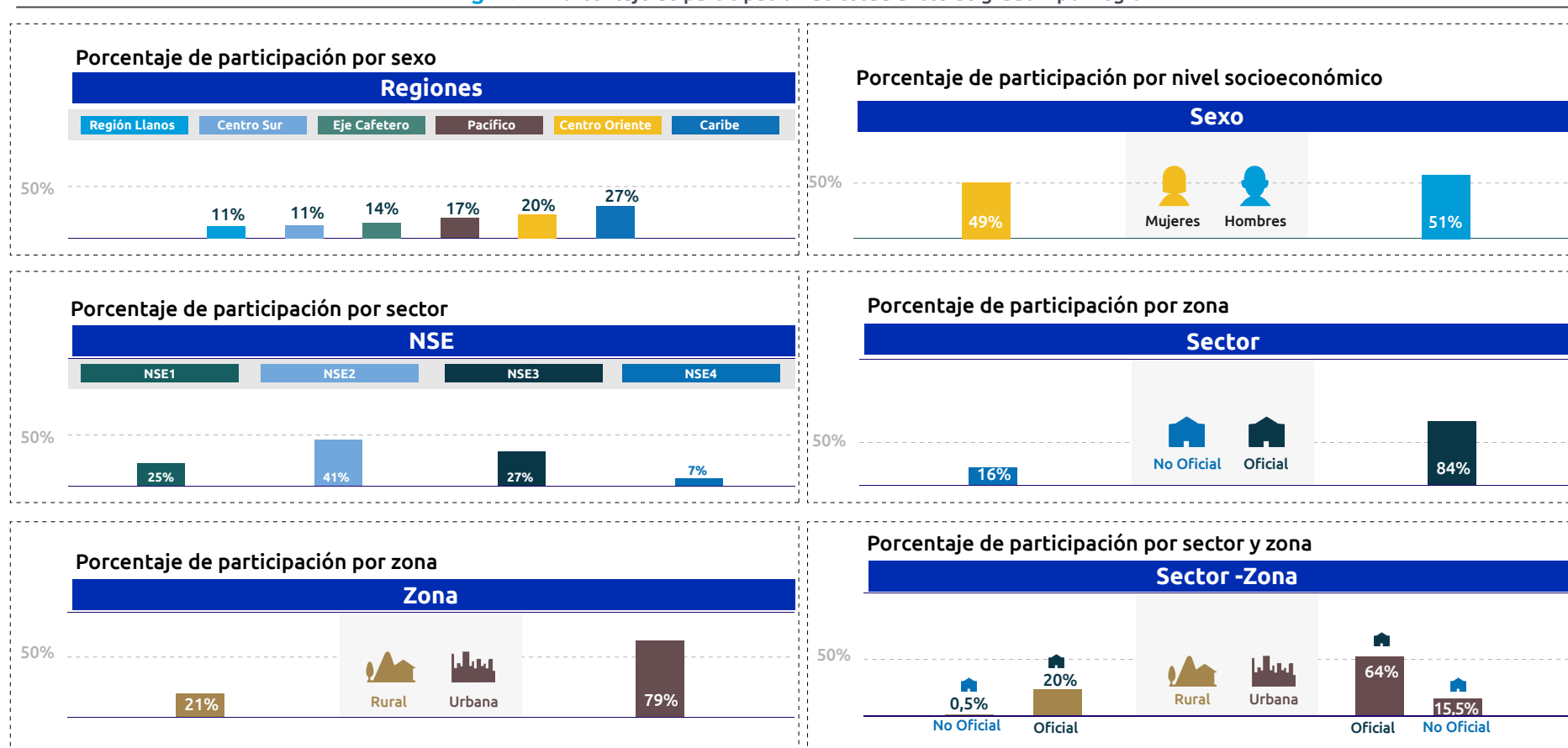
Cuestionario de habilidades socioemocionales	Cuestionario de factores asociados - estudiantes	Cuestionario de factores asociados - docentes	Cuestionario de factores asociados - rectores
62	42	36	37

Caracterización de los participantes

43.169 estudiantes respondieron la prueba Saber 7°. La Figura 2 indica el porcentaje de participación en la muestra de los estudiantes para las variables de región, sexo, nivel socioeconómico, además de las características de los establecimientos educativos en las variables sector, zona y la combinación de sector-zona. Respecto a los estudiantes, a nivel de regiones, la que tuvo una mayor participación fue la región Caribe, mientras que las regiones de Llanos y Centro Sur las que menor participación tienen, con un 16% menos que Caribe. Por otra parte, el porcentaje de hombres participantes es 2% mayor que el de mujeres y la mayor parte de los estudiantes fue clasificada en el nivel socioeconómico 2, así como la participación de los niveles 1 y 3 es similar.

Respecto a las características de los establecimientos educativos, se observa que la cantidad de colegios oficiales es cinco veces mayor que la cantidad de no oficiales, así como la de establecimientos urbanos es tres veces mayor al de rurales. En concordancia con estos datos, al hacer el cruce entre las variables sector y zona, más del 60% de los establecimientos son oficiales ubicados en zona urbana, mientras que solo el 1% son no oficiales ubicados en zona rural. Estos datos son el reflejo de la distribución de estas variables a nivel nacional, dado que la muestra seleccionada es representativa del país.

Figura 2. Porcentaje de participación de estudiantes de grado 7 por región



CAPÍTULO 2.

Análisis psicométrico de las pruebas

Metodología de análisis psicométrico para las competencias básicas

Los resultados se presentan en términos de porcentajes de respuestas correctas, omisiones, así como algunos estadísticos y parámetros en el contexto de las dos teorías dominantes de la medición en educación y psicología: la Teoría Clásica de los Test (TCT) y la Teoría de Respuesta al Ítem (TRI). La primera fue una propuesta pionera para el análisis de las puntuaciones de los test y dominó este campo por cerca de 70 años. A través de modelos lineales, la TCT analiza las propiedades de los test. Según Muñiz (2010), esta teoría cuenta con tres supuestos que se describen a continuación:

- La puntuación verdadera de una persona en un test es igual a su puntuación promedio si se le aplicara infinitas veces.
- No hay relación entre la cuantía de la puntuación verdadera y la cantidad de error que afecta a dichas puntuaciones.
- Los errores de medida en un test no están relacionados con los errores en otro test para un mismo grupo de personas.

La TCT ha recibido algunas críticas, entre las que se encuentran la imposibilidad de comprobar los supuestos descritos anteriormente y de comparar los resultados de una persona en diferentes instrumentos que miden lo mismo y la variación de las propiedades de los test según la muestra de personas que se use para estimarlas (Muñiz, 2010).

La segunda, la teoría de respuesta al ítem (TRI), entre otras cosas, proporciona algunas soluciones a las críticas de la TCT. La TRI se sirve de modelos matemáticos más complejos para analizar los ítems que componen las pruebas y usa una sola escala de medida: la escala de habilidad. Esta teoría tiene los siguientes supuestos (Muñiz, 1997, p. 19):

- Hay una relación funcional entre los valores de la variable a medir y la probabilidad de respuesta a nivel del ítem. Esta relación se representa en la curva característica de cada ítem y se describe, en general, en términos de uno o más de los siguientes parámetros según el modelo logístico utilizado (Rasch, de uno, de dos o de tres parámetros): discriminación (parámetro a), dificultad (parámetro b) y probabilidad de respuesta por azar (parámetro c). Los primeros dos parámetros se describirán más adelante.
- Los ítems miden, esencialmente, una dimensión. Este supuesto es llamado, generalmente, “unidimensionalidad” y se entiende que el ítem debe medir únicamente la dimensión, factor y/o aspecto para el que fue construido. Se reconoce que, en la práctica, este requerimiento puede ser difícil de cumplir (por ejemplo, que un ítem de matemáticas no se vea influido por el nivel de lectura del evaluado). Por lo tanto, desde la construcción, se debe asegurar que el ítem se enfoque en medir, principalmente, el aspecto para el que fue construido.

- La respuesta a un ítem no es influida por la respuesta a los otros. Este supuesto es conocido como “independencia local” y se considera que se cumple matemáticamente si el supuesto de unidimensionalidad es satisfecho.

El énfasis en cada una de las teorías es diferente (test vs. ítem), así como los modelos matemáticos con los que hacen los análisis (lineales vs. logísticos, de ojiva normal u otros). También, hay otras diferencias en el ámbito de lo práctico, como la escala en la que se expresan las puntuaciones o los tamaños de muestra necesarios para hacer los cálculos. En este punto, la TRI presenta un criterio más difícil de cumplir, ya que requiere de tamaños de muestra de, al menos, 500 evaluados para sus cálculos (Muñiz, 2010).

Aunque se puede considerar que la TRI es la teoría idónea para el análisis psicométrico, lo cierto es que las dos son usadas, ya que, como lo dice Muñiz (1997), “la combinación de ambos acercamientos, clásico y TRI, en función de las circunstancias, parece lo más recomendable como sucede habitualmente en la práctica” (p. 57). Por esta razón, con el objetivo de brindar un panorama general de la calidad de los instrumentos aplicados, los estadísticos o parámetros psicométricos que se presentan para describir el piloto objeto de este informe provienen de las dos teorías. A continuación, se describe cada uno de ellos:

Porcentaje de omisiones: Indica el porcentaje promedio de preguntas que no fueron contestadas. Algunas interpretaciones de este indicador están relacionadas con el tiempo límite para responder las pruebas (no hay suficiente tiempo para responder las

preguntas), problemas en el proceso de respuesta por parte del evaluado (saltar la pregunta o el espacio para consignar la respuesta, no saber la respuesta o negarse a contestar, entre otros), dificultades con las preguntas del cuadernillo (no se da información para responder, no hay respuesta posible, inconvenientes para leer la pregunta, entre otros) o pérdida de información en la base de datos (De Leeuw et al., 2003).

Porcentaje de respuestas correctas: Indica el porcentaje promedio de respuestas correctas a las preguntas. Este indicador es una forma de ver la dificultad de las preguntas en el grupo de evaluados, aunque no es un estimador robusto de esta, como lo es el indicador de dificultad de la TRI.

Alpha de Cronbach: Es una medida de la consistencia interna de un test o de la forma piloteada, en este caso. En palabras de Cervantes (2005), “separa del conjunto la variación que corresponde a factores comunes de los ítems y la que corresponde a factores únicos de cada uno de ellos” (p. 10). Este estadístico es uno de los más usados en la TCT para evaluar la confiabilidad de un test (cuando este es unidimensional) o de una parte dentro este (cuando está compuesto de diferentes pruebas o escalas). Algunos de los factores que pueden afectar el resultado de este estadístico son el tamaño y heterogeneidad de la muestra o el número de ítems que componen el test o la escala (Morales, 2007, p. 203).

Correlación punto biserial: Es una medida que da cuenta de la relación entre la respuesta a un ítem y el resultado en la prueba (Pardo et al., 2005). En general, se emplea como un indicador de la discriminación

de los ítems, es decir, sirve para identificar el grado en que el ítem permite diferenciar las personas con un nivel alto del objeto de medición (atributo, rasgo, conocimiento, habilidad, etc.), de aquellas con un bajo nivel (Mahias y Polloni, 2019; Argüello, 2019). Es de común uso en la TCT.

Parámetro de discriminación de la TRI: Este indicador comparte la misma interpretación de la correlación punto biserial, aunque su cálculo no se basa en una correlación. Este parámetro “representa el cambio en la probabilidad de contestar correctamente el ítem conforme cambia el nivel de habilidad, en el punto correspondiente a la dificultad” (Icfes, 2018b, p. 3). La interpretación de este parámetro es más precisa en los valores cercanos al valor de la habilidad asociado al parámetro b del ítem (Baker, 2001).

Parámetro de dificultad de la TRI: Este indicador “representa la habilidad correspondiente a la probabilidad de contestar correctamente el ítem.” (Icfes, 2018b, p. 3), con una probabilidad del 50%. En el caso de las pruebas Saber, sus valores están estandarizados con media cero (0) y desviación estándar uno (1). Este es un parámetro que indica la ubicación del ítem en la escala de habilidad de la TRI (Baker, 2001), por lo tanto, a valores más bajos, el ítem es más fácil y, a valores más altos, indicaría mayor dificultad para los evaluados.

Los valores del porcentaje de omisiones y de respuestas correctas estarán representados en diagramas de barras, mientras que los valores para los demás indicadores serán resumidos en tablas. Con respecto a estos últimos, la **Tabla 5** presenta algunos valores de referencia para clasificar sus resultados.

Tabla 5. Valores de referencia para los estadísticos y parámetros psicométricos

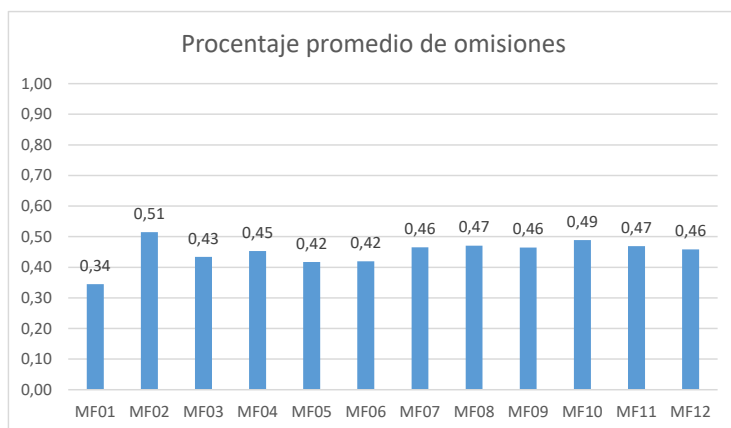
Indicador	Valor de referencia	Interpretación
Alpha de Cronbach (Kline, 2000; DeVellis, 2016)	$\geq 0,9$	Excelente
	$0,8 \leq < 0,9$	Buena
	$0,7 \leq < 0,8$	Aceptable
	$0,6 \leq < 0,7$	Cuestionable
	$0,5 \leq < 0,6$	Pobre
Correlación punto biserial (Ebel, 1965)	Mayor o igual a 0,4	El ítem discrimina muy bien.
	Entre 0,30 y 0,39	El ítem discrimina bien.
	Entre 0,20 y 0,29	El ítem tiene una discriminación suficiente.
	Menor a 0,20	El ítem no discrimina.
Discriminación TRI (Baker, 2001)	+ Infinito	Perfecta
	$> 1,70$	Muy alta
	1,35 - 1,69	Alta
	0,65 - 1,34	Moderada
	0,35 - 0,64	Baja
	0,01 - 0,34	Muy baja
	0	Ninguna
Dificultad TRI (Elaboración propia)	> 2	Muy difícil
	0 a 2	Difícil
	-2 a 0	Fácil
	< -2	Muy fácil

Con base en estos criterios, a continuación, se realizará la descripción de los resultados de análisis psicométrico para los ítems del piloto de Saber 7°.

Matemáticas

Para la prueba de Matemáticas se conformaron 12 formas combinatorias de bloques de ítems, las cuales se toman como referencia para mostrar los resultados de los análisis. La [Figura 3](#) muestra que los porcentajes promedio de omisión se encuentran entre 0,42% y 0,51% para 11 de las 12 formas, destacando que dicho porcentaje es de 0,34 para la forma 1, mostrando que es la que menos omisión tuvo.

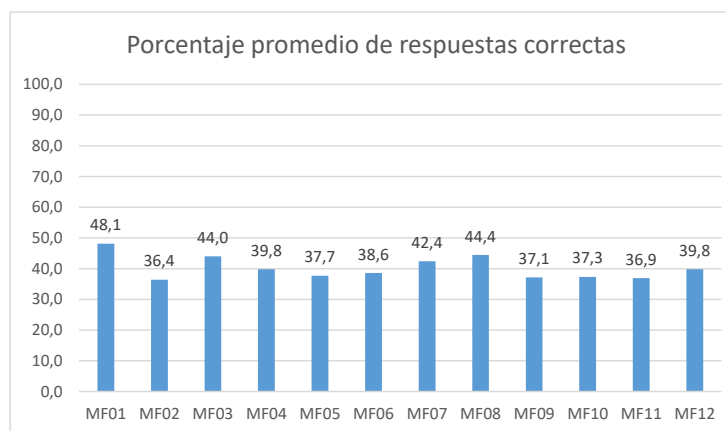
Figura 3. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Matemáticas



Por otra parte, la [Figura 4](#) muestra el porcentaje promedio de respuestas correctas por forma. Se puede observar que, en ningún caso, el porcentaje fue mayor al 50%, lo que indica que, en general, menos de la mitad de los estudiantes respondieron correctamente a las preguntas de la prueba de Matemáticas. Un elemento más que se puede tener en cuenta es que, para ocho de las 12 formas, el porcentaje promedio de respuestas correctas es menor al 40%, aunque mayor al 36%.

En la [Tabla 6](#), se encuentran los valores para los estadísticos y parámetros estimados, cuyos valores de referencia se encuentran en la [Tabla 5](#). Según estos,

Figura 4. Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba de Matemáticas



se puede observar que los valores del coeficiente Alpha de Cronbach indicarían que seis formas estarían clasificadas en la categoría “aceptable” y la forma 1 estaría clasificada en la categoría “buena”, mientras las restantes estarían clasificadas como "cuestionable", en relación con la consistencia interna.

Por otra parte, respecto a las medidas de discriminación aplicadas, se observa, en el contexto de la TCT, que cuatro de las 12 están clasificadas dentro de la categoría “discrimina bien”, mientras que las formas restantes se clasifican en la categoría "discriminación suficiente", de la correlación biserial puntual. Este resultado no se corresponde con los valores para el indicador de discriminación de la TRI, donde todas las formas están clasificadas en la categoría “baja”. Es importante tener en cuenta que los resultados de los dos indicadores de discriminación están dados en términos de promedio, por lo que es posible que, dentro de las formas evaluadas, haya ítems que tengan mejores valores de discriminación.

El indicador de dificultad de la TRI muestra que, en promedio, las preguntas contenidas en las formas de la prueba fueron difíciles para los estudiantes, por lo que cinco se encuentran clasificadas en la categoría “difícil” y las restantes en la categoría “muy difícil”. Este resultado se relaciona con los datos presentados en la [Figura 4](#).

Tabla 6. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Matemáticas

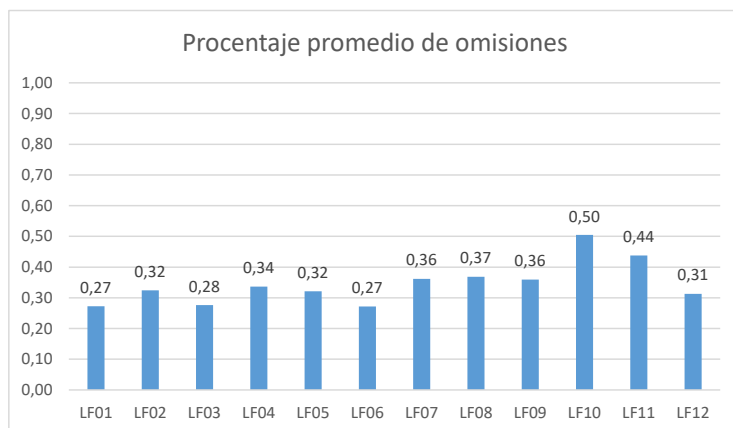
Forma	Alpha de Cronbach	Promedio correlación biserial	Discriminación TRI	Dificultad TRI
MF01	0,82	0,39	0,60	1,00
MF02	0,65	0,23	0,37	2,77
MF03	0,75	0,31	0,48	3,80
MF04	0,66	0,24	0,41	3,49
MF05	0,71	0,27	0,43	1,78
MF06	0,71	0,27	0,39	4,40
MF07	0,74	0,30	0,44	2,12
MF08	0,76	0,32	0,56	0,92
MF09	0,66	0,24	0,38	3,90
MF10	0,66	0,25	0,42	3,30
MF11	0,69	0,26	0,38	1,81
MF12	0,73	0,29	0,43	1,72

Los datos presentados para la prueba de Matemáticas muestran que, en general, se les dificultó a los estudiantes; además, presenta unos valores de discriminación muy bajos. Por otra parte, la confiabilidad de las formas, evaluada con el Alpha de Cronbach, se encuentra entre aceptable y buena.

Competencias Comunicativas en Lenguaje: Lectura

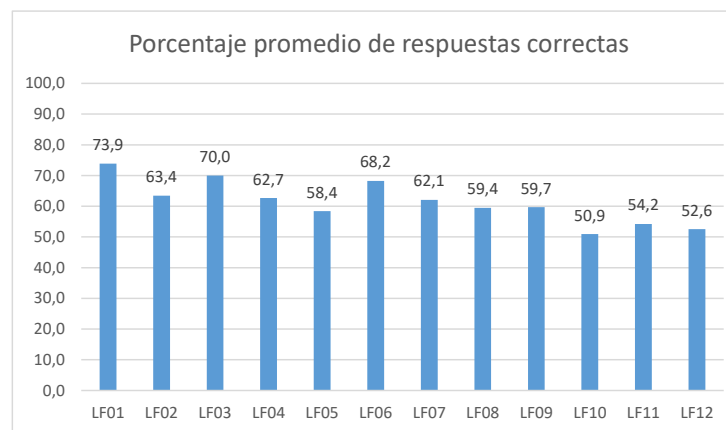
Como en la prueba de Matemáticas, esta agrupó los bloques de ítems en 12 formas, de las cuales se presentan los resultados a continuación. La [Figura 5](#) muestra que 11 de las doce formas tienen porcentajes promedio de omisión entre 0,27% y 0,44% y solo la forma 10 alcanza un 0,5% de omisiones.

Figura 5. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Competencias Comunicativas en Lenguaje: Lectura



La [Figura 6](#) muestra que el porcentaje de respuestas correctas, para las 12 formas, se encuentra entre 50,9% (forma 10) y el 73,9% (forma 1). Este rango de valores indica que algunas de las formas resultaron más difíciles que otras, principalmente, las formas 10, 11 y 12.

Figura 6. Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba de Competencias Comunicativas en Lenguaje: Lectura



En la [Tabla 7](#), se observa que, según los valores del Alpha de Cronbach, todas las formas se podrían clasificar en la categoría “buena” respecto a la consistencia interna. Por otra parte, aunque el indicador de la TCT (correlación biserial), para discriminación, indica que los ítems discriminan muy bien, el indicador basado en la TRI señala que 8 de las 12 formas tendrían una discriminación promedio moderada, mientras que las restantes se clasificarían como de discriminación baja. Es importante tener en cuenta que el análisis con la TRI es mucho más complejo y exigente, por lo que el que se presenten estas diferencias no puede ser considerado como un resultado incoherente. También, es importante recordar que son promedios, lo que implica que algunos de los ítems pueden tener una mejor discriminación de lo que se muestra en este resumen.

Respecto al indicador de dificultad basado en la TRI, se observa que los resultados son consistentes con

los presentados en la [Figura 6](#), dado que 9 formas se clasifican en la categoría “fácil”, mientras que las restantes se clasifican en la categoría “difícil”. Estas tres formas son la 10, 11 y 12, que fueron señaladas anteriormente como las que tuvieron un menor porcentaje de respuestas correctas.

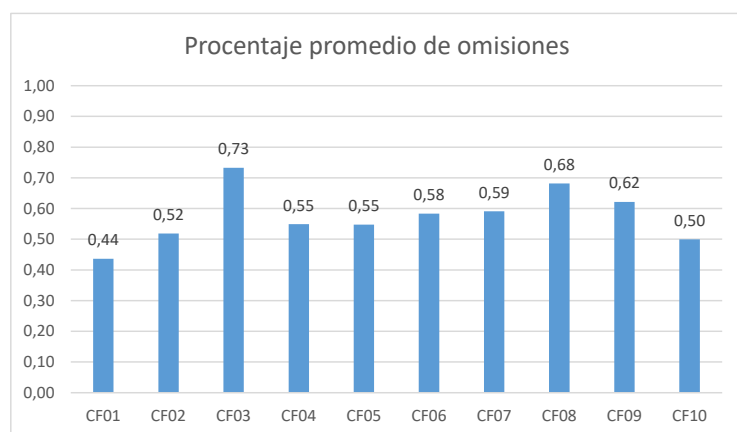
Tabla 7. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Competencias Comunicativas en Lenguaje: Lectura

Forma	Alpha de Cronbach	Promedio correlación biserial	Discriminación TRI	Dificultad TRI
LF01	0,87	0,55	0,75	-1,27
LF02	0,82	0,42	0,58	-0,85
LF03	0,83	0,48	0,67	-1,16
LF04	0,88	0,54	0,77	-0,43
LF05	0,86	0,46	0,61	-0,16
LF06	0,88	0,54	0,75	-0,88
LF07	0,90	0,56	0,78	-0,48
LF08	0,89	0,54	0,76	-0,12
LF09	0,86	0,48	0,67	-0,41
LF10	0,86	0,45	0,65	0,22
LF11	0,87	0,47	0,64	0,09
LF12	0,85	0,44	0,57	0,32

Ciencias Naturales y Educación Ambiental

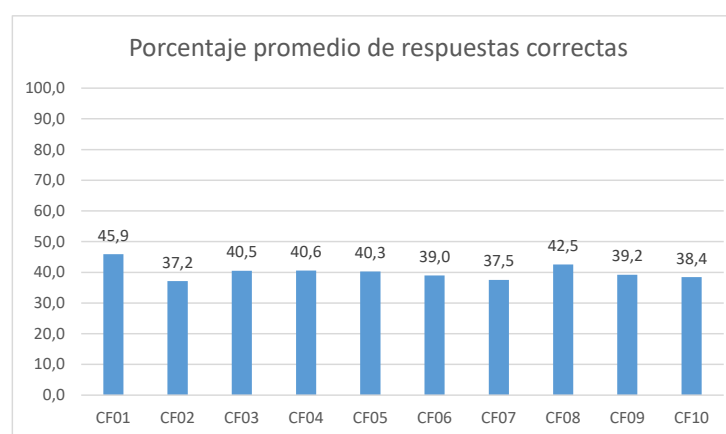
Para esta prueba, se combinaron los bloques de ítems en 10 formas, de las cuales se presentan los resultados a continuación. La [Figura 7](#) evidencia que 6 de las formas tienen porcentajes de omisión entre 0,5% y 0,59%, tres están por encima del 0,6% y la forma 1 presenta el menor porcentaje de omisión.

Figura 7. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Ciencias Naturales y Educación Ambiental



Entre tanto, en la [Figura 8](#), se advierte que el porcentaje promedio de respuestas correctas para esta prueba fue menor al 46% para todas las formas. Entre las formas 1 y 2, las cuales representan las de mayor y menor porcentaje de respuestas correctas, respectivamente, hay un 8,7% de diferencia, dato que muestra que las formas tuvieron una dificultad similar, tal como se observa en el gráfico.

Figura 8. Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba de Ciencias Naturales y Educación Ambiental



Para esta prueba, ocho de las formas tienen una consistencia interna aceptable, mientras que para las restantes esta es buena (ver [Tabla 8](#)). Respecto a los indicadores de discriminación, el que está basado en la TCT (correlación biserial) indica que nueve de las formas discriminan bien y solo la forma 1 discrimina muy bien. Por otra parte, el indicador basado en la TRI expresa que todas las formas tienen una baja discriminación promedio.

El indicador de dificultad de la TRI muestra que dos de las formas tienen valores por encima de 2, lo que las clasifica como muy difíciles, mientras que las restantes se clasifican como difíciles. Este resultado es similar al presentado en la [Figura 8](#), con lo que se confirma que todas las formas resultaron difíciles para los estudiantes de grado 7.

Tabla 8. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Ciencias Naturales y Educación Ambiental

Forma	Alpha de Cronbach	Promedio correlación biserial	Discriminación TRI	Dificultad TRI
CF01	0,85	0,47	0,64	0,42
CF02	0,72	0,30	0,42	1,97
CF03	0,79	0,37	0,49	0,77
CF04	0,77	0,35	0,50	1,46
CF05	0,79	0,37	0,46	3,35
CF06	0,76	0,35	0,46	1,54
CF07	0,73	0,31	0,40	2,66
CF08	0,79	0,39	0,52	0,86
CF09	0,80	0,39	0,48	1,95
CF10	0,74	0,32	0,43	1,15

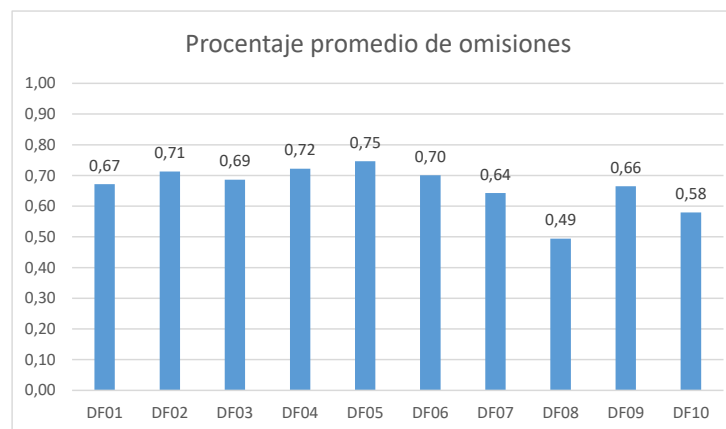
Competencias Ciudadanas

Las competencias ciudadanas son evaluadas a través de dos pruebas: Pensamiento Ciudadano y Acciones y Actitudes. Los datos que se presentan para las dos pruebas son diferentes, dado que las escalas de calificación nos son las mismas. Por lo tanto, en primer lugar, se presentan los resultados de la prueba de Pensamiento Ciudadano, que se evaluarán con las mismas categorías que se utilizaron para las pruebas anteriores, mientras que, para la prueba de Acciones y Actitudes, se presentarán porcentajes por omisión por evidencias, porcentajes de respuesta por opción de respuesta, según las evidencias y valores del Alpha de Cronbach para cada afirmación y evidencia de la prueba.

Pensamiento ciudadano

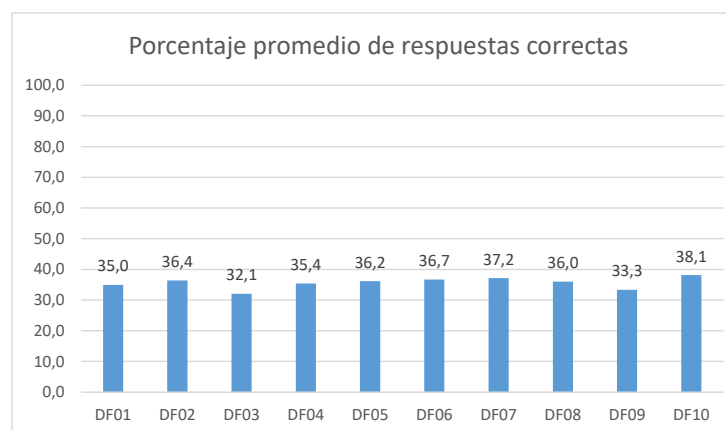
Como para la prueba de Ciencias Naturales y Educación Ambiental, esta prueba tiene 10 formas. Como se observa en la [Figura 9](#), los porcentajes de omisiones varían entre 0,49% (forma 8) y 0,75% (forma 5). Cuatro de las formas tienen un porcentaje de omisión mayor o igual al 0,7% y cuatro tienen porcentajes entre 0,64% y 0,69%. Estos datos mostrarían que la prueba es la que tiene los mayores porcentajes de omisión, entre las pruebas que se califican según una respuesta correcta.

Figura 9. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Pensamiento Ciudadano



La [Figura 10](#) deja ver el porcentaje de respuestas correctas para las 10 formas. Se observa que la diferencia entre la forma con mayor número de respuestas correctas (forma 10) y la menor (forma 3) es de 6% y, en ningún caso, este porcentaje es mayor a 38,1%, dato que indicaría que esta prueba, en general, fue la más difícil de Saber 7°.

Figura 10. . Porcentaje promedio de respuestas correctas para las formas piloteadas de la prueba Pensamiento Ciudadano



La [Tabla 9](#) muestra que 9 de las formas tienen una consistencia interna aceptable y la forma 10 sería clasificada en la categoría "buena", según los valores del Alpha de Cronbach. Respecto a la discriminación, de acuerdo con el indicador basado en la TCT (correlación biserial), 9 de las 10 formas contienen ítems que, en promedio, discriminan bien, mientras que, conforme el indicador basado en la TRI, en promedio, la discriminación de los ítems es baja.

El parámetro de dificultad de la TRI evidencia que 8 de las formas se clasifican como difíciles, mientras que las dos restantes estarían clasificadas como muy difíciles. Aunque los resultados son consistentes con los de la [Figura 10](#), no muestran una mayor dificultad que los de otras pruebas, contrario a lo que se comentó respecto a los porcentajes de respuestas correctas.

Tabla 9. Estadísticos y parámetros psicométricos para las formas piloteadas de la prueba de Pensamiento Ciudadano

Forma	Alpha de Cronbach	Promedio correlación biserial	Discriminación TRI	Dificultad TRI
DF01	0,73	0,32	0,48	1,79
DF02	0,78	0,37	0,48	1,16
DF03	0,71	0,30	0,42	2,53
DF04	0,77	0,35	0,51	1,03
DF05	0,70	0,29	0,44	1,16
DF06	0,76	0,35	0,46	1,14
DF07	0,79	0,38	0,49	1,07
DF08	0,76	0,35	0,45	1,09
DF09	0,74	0,33	0,46	2,11
DF10	0,80	0,39	0,57	0,84

Acciones y Actitudes

Como se comentó al inicio de esta sección, la escala de respuesta de la prueba de Acciones y Actitudes es diferente de las anteriormente presentadas, ya que esta prueba utiliza una escala Likert de frecuencia para las preguntas relacionadas con Acciones, y una escala Likert de acuerdo para las preguntas relacionadas con Actitudes. Otro elemento para tener en cuenta es que la calificación de esta prueba se realiza sobre la teoría clásica de los test, por lo que se presentan solo estadísticos descriptivos y valores de Alpha de Cronbach para los ítems agrupados en las evidencias y afirmaciones de la prueba. A continuación, se exponen los resultados.

La **Figura 11** expone el porcentaje promedio de omisiones a las preguntas de cada una de las evidencias. Se observa que, en general, los mayores promedios de omisión se encuentran para las evidencias relacionadas

con Actitudes, donde solo la evidencia 1 de la afirmación 1 tiene un porcentaje promedio por debajo al de las evidencias relacionadas con Acciones.

También se puede destacar que las evidencias de las afirmaciones 3 y 4 relacionadas con Actitudes, además de la evidencia 4 de la afirmación 2, son las que tienen los porcentajes más altos de omisión de la prueba, los cuales se encuentran entre 2,19% y 2,75%. Por otra parte, 5 de las 6 evidencias de Acciones tienen porcentajes de omisión iguales o menores al 1%, los cuales, sin embargo, resultan mayores que los de los resultados presentados para una gran parte de las formas en las demás pruebas de Saber 7°.

La **Figura 12** registra el porcentaje promedio por opción de respuesta para los dos componentes de la prueba de Acciones y Actitudes. Es importante tener en cuenta

que la calificación de las opciones puede ser directa y se asigna el menor valor a las opciones que indican menor frecuencia o más desacuerdo. A partir de esta valoración, aumenta la calificación. Por otra parte, la calificación puede ser inversa, en la que la mayor calificación se les asigna a las opciones que indican menor frecuencia o más desacuerdo y, a partir de estas, se va disminuyendo la calificación.

Para las preguntas de esta prueba, se observa que hay consistencia en la respuesta de los evaluados según la escala y sentido de calificación, dado que los porcentajes de estudiantes que las seleccionan son mayores, de acuerdo con el aumento esta última. La única excepción a este comportamiento son las opciones de calificación directa relacionada con Acciones, donde la opción “casi siempre” obtiene un 3,2% más de respuestas que la opción “siempre”.

Figura 11. Porcentaje promedio de omisiones para las formas piloteadas de la prueba de Acciones y Actitudes

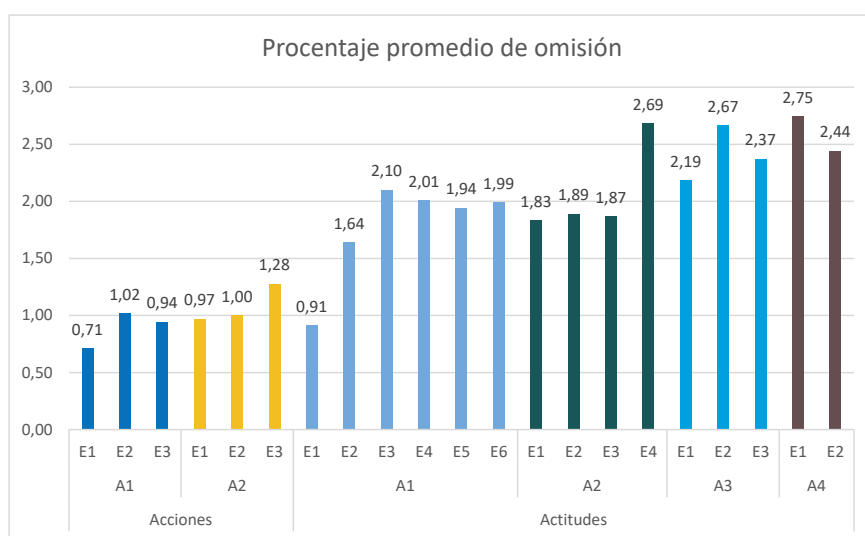
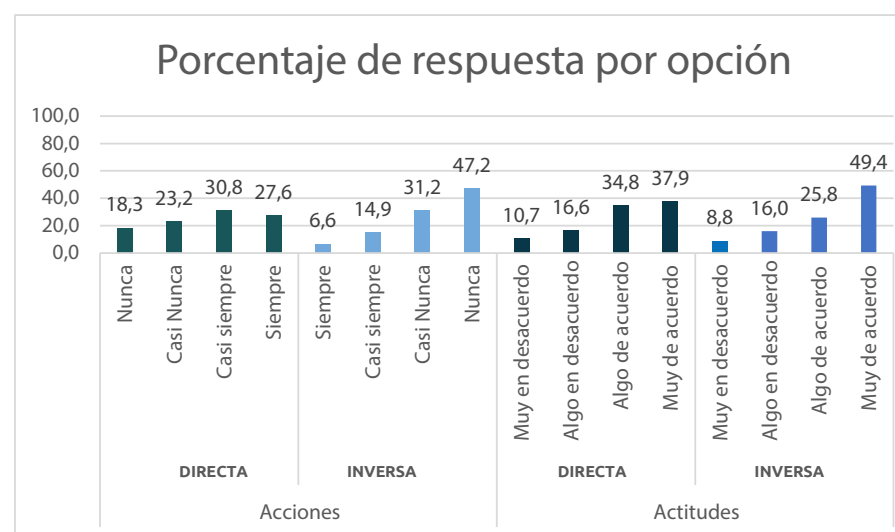


Figura 12. Porcentaje promedio de respuestas, por opción, para las formas piloteadas de la prueba de Acciones y Actitudes



La Tabla 10 puntualiza los valores del coeficiente Alpha de Cronbach para las afirmaciones y evidencias de Acciones y Actitudes. Los resultados se presentan para aquellas afirmaciones y evidencias en las que haya más de 3 ítems que hagan parte de la calificación, por lo que no se muestran resultados para todas las afirmaciones y todas las evidencias de la prueba.

Se observa que solo las afirmaciones relacionadas con Acciones, así como la evidencia 1.2 de la misma competencia, presentan valores entre 0,7 y 0,8 del coeficiente, lo que se traduce en una consistencia interna aceptable. Para las demás afirmaciones y evidencias, los valores no permiten inferir que haya una medición uniforme por medio de los ítems que los componen.

Conclusiones sobre resultados psicométricos de las competencias básicas

Con el fin de entender adecuadamente los resultados psicométricos de las competencias básicas para la prueba Saber 7°, se exponen algunas consideraciones teóricas para tener en cuenta.

La diferencia más importante entre la Teoría Clásica de los Test (TCT) y la Teoría de Respuesta al Ítem (TRI), es que en la TCT se utiliza una estimación común de la precisión de la medición que se supone igual para todos los individuos, independientemente de sus niveles de atributos, mientras que, por el lado de la TRI, la precisión de la medición depende del valor del atributo latente (Jabrayilov et al., 2016). Esto puede resultar en diferencias entre la TCT y la TRI con respecto a sus conclusiones sobre la importancia

Tabla 10. Valores del Alpha de Cronbach para afirmaciones y evidencias de la prueba de Acciones y Actitudes

Competencia	Afirmación	Alpha	Evidencia	Alpha
Acciones	1. Cuida las comunidades a las que pertenece a través de sus acciones cotidianas.	0,77	1.1. Privilegia la resolución pacífica de los conflictos, tanto propios como de otros, en sus comunidades cercanas.	0,56
			1.2. Cuida el respeto de los derechos en las relaciones interpersonales.	0,70
	2. Participa en la construcción colectiva de los asuntos públicos en las comunidades a las que pertenece.	0,79	2.2. Participa en procesos de toma de decisiones que afectan el bien común.	0,59
			2.3 Participa en acciones políticas y comunitarias a través de espacios alternativos.	0,63
Actitudes	1. Valora los derechos humanos.	0,52	NA	
	2. Valora las responsabilidades inherentes a la pertenencia a comunidades y la importancia de las instituciones.	0,65	2.2. Valora lo público.	0,55
	4. Valora la posibilidad de incidir en la comunidad como motivación de las acciones ciudadanas.	0,58	NA	

estadística de los parámetros calculados, ya que los indicadores que se basan en la TCT se van a concentrar en analizar una prueba o escala en general, en tanto que los indicadores de TRI van a enfocar los análisis al nivel del ítem. Así, los valores entregados en este informe pueden ser tomados como una referencia, pero se debe tener en cuenta que los ítems están distribuidos en la escala de habilidad para abarcar todos los niveles de ejecución de los participantes.

Observando el porcentaje de omisiones a lo largo de todas las formas en las cuatro competencias básicas, se puede considerar que, en general, fue bajo, dado que llega, como máximo, al 0,75%, en la forma 5 de la prueba de Pensamiento Ciudadano. Sin embargo, para la prueba de Acciones y Actitudes, donde la forma de calificación es diferente, estos llegan hasta 2,75%.

Los resultados generales para las pruebas de Matemáticas, Lectura, Ciencias Naturales y Pensamiento Ciudadano muestran que, en general, los indicadores de consistencia interna (Alpha de Cronbach), discriminación (correlación biserial), y dificultad (porcentaje de respuestas correctas) basados en la TCT son buenos.

Por su lado, los resultados basados en los indicadores de discriminación (parámetro a) y dificultad (parámetro b) basados en la TRI son mixtos. Puntualmente, mientras que los rangos de dificultad tienden a ubicarse dentro del rango entre 0 y 2 (difícil), los valores para los indicadores de discriminación de la TRI están entre 0,35 y 1,34, cayendo, en su mayoría, dentro de la categoría “baja”, con excepción de la prueba de Lectura en donde la mayoría entra en la categoría “moderada”.

Respecto a la dificultad, los dos parámetros (porcentaje promedio de respuestas correctas de la TCT y el parámetro b de la TRI) tienden a ser similares, lo que resulta coherente con lo reportado en el estudio de Macdonald y Paunonen (2002) con datos simulados. Tanto los resultados del porcentaje promedio de respuestas correctas, como el parámetro de dificultad, muestran que la prueba de Pensamiento Ciudadano fue la más difícil para los estudiantes, mientras que la de Lectura fue la más fácil.

Con respecto a la discriminación, los dos parámetros (correlación biserial puntual de la TCT y el parámetro a de la TRI) tienden a ser disímiles. Mientras que el parámetro de la TRI es, en general, más severo, el parámetro de la TCT tiende a ser más indulgente (Macdonald & Paunonen, 2002). Así, como se indicó arriba con el parámetro de discriminación, la mayoría de las formas de las pruebas fue considerada con un nivel de discriminación “baja”, excepto para la prueba de Lectura. Pero, al observar la correlación biserial puntual, solo para las formas de Matemáticas se considera que “los ítems tienen una discriminación suficiente”, mientras que la mayoría de las formas de las otras pruebas caen en el rango en el cual “el ítem discrimina bien” o superior.

Los resultados expuestos indican que las formas piloteadas para las pruebas Saber 7°, en general, muestran valores positivos con base en los criterios de evaluación, con la excepción de la discriminación basada en la TRI. Sin embargo, con respecto a este último criterio, se debe tener en cuenta que la cifra presentada en este informe es el resultado de un promedio, por lo que no podría tomarse como un criterio único que hable de la calidad de los ítems de las pruebas.

CAPÍTULO 3.

Cuestionarios auxiliares

Cuestionarios auxiliares

Los cuestionarios auxiliares hacen parte de las pruebas Saber 3°, 5°, 7° y 9°. La aplicación de estos instrumentos contribuye a tener un panorama global de los estudiantes evaluados, puesto que los elementos involucrados en la relación de los estudiantes con sus compañeros, docentes y cuidadores probablemente son determinantes en el proceso de aprendizaje. Estos cuestionarios indagan acerca de las habilidades sociales y emocionales de los estudiantes, así como del relacionamiento de los docentes con sus pares y la percepción al interior de la institución educativa. Además, cuestionan y analizan las características sociales y económicas del hogar de las personas evaluadas. Con este fin, se elaboraron tres cuestionarios auxiliares - dirigidos a los estudiantes -que abarcan temáticas relacionadas con habilidades socioemocionales, factores asociados al aprendizaje y el contexto socioeconómico de los evaluados. En el caso de factores asociados, se incluye un cuestionario para docentes y otro para rectores. La información obtenida constituye un insumo para avanzar en el desarrollo de una evaluación de la calidad educativa integral, ya que incluye componentes complementarios a los tradicionales. A continuación, se describen las generalidades de cada uno de estos cuestionarios.

Características de los cuestionarios auxiliares

Características del cuestionario de habilidades socioemocionales

Las habilidades sociales y emocionales se refieren a las capacidades que tienen las personas para manejar sus pensamientos, emociones y conductas. Además, permiten identificar qué tan bien se adaptan las personas a su entorno, algo fundamental no solo para su bienestar y desarrollo, sino también para la sociedad. Este cuestionario se enfocó en las siguientes dimensiones:

- Autoconciencia emocional: El reconocimiento de las emociones básicas en sí misma y las demás personas
- Autorregulación emocional: El manejo de las emociones propias y la respuesta adecuada a las emociones de los demás
- Automotivación: La confianza en las habilidades propias para alcanzar las metas propuestas

En la [Tabla 11](#), se expone la estructura del cuestionario por dimensión, evidencia y número de tareas del cuestionario de habilidades socioemocionales. Tal como se observa, para grado séptimo, las evidencias se concentran, principalmente, en las dimensiones de autorregulación emocional y automotivación. En estas personas, hay una diferencia en el desarrollo social y emocional con respecto a tercero o quinto, momento en que se consolidan estos procesos.

Este cuestionario se aplicó en dos formas de cuadernillo (A y B) con el fin de pilotear todos los ítems sin agotar a los estudiantes evaluados. Además, estuvo conformado por preguntas de selección dicotómica (sí/no), selección de emociones, selección de acciones y en escala Likert (escala de frecuencia y escala de acuerdo con 4 niveles) con única respuesta. En la [Tabla 12](#), se resume el número de ítems por cada una de las dimensiones por forma para el grado 7°.

Tabla 11. Estructura del cuestionario de habilidades socioemocionales por dimensión, evidencia y número de tareas para el grado séptimo

Dimensión	Evidencia	Número de tareas grado séptimo
1. Autoconciencia emocional	1.1 Reconoce las emociones básicas en sí mismo	1
	1.2 Reconoce las emociones básicas en otras personas	1
2. Autorregulación emocional	2.1 Controla sus emociones	3
	2.2 Toma perspectiva frente a las emociones de los demás	3
	2.3 Empatiza con otros	2
	2.4 Expresa sus sentimientos y emociones	1
3. Automotivación	3.1 Tiene confianza en su capacidad para desarrollar tareas y alcanzar objetivos	4
	3.2 Trabaja en equipo	3

Tabla 12. Número de ítems por dimensión y forma del Cuestionario de Habilidades Socioemocionales

Dimensión	Grado 7°	
	Forma A	Forma B
Autoconciencia emocional	6	6
Autorregulación emocional	31	32
Automotivación	25	24
Total	62	62

Características del cuestionario de Factores asociados

Este cuestionario ayudó identificar algunas variables que pueden influir en el rendimiento escolar de los estudiantes en las pruebas. Dichos factores pueden llegar a incidir en su potencial o dificultar el aprendizaje. En dichos procesos de aprendizaje, participan diferentes actores del sistema educativo que se relacionan entre sí, lo que genera un ambiente escolar en el cual confluyen las actitudes, los comportamientos y las acciones de todos ellos. Debido a que estas variables pueden influir en el desempeño académico de los estudiantes, estos cuestionarios recolectan información que permite consolidar una evaluación más integral. Además, se proponen como un insumo para actividades de investigación relacionadas con el desarrollo de políticas públicas encaminadas a mejorar la calidad educativa nacional. Este cuestionario es aplicado a estudiantes; sus componentes se desglosan de la siguiente manera y la distribución por componente y evidencia se presenta en la Tabla 13:

- **Percepción del colegio:** Incluye las percepciones de la relación entre el estudiante y el docente, así como el gusto por el colegio.
- **Preferencias del aprendizaje:** Evalúa el aprendizaje colaborativo en el aula.
- **Entorno:** Evalúa el contexto de riesgo en la institución educativa.
- **Ambiente en la clase de Lenguaje:** Incluye la evaluación del ambiente en el salón, el estilo del docente y los aspectos democráticos en el aula en la clase de Lenguaje.
- **Ambiente en la clase de Matemáticas** Incluye la evaluación del ambiente en el salón, el estilo del docente y los aspectos democráticos en el aula en la clase de Matemáticas.
- **Involucramiento parental:** Incluye las dinámicas de la relación entre padres, madres y/o cuidadores con el/la estudiante en juego y diálogo.

Tabla 13. Estructura del cuestionario de factores asociados por componente y evidencia

Componente	Evidencia	Número de tareas grado séptimo
1.Percepción del colegio	1.1 Relación estudiante docente	1
	1.2 Gusto por el colegio	1
2.Preferencias del aprendizaje	2.1 Aprendizaje colaborativo	1
3.Entorno	3.1 Contexto de riesgo	1
4.Ambiente clase de Lenguaje	4.1 Ambiente en el salón de lenguaje	1
	4.2 Estilos docentes de lenguaje	1
	4.3 Democracia en el aula de lenguaje	1
5.Ambiente clase de Matemáticas	5.1 Ambiente en el salón de matemáticas	1
	5.2 Estilos docentes de matemáticas	1
	5.3 Democracia en el aula matemáticas	1
6.Involucramiento parental	6.1 Juego	1
	6.2 Diálogo	1

Los ítems se presentan con opciones de respuesta de selección dicotómica directa o inversa (sí/no), o escala Likert (frecuencia y de escala de acuerdo, en 4 niveles) con única respuesta. En la [Tabla 14](#), se resume el número de ítems por componente. Se observa que, en las formas A y B, el número de ítems es simétrico, teniendo en cuenta el enfoque de clase para cada una. El número de ítems que se observa en la [Tabla 16](#) hace referencia a la cantidad de preguntas de cada componente en los cuadernillos de la forma A o B. Las preguntas se relacionan con las percepciones de los estudiantes de su institución educativa, la incidencia y/o utilidad del trabajo colaborativo en el aula, los riesgos asociados con el entorno en el que se desenvuelven los estudiantes, el ambiente en el aula, la democracia en el aula en términos de participación e interlocución entre estudiante-docente, y el grado de involucramiento que tienen los padres y/o cuidadores de cada estudiante frente al juego y el diálogo. En los campos de la tabla donde aparecen las letras NA (No aplica), la forma A corresponde al cuestionario con énfasis en la clase de lenguaje y, la forma B es para la clase de matemáticas.

Tabla 14. Número de ítems por componente del cuestionario de factores asociados

Dimensión	Grado 7°	
	Forma A	Forma B
1. Percepción del colegio	8	8
2. Preferencias del aprendizaje	7	7
3. Entorno	5	5
4. Ambiente clase de Lenguaje	14	NA
5. Ambiente clase de Matemáticas	NA	14
6. Involucramiento parental	8	8

Características del cuestionario socioeconómico

El Cuestionario Socioeconómico (SE) recolecta información sobre el contexto social y económico de los examinandos, lo cual permite caracterizar a la población evaluada. Para ello, los estudiantes de grado séptimo respondieron un cuestionario con un determinado número de preguntas asociadas a cada uno de los cinco componentes, tal y como se muestra en la [Tabla 15](#):

Las preguntas que respondieron los estudiantes se relacionan con el número de horas que leen en casa, dedicación a prácticas deportivas, forma y tiempo de desplazamiento a la institución educativa, posesión en el hogar de ciertos bienes, entre otras. Para ello, los estudiantes seleccionaron una opción de respuesta para cada pregunta; dichas opciones de respuesta podían ser de tipo dicotómico (Sí o No), frecuencia de días, número de años, grado escolar, entre otras.

Tabla 15. Número de preguntas del Cuestionario Socioeconómico para grado séptimo

Dimensión	Grado 7°
Demográfico	3
Prácticas educativas del hogar	3
Desplazamiento a la institución educativa	2
Características socioeconómicas y culturales del hogar	22
Antecedentes escolares	3
Total	33

Metodología de análisis psicométrico para los Cuestionarios Auxiliares

Se realizaron análisis psicométricos de los ítems de cada uno de los cuestionarios auxiliares por grado y forma de cuestionario, con el fin de validar las inferencias realizadas a partir de los datos arrojados. Las fuentes de evidencia de validez presentadas en el informe acerca de la estructura interna de los cuadernillos se presentan a través de los estadísticos univariados relacionados con el porcentaje de distribución de frecuencia de respuestas, análisis de confiabilidad a través del estadístico Alfa de Cronbach y análisis de dimensionalidad.

Para el Cuestionario de Habilidades Socioemocionales y el Cuestionario de Factores Asociados, el análisis de distribución de frecuencias de respuesta se centró en dos criterios, con base en el esquema de valor definido para cada pregunta, los dos criterios verifican que las preguntas no sean muy difíciles y que no existan opciones de respuesta con frecuencias muy bajas. Los resultados que se presentan a continuación se describirán en términos de porcentaje de preguntas que cumplieron o no los criterios. El primer criterio corresponde con la obtención por pregunta de un porcentaje de selección igual o superior al 40% en la opción de respuesta con mayor puntuación definida en el esquema de valor. Esto significa que, por ejemplo, cuando se presenta un 100% en este criterio se hace referencia a que, en la totalidad del grupo de preguntas, la opción de respuesta con mayor valor fue la más seleccionada en cada pregunta. El segundo criterio hace referencia a una opción de respuesta con un porcentaje de selección mayor al 1%, es decir que, cuando se presenta un 100% de cumplimiento en este criterio, significa que, por pregunta, las opciones que no tenían mayor valor en el esquema de valor también fueron atractivas y seleccionadas por los estudiantes.

En términos de utilidad, estos criterios permiten analizar si la respuesta de los estudiantes se comporta de acuerdo con lo esperado desde el diseño de los cuestionarios, donde se desarrolló un esquema de valor que corresponde a la asignación de puntuaciones de las opciones de respuesta para cada ítem. Así mismo, el segundo criterio permite identificar si alguna opción de respuesta se comporta como una alternativa difícilmente descartable para los evaluados.

Para el análisis de distribución de frecuencias de respuesta de los Cuestionarios de Factores Asociados se emplearon los mismos criterios que en el Cuestionario de Habilidades Socioemocionales, con excepción de un aumento en el porcentaje del segundo criterio, puesto que la distribución de la selección de opciones de respuesta fue más variada para las preguntas de estos cuestionarios. Así, el segundo criterio hace referencia a una opción de respuesta con un porcentaje de selección mayor al 2%.

Una consideración adicional para tener en cuenta consiste en la distribución de componentes en el Cuestionario de Factores Asociados pues, como se abordó anteriormente, las formas A y B comparten ítems comunes. En tal caso, se realiza un análisis conjunto para los ítems comunes de las formas, y otro análisis para cada uno de los componentes diferenciados por clases (matemáticas y lenguaje, pues no son comparables los ítems entre sí y requieren que sean analizados de manera separada.

Por otro lado, para establecer el funcionamiento psicométrico de los componentes del Cuestionario Socioeconómico para el grado séptimo y de los ítems asociados a cada uno de estos, se realizó un análisis

de distribución de frecuencias de respuesta para cada ítem, es decir, se revisaron las respuestas que todos los estudiantes marcaron en cada una de las preguntas. Una vez se consolidaron todos estos datos para los estudiantes de grado séptimo, se observó cómo fue la distribución de esas respuestas, en términos porcentuales. El análisis de la distribución de frecuencias de respuesta se enfocó en dos criterios, diferentes a los mencionados anteriormente, y los resultados se presentan en términos del porcentaje de preguntas que cumplieron o no los siguientes criterios:

El primer criterio corresponde a aquellos ítems que obtuvieron una selección de respuesta mayor al 70% en alguna de sus opciones. Por ejemplo, si un componente tiene un porcentaje de cumplimiento del criterio del 100%, esto significa que todos los ítems que hacen parte del componente tienen, en alguna de sus opciones de respuesta, un porcentaje de selección mayor al 70%. Dado que este es un cuestionario que busca caracterizar a la población, este resultado indica que, para el componente, hay una característica que es recurrente en la población.

El segundo criterio hace referencia a aquellos ítems que presentan alguna de sus opciones de respuesta con un porcentaje de selección menor al 2%. Por ejemplo, cuando se presenta un valor del 100% de cumplimiento de este criterio, y dado que este es un cuestionario que busca caracterizar a la población, este resultado indica que, para el componente, hay una característica que es poco recurrente en la población.

Resultados

Cuestionario de habilidades socioemocionales

Este apartado se encuentra dividido en dos secciones: 1) análisis de distribución de frecuencia de las respuestas, y 2) análisis de confiabilidad y dimensionalidad.

Análisis de distribución de frecuencias de respuesta del cuestionario de habilidades socioemocionales

Con base en los criterios anteriormente mencionados, en la [Tabla 16](#), se presenta el resultado del análisis de distribución de frecuencias de respuesta por forma (A y B) del Cuestionario de Habilidades Socioemocionales. En esta tabla, se puede observar el desagregado de la cantidad de ítems totales y el número de ítems por forma que cumple cada criterio, además de la tasa de cumplimiento en relación con el total.

En cuanto al primer criterio, se observó un mayor porcentaje de cumplimiento en la forma A, en la cual el 92% de los ítems obtuvo un porcentaje de selección igual o superior al 40% en la opción de respuesta esperada; es decir, de las 62 preguntas de esta forma, 57 fueron respondidas de acuerdo con lo planteado en el esquema de valor. Con respecto a la forma B, se observó que 8 preguntas, de las 62, es decir el 13% no fueron respondidas de acuerdo con lo esperado. En cuanto al segundo criterio, en la forma A, el 95% de los ítems presentaron opciones de respuesta con un porcentaje de selección mayor a 1%; y, en la forma B, este porcentaje fue de 93,5%.

Tabla 16. Análisis de distribución de frecuencias de respuesta por forma del Cuestionario de Habilidades Socioemocionales – grado séptimo

Forma	Cantidad de ítems	Cumple el criterio	Criterio 1		Criterio 2	
Forma A	62	Sí	57	92%	59	95%
		No	5	8%	3	5%
Forma B	62	Sí	54	87%	58	93,5%
		No	8	13%	5	6,5%

De otra parte, en la [Tabla 17](#), se presenta el análisis de distribución de frecuencias de respuesta por dominios y evidencias del cuestionario, donde se muestra el porcentaje

de preguntas que cumplieron con el primer y segundo criterio, por evidencia. De las ocho evidencias que componen los cuestionarios de séptimo, la forma A cumplió el criterio 1 al 100% en 3 evidencias, mientras que la forma B lo hizo en 4 evidencias. Para las evidencias que no alcanzaron el 100%, se observó que algunos ítems tuvieron porcentajes cercanos al 40% en la respuesta deseada, pero no alcanzaron el umbral de la condición; en este caso, fueron 5 ítems para la forma A y 8 ítems para la forma B que no cumplieron el criterio 1. Respecto al segundo criterio, en esta misma tabla, se observa que 6 evidencias de la forma A cumplieron al 100% la condición de tener opciones de respuesta con porcentajes de selección superiores al 1%, mientras que 4 llegaron al 100% de cumplimiento en la forma B. Es de resaltar que, en ambas formas, la evidencia 3.2 Trabaja en equipo cumplió al 100% ambos criterios.

Tabla 17. Análisis de distribución de frecuencias de respuesta por evidencia del Cuestionario de Habilidades Socioemocionales –grado séptimo

Dominio	Evidencia	Forma A			Forma B		
		Número de ítems	Cumple el criterio 1	Cumple el criterio 2	Número de ítems	Cumple el criterio 1	Cumple el criterio 2
1.Autoconciencia emocional	1.1 Reconoce las emociones básicas en sí mismo.	3	100%	67%	3	100%	67%
	1.2 Reconoce las emociones básicas en otras personas.	3	67%	100%	3	100%	67%
2.Autorregulación emocional	2.1 Controla sus emociones.	12	92%	100%	11	82%	100%
	2.2. Toma perspectiva frente a las emociones de los demás.	9	100%	89%	9	100%	89%
	2.3 Empatiza con otros.	8	87%	100%	8	75%	100%
	2.4 Expresa sus sentimientos y emociones.	2	50%	100%	3	33%	75%
3.Automotivación	3.1 Tiene confianza en su capacidad para desarrollar tareas y alcanzar objetivos.	15	93%	100%	15	87%	100%
	3.2 Trabaja en equipo.	10	100%	100%	10	100%	100%

Análisis de confiabilidad y dimensionalidad

El análisis de confiabilidad se realizó con base en las estimaciones de Alpha de Cronbach, mediante la comparación de las correlaciones entre los ítems que forman parte de cada tarea con el total de la prueba. Si bien para la investigación en general, el valor ideal esperado para estas estimaciones de Alfa (α) de Cronbach es entre 0,80 y 0,90 para indicar que hay consistencia interna, no hay acuerdo en los valores aceptables para el tipo de pruebas denominadas de ejecución típica, porque indagan sobre áreas no cognitivas de la conducta humana (Abal et al., 2010, Taber, 2018). Por lo tanto, para este análisis inicial, se aceptaron valores mínimos de alfa desde 0,60. Para el instrumento en general, se encontró que la forma A presentó un Alfa de Cronbach de 0,75, mientras que la Forma B tuvo un valor de 0,71.

En la [Tabla 18](#), se pueden observar las estimaciones de consistencia interna para ambas formas de prueba. Se evidencia que, para este pilotaje de las preguntas para el grado 7°, las estimaciones de mayor puntuación en consistencia interna y que se consideran dentro del rango aceptable fueron el Índice General de Habilidades socioemocionales y el Índice Automotivación; de igual forma, los subíndices con mayores estimaciones fueron Empatía y Trabajo en equipo. La robustez de los índices de Autoconciencia emocional, Autorregulación emocional, así como del subíndice de Autoeficacia, debe seguir siendo fortalecidos para esta medición con el fin de que permita, a futuro, brindar mediciones más consistentes.

Para identificar las evidencias de validez de la estructura interna de lo que se quiere medir, se realizó un Análisis Factorial Exploratorio (AFE) que permitió indagar acerca de la covariación empírica de las respuestas a las preguntas de los cuestionarios que conforman dimensiones, definidas por la significación común de esas respuestas en la muestra de estudiantes evaluados. En este análisis, se consideraron las cargas mayores a 0,30 en valor absoluto (Pett et al., 2003).

En el análisis de dimensionalidad, para la forma A, se observó que el 93,5% de los ítems obtuvieron cargas factoriales iguales o superiores a 0,3, y solamente 4 preguntas no alcanzaron los valores esperados. En cuanto a la forma B, el 92,3% alcanzó o superó el rango de carga factorial esperada y 6 ítems lograron dicho peso factorial. En ambos casos, para las preguntas cuyo peso no permite evidenciar su agrupación, se requiere una revisión interna para garantizar que la dimensionalidad sea representativa del constructo.

Tabla 18. Análisis confiabilidad por índices y subíndices del Cuestionario de Habilidades Socioemocionales – grado séptimo

Índices/Subíndices	Consistencia interna Alpha de Cronbach	
	Forma A	Forma B
Índice General de Habilidades socioemocionales	0,62	0,57
Índice Autoconciencia emocional	0,25	0,2
Índice Autorregulación emocional	0,45	0,4
Subíndice Empatía	0,61	0,58
Índice Automotivación	0,62	0,58
Subíndice Autoeficacia	0,41	0,36
Subíndice Trabajo en equipo	0,63	0,62

Cuestionario de Factores asociados

A partir de los criterios anteriores, en la [Tabla 19](#), se presenta el resultado del análisis de distribución de frecuencias de respuesta por forma del cuestionario de factores asociados. En esta tabla, se puede observar el desagregado de la cantidad de ítems totales y el número de ítems por forma que cumple cada criterio, además de la tasa de participación en relación con el total.

Para la forma A, se presentó una distribución en el criterio 1 del 70%, lo que quiere decir que obtuvo un porcentaje de selección igual o superior al 40% en la opción de respuesta con mayor puntuación. Así las cosas, de las 42 preguntas de esta forma, 29 fueron respondidas de acuerdo con lo planteado en el esquema de valor. En la forma B, el criterio 1 se cumplió en el 65% de los ítems, es decir que 27 de los 42 ítems se respondieron según lo planteado en el esquema de valor. En cuanto al segundo criterio, en la forma A, el 100% de los ítems presentó alguna opción de respuesta con un porcentaje de selección mayor a 1%; y, en la forma B, también corresponde al 100%.

En la [Tabla 20](#), se presenta el análisis de distribución de frecuencias de respuesta por componentes y evidencias del cuestionario, así como el porcentaje de preguntas que cumplieron con el primer y segundo criterio por evidencia. Se encontró que, de las doce evidencias que componen el cuestionario de factores asociados de séptimo, solo 3 cumplieron al 100% el criterio 1, esto se presentó en ambas formas. Con respecto al segundo criterio, en la [Tabla 20](#), se observa que todas las evidencias en las dos formas cumplieron la condición de tener opciones de respuesta con

porcentajes de selección superiores al 1%. Finalmente, los espacios sin un porcentaje correspondiente se explican porque la forma A es el cuestionario para analizar los componentes en clase de Matemáticas y, la forma B, en clase de Lenguaje.

Tabla 20. Análisis de distribución de frecuencias de respuesta por componentes y evidencias del cuestionario de factores asociados

Componente	Evidencia	Forma A			Forma B		
		Número de ítems	Cumple criterio 1	Cumple criterio 2	Número de ítems	Cumple criterio 1	Cumple criterio 2
1.Percepción del colegio	1.1 Relación estudiante docente	4	100%	100%	4	100%	100%
	1.2 Gusto por el colegio	4	100%	100%	4	100%	100%
2.Preferencias del aprendizaje	2.1 Aprendizaje colaborativo	7	100%	100%	7	100%	100%
3.Entorno	3.1 Contexto de riesgo	5	80%	100%	5	80%	100%
4.Ambiente clase de Lenguaje	4.1 Ambiente en el salón de matemáticas	5	20%	100%			
	4.2 Estilos docentes matemáticas	5	80%	100%			
	4.3 Democracia en el aula matemáticas	4	50%	100%			
5.Ambiente clase de Matemáticas	5.1 Ambiente en el salón lenguaje				5	0%	100%
	5.2 Estilos docentes lenguaje				5	60%	100%
	5.3 Democracia en el aula lenguaje				4	50%	100%
6.Involucramiento parental	6.1 Juego	2	0%	100%	2	0%	100%
	6.2 Diálogo	6	50%	100%	6	50%	100%

Tabla 19. Análisis de distribución de frecuencias de respuesta por forma del cuestionario de factores asociados

Forma	Cantidad de ítems	Cumple el criterio	Criterio 1		Criterio 2	
			Número de ítems	Porcentaje	Número de ítems	Porcentaje
Forma A	42	Sí	29	70%	42	100%
		No	13	30%	0	0%
Forma B	42	Sí	27	65%	42	100%
		No	15	35%	0	0%

Análisis de confiabilidad para el cuestionario de factores asociados

El análisis de confiabilidad se realizó con base en las estimaciones de Alfa de Cronbach y se compararon las correlaciones entre los ítems que forman parte de cada tarea con el total. Se realizó este análisis para las dos formas del cuestionario de estudiantes de grado séptimo, con el objetivo de observar el coeficiente Alpha de Cronbach, el cual mide la consistencia interna alcanzada por evidencia.

En la [Tabla 21](#), se observa que, para ambas formas, el comportamiento del Alpha es similar, siendo involucramiento parental el componente con el valor más alto (0,79), lo que indica que este es el componente que presenta la mejor consistencia interna. No obstante, en la misma tabla, es posible observar que la mayoría de los componentes tienen un valor del Alpha muy cercano al de involucramiento parental, lo cual evidencia una consistencia interna buena en los componentes. El componente entorno es el dominio con el valor más bajo; no obstante, es un valor del Alpha aceptable.

Además, para identificar las evidencias de validez del constructo a medir, se realizó un análisis factorial exploratorio (AFE), el cual permitió indagar acerca de la covariación empírica de las respuestas a las preguntas de los cuestionarios que conforman dimensiones,

definidas por la significación común de esas respuestas en la muestra de estudiantes evaluados. En este ejercicio, se obtuvieron muy buenos resultados, pues los ítems presentaron una agrupación en los factores esperados con cargas entre 0,6 y 0,9.

Cuestionario socioeconómico

De acuerdo con los dos criterios de análisis presentados, se identificó el comportamiento de las respuestas de los estudiantes para cada pregunta del cuestionario. En la [Tabla 22](#), se resume el número total de ítems por cada uno de los componentes y el porcentaje de ítems que cumplieron con cada uno de los criterios establecidos para los evaluados de grado séptimo.

En general, se observa que el criterio 1 no se cumple en la totalidad de los ítems que conforman el componente demográfico, prácticas educativas del hogar y desplazamiento a la institución educativa. Esto significa que los ítems para esos componentes no tienen, en ninguna de sus opciones de respuesta, un porcentaje de selección mayor al 70%. En consecuencia, para estos componentes no hay una característica que sea recurrente en la población. Por otro lado, para los componentes de características socioeconómicas y culturales, y antecedentes escolares, el 18,18% y el 33% de los ítems, respectivamente, cumplen el criterio 1, de modo que estos componentes tienen ítems con alguna opción de respuesta con un porcentaje de selección mayor a 70%. Esto significa que hay una característica recurrente en la población. Estos hallazgos son similares a los encontrados en la aplicación 2021 de la prueba.

Respecto al criterio 2, se encontró que, para los componentes de prácticas educativas del hogar, características socioeconómicas y culturales del hogar, y antecedentes escolares, la mayoría de los ítems no cumplen con el criterio 2, lo cual indica que hay pocos ítems que tengan alguna opción de respuesta con un porcentaje de selección menor al 2%. Por otro lado, para el componente demográfico, se encontró que el 66,66% de los ítems sí cumplen el criterio 2, lo cual indica que hay opciones de respuesta de los ítems que tuvieron un porcentaje de selección menor al 2%, por lo que hay características que son poco recurrentes en la población evaluada. Estos hallazgos son similares a los encontrados en la aplicación 2021 de la prueba.

Tabla 21. Valores del Alpha de Cronbach para los componentes del cuestionario de factores asociados

Componente	Alpha de Cronbach	
	Forma A	Forma B
1. Percepción del colegio	0,78	0,78
2. Preferencias del aprendizaje	0,77	0,77
3. Entorno	0,65	0,65
4. Ambiente clase de Lenguaje	0,78	
5. Ambiente clase de Matemáticas		0,77
6. Involucramiento parental	0,79	0,79

Análisis de confiabilidad del cuestionario socioeconómico

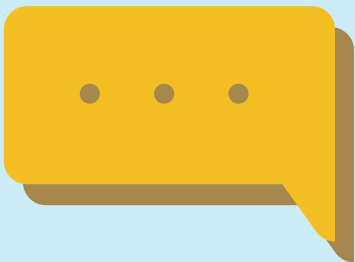
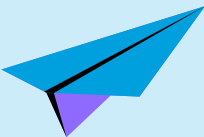
Para revisar la consistencia interna de este cuestionario, se recurrió al Alpha de Cronbach (α), el cual proporciona información sobre qué tan consistente es el instrumento, teniendo en cuenta los ítems o los bloques que lo conforman. Para el caso de grado séptimo, la confiabilidad del instrumento es de 0,56, la cual es considerada como baja; sin embargo, esto podría ocurrir debido a que, como se mencionó anteriormente, este cuestionario indaga por varios elementos sobre el entorno social y económico de los estudiantes, y no mide una habilidad específica. Además, esta medida de consistencia interna del instrumento, aunque no tiene los valores más adecuados, permite entender la estructura del cuestionario. Para finalizar, el análisis de dimensionalidad de este cuestionario no se llevó a cabo, debido a que busca caracterizar a la población en diferentes aspectos sociales y económicos, y no realizar la medición de un constructo, como el caso del Cuestionario de Habilidades Socioemocionales y de Factores Asociados.

Tabla 22. Análisis univariados por componente del Cuestionario socioeconómico para grado séptimo

Componente	Número de ítems	% de ítems que cumplen el criterio 1		% de ítems que cumplen el criterio 2	
		Sí	No	Sí	No
Demográfico	3	0%	100%	66,66%	33,33%
Prácticas educativas del hogar	3	0%	100%	0%	100%
Desplazamiento a la institución educativa	2	0%	100%	50%	50%
Características socioeconómicas y culturales del hogar	22	18,18%	81,81%	13,63%	86,36%
Antecedentes escolares	3	33,33%	66,66%	0%	100%
Total	33	-	-	-	-

Referencias

- ▶ Abal, F. J. P., Lozzia, G. S., Aguerri, M. E., Galibert, M. S., & Attorresi, H. F. (2010). La escasa aplicación de la teoría de respuesta al ítem en tests de ejecución típica. *Revista Colombiana de Psicología*, 19(1), 111-122.
- ▶ Argüello, J. (2019). LERTAP 5 (Paquete de análisis de pruebas del laboratorio de investigaciones educativas): software para el análisis de ítems. *Medir y Evaluar*, 3, 62-71. https://www.humanas.unal.edu.co/lab_psicometria/download_file/view/214/244
- ▶ Baker, F. (2001). *The basics of item response theory*. Office of Educational Research and Improvement. <https://files.eric.ed.gov/fulltext/ED458219.pdf>
- ▶ De Leeuw, E., Hox, J. y Huisman, M. (2003). Prevention and treatment of item non response. *Journal of Official Statistics*, 19(2), 153-176. <https://www.scb.se/contentassets/ca21efb41fee47d293bbee5bf7be7fb3/prevention-and-treatment-of-item-nonresponse.pdf>
- ▶ DeVellis, R. F. (2016). *Scale development: Theory and applications*. Sage.
- ▶ Ebel, R. (1965). *Measuring Educational Achievement*. Prentice-Hall.
- ▶ Instituto Colombiano para la Evaluación de la Educación [Icfes] (2018a). *Guía introductoria al diseño centrado en evidencias*.
- ▶ Instituto Colombiano para la Evaluación de la Educación [Icfes] (2018b). *¿Cómo se generan los puntajes en las pruebas Saber del Icfes?*, Saber al detalle. <https://www.icfes.gov.co/documents/39286/2231027/Boletin+1+-+como+se+generan+los+puntajes+en+las+pruebas+.pdf/c89f8415-140c-d928-77d2-d13b0fe22875?version=1.0&t=1647958798134>
- ▶ Instituto Colombiano para la Evaluación de la Educación [Icfes] (2022a). *Informe de aplicación de las pruebas Saber 3°, 5°, 7° y 9° modalidad lápiz y papel para la población general y electrónica para población con discapacidad*.
- ▶ Instituto Colombiano para la Evaluación de la Educación [Icfes] (2022b). *Piloto muestral y controlado 2022 – Saber 7°*.
- ▶ Jabrayilov, R., Emons, W. H. M., & Sijtsma, K. (2016). Comparison of Classical Test Theory and Item Response Theory in Individual Change Assessment. *Applied Psychological Measurement*, 40(8), 559–572. <https://doi.org/10.1177/0146621616664046>
- ▶ Kline, P. (2000). *The handbook of psychological testing* (2nd ed.). Taylor & Frances/Routledge.
- ▶ Mahias, P. y Polloni, M. (2019). *Desarrollo de instrumentos de evaluación: pruebas*. <https://www.inee.edu.mx/wp-content/uploads/2019/08/P2A354.pdf>
- ▶ Macdonald, P., & Paunonen, S. V. (2002). A Monte Carlo Comparison of Item and Person Statistics Based on Item Response Theory versus Classical Test Theory. *Educational and Psychological Measurement*, 62(6), 921–943. <https://doi.org/10.1177/0013164402238082>
- ▶ Mislavy, R., Steinberg, L. y Almond, R. (2003). *On the Structure of Educational Assessments, CSE Technical Report* (Informe n° CSE-TR-597). Office of Educational Research and Improvement. <https://files.eric.ed.gov/fulltext/ED480557.pdf>
- ▶ Morales, P. (2007). *La fiabilidad de los tests y las escalas*. <https://matcris5.files.wordpress.com/2014/04/fiabilidad-tests-y-escalas-morales-2007.pdf>
- ▶ Muñoz, J. (1997). *Introducción a la teoría de respuesta al ítem*. Ediciones Pirámide.
- ▶ Muñoz, J. (2010). Las teorías de los test: teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo*, 31(1), 57-66. <https://www.papelesdelpsicologo.es/pdf/1796.pdf>
- ▶ Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). Making sense of factor analysis: The use of factor analysis for instrument development in health care research. Sage.
- ▶ Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273-1296.



Queremos conocer tu opinión
sobre este documento, escríbenos a:

analisisydifusion@icfes.gov.co

