



Documento
metodológico:

Análisis de consecuencias de uso Saber 3°, 5° y 9°

Aplicación 2021

[Acceder >](#)



La educación
es de todos

Mineducación

Documento
metodológico:

Análisis de consecuencias de uso Saber 3°, 5° y 9°

Aplicación 2021

Cite este documento así:

APA 7

Instituto Colombiano para la Evaluación de la Educación (Icfes). (2022). Documento metodológico: Análisis de consecuencias de uso Saber 3°, 5° y 9°. Aplicación 2021.

Icontec

INSTITUTO COLOMBIANO PARA LA EVALUACIÓN DE LA EDUCACIÓN. Documento metodológico: Análisis de consecuencias de uso Saber 3°, 5° y 9°. Aplicación 2021. Bogotá: Icfes, 2022.

Presidente de la República
Iván Duque Márquez

Ministra de Educación Nacional
María Victoria Angulo González

Viceministra de Educación
Preescolar, Básica y Media
Constanza Alarcón Párraga

Directora General
Mónica Ospina Londoño

Secretario General
Ciro González Ramírez

Directora de Evaluación
Natalia González Gómez

Director de Producción y Operaciones
Oscar Orlando Ortega Mantilla

Director de Tecnología e información
Sergio Andrés Soler Rosas

Subdirector de Diseño de Instrumentos
Luis Javier Toro Baquero

Subdirectora de Análisis y Divulgación
Mara Brigitte Bravo Osorio

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
María Mercedes Corcho Caro

Jefe Oficina Asesora de
Comunicaciones y Mercadeo
María del Rocío Gutiérrez Araujo

Jefe Oficina Asesora de Gestión de
Proyectos de Investigación
Clara Lorena Trujillo Quintero

Jefe Oficina Asesora de Planeación
Luis Alberto Colorado Aldana

Jefe Oficina de Control Interno
Adriana Bello Cortés

Jefe Oficina Asesora Jurídica
Ana María Cristina de la Cuadra

Asesora Unidad Atención al Ciudadano
Alba Liliana Abril Daza

Elaboración del documento
Claudia Fernanda Castañeda Ávila
Claudia Johanna García Carvajal
Julia Liliana Ríos Herrera
Luisa Fernanda Rincón Caicedo
Jonnathan David Rico Marín
Jackeline Salamanca Durán

Revisión
Mara Brigitte Bravo Osorio
Paul William Cifuentes Velásquez

Diseño, diagramación e ilustración
Carolina García Aponte
Mónica Liliana López

Corrección de estilo
Paul Cifuentes Velásquez

ISBN: 978-958-11-1000-1
Bogotá D.C., junio 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 ©.



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 y derechos de autor a favor del Icfes. 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 material. 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.





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.

En todo caso, cuando se haga uso parcial o total de los contenidos de esta publicación, el usuario deberá 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 y de propiedad exclusiva del Instituto Colombiano para la Evaluación de la Educación (Icfes). Por tanto, cuando su uso pueda causar confusión, los terceros no podrán usar las marcas de propiedad 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.

Tabla de contenidos

	Introducción	6
	1. Marco conceptual: Consecuencias de uso y conceptos relacionados	7
	1.1. Validez de los test	8
	1.2. Consecuencias de uso en las pruebas	9
	1.3. Consecuencias de uso en el contexto de las pruebas Saber 3°, 5° y 9°	11
	2. Propuesta metodológica	12
	2.1 Marco metodológico	13
	2.2 Recolección de datos	14
	2.2.1 Fase 1: Encuesta	14
	2.2.2. Fase 2: Entrevistas	15
	Referencias	16

Índice de figuras

Índice de tablas

Figura 1. Fuentes de validez según los estándares para pruebas educativas y psicológicas.....	8
Figura 2. Modelo de consecuencias de una prueba propuesto por Pan (2009).....	10
Figura 3. Esquema de categorías para el estudio de consecuencias de uso.....	13

Tabla 1. Esquema de categorías para el estudio de consecuencias de uso.....	14
Tabla 2. Distribución de las entrevistas por tipo de participante	15

Introducción

En el marco del contrato interadministrativo CO1. PCCNTR.3081134, suscrito entre el Ministerio de Educación Nacional (MEN) y el Instituto Colombiano para la Evaluación de la Educación (Icfes), se propuso, para las pruebas Saber 3°, 5° y 9° del 2021, la aplicación de una metodología cualitativa que tenga como fin analizar las consecuencias de uso. Esta actividad compromete dos documentos: uno con la metodología a aplicar y otro con los resultados de la metodología.

La evidencia basada en las consecuencias de uso es una de las cinco fuentes de evidencia de validez enunciadas en los estándares para pruebas educativas y psicológicas (American Educational Research Association [AERA], et al., 2014). Así las cosas, la primera parte de este documento expone los conceptos relacionados con la validez de los test y las

consecuencias de uso, planteadas en los estándares. A partir de esto, se busca que el lector comprenda por qué la validez es un tema central en la construcción de instrumentos de medición y la importancia de las consecuencias de uso de los resultados de cualquier instrumento estandarizado. Esta sección cierra con una breve exposición sobre las consecuencias de uso en el ámbito de las pruebas Saber 3°, 5° y 9°.

La segunda parte de este documento desarrolla la propuesta metodológica para la recolección de información que permitirá identificar las potenciales consecuencias de uso de los resultados de las pruebas Saber 3°, 5° y 9°. Además, se exponen las dos fases propuestas (encuesta y entrevistas) y se describen el alcance, los participantes y la forma de análisis para cada uno de los tipos de datos recolectados.



CAPÍTULO 1.

Marco conceptual: Consecuencias de uso y conceptos relacionados

Las pruebas estandarizadas se emplean con frecuencia en el ámbito educativo, principalmente como instrumentos para la evaluación externa de grandes cantidades de estudiantes, por ejemplo, los exámenes Saber del Icfes. Estas evaluaciones a gran escala sirven para recolectar información sobre el desempeño general del sistema educativo, datos que contribuyen en la toma de decisiones de política educativa. Por lo tanto, los principales destinatarios de los resultados son los tomadores de decisiones (Clarkey Luna, 2021). Por otro lado, hay evaluaciones masivas que tienen como propósito la certificación o promoción de estudiantes, lo cual implica consecuencias para sus vidas y en las decisiones de docentes e instituciones educativas sobre el currículo; estas son conocidas como evaluaciones de alto impacto (Clarke y Luna, 2021).

En ambos casos, es importante indagar por los usos e interpretaciones de las pruebas de los actores en el campo educativo. De esta manera, las instituciones que miden la calidad educativa pueden determinar las consecuencias en diferentes niveles del sistema y la percepción de los usuarios de resultados sobre el impacto en la calidad educativa. Con estos datos, se establece si las percepciones y usos están acorde con lo previsto por los constructores de los instrumentos.

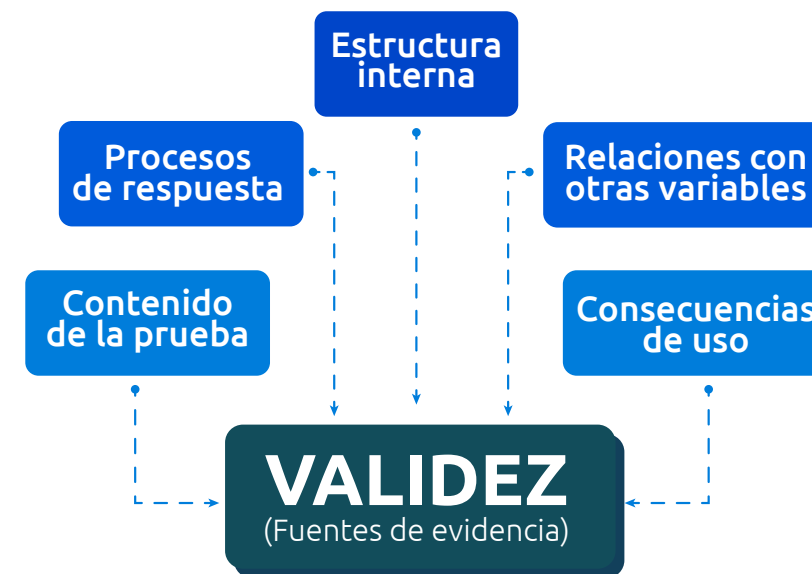
1.1. Validez de los test

En la actualidad, uno de los referentes más importantes, a nivel mundial, para la construcción y desarrollo de instrumentos de medición en educación y psicología son los estándares para pruebas educativas y psicológicas (AERA, et al., 2014). Estos, además de brindar un extenso catálogo de criterios para la evaluación de la calidad de los instrumentos de medición, también aportan un contexto teórico que expone uno de los conceptos más importantes en el ámbito de la medición: la validez, la cual se puede definir como el “grado en el cual evidencia y teoría soportan las interpretaciones de las puntuaciones de la prueba implicadas en los usos propuestos

para la misma.” (AERA et al., 2014, p. 11). Según esta definición, la validez no es una propiedad de los instrumentos, sino una cuestión de grado, lo que implica que no se puede indicar que un instrumento es válido o no, sino en qué grado la interpretación de las puntuaciones es válida en un contexto.

Con el fin de determinar ese grado de validez, los estándares proponen la recolección de evidencia a partir de cinco fuentes (Figura 1).

Figura 1. Fuentes de validez según los estándares para pruebas educativas y psicológicas



Nota: Adaptado de The Standards for Educational and Psychological Testing por American Educational Research Association, American Psychological Association y National Council on Measurement in Education, 2014.

La evidencia basada en el contenido de la prueba se enfoca en demostrar “la adecuación con la que el contenido de la prueba representa el dominio de contenido y la relevancia del dominio de contenido para la interpretación propuesta de los puntajes de la prueba” (AERA et al., 2014, p.15). Por lo tanto, se incluyen las metodologías utilizadas para construir las estructuras de los instrumentos y los ítems, el juicio de expertos y los procedimientos de revisión de contenido.

La evidencia basada en procesos de respuesta implica “análisis teóricos y empíricos de los procesos de respuesta de los examinandos, [los cuales] pueden proporcionar evidencia respecto de la adecuación entre el constructo y la naturaleza detallada del desempeño o respuesta efectivamente empleada por los examinandos” (AERA et al., 2014, p. 16). En la recolección de evidencia de este tipo, generalmente se emplean técnicas de análisis a las respuestas de un individuo, como la entrevista cognitiva o el análisis de la latencia de respuesta.

Por otro lado, con la recolección de evidencia basada en la estructura interna, se busca indicar “el grado en que las relaciones entre ítems de la prueba y componentes de la prueba se ajustan al constructo sobre el que se basan las interpretaciones propuestas de puntajes de la prueba” (AERA et al., 2014, p. 16): En otras palabras, los análisis realizados muestran la relación que hay entre los datos, recolectados con el instrumento, y el marco conceptual. En general, las técnicas usadas para el análisis son de naturaleza estadística, como el análisis factorial, los análisis de funcionamiento diferencial de los ítems o el coeficiente Alpha de Cronbach.

Con respecto a la evidencia basada en relaciones con otras variables, esta surge en el contexto de que “la interpretación prevista para un uso dado implica que el constructo debería relacionarse con algunas otras variables y, como resultado, se analiza la relación de los puntajes de la prueba con variables externas a la prueba, lo que proporciona otra fuente importante de evidencia de validación” (AERA et al., 2014, p. 17). Si las variables externas son otras medidas que evalúan el mismo constructo, se hablaría de evidencia convergente, mientras que, si son medidas que evalúan

constructos diferentes, se hace referencia a evidencia discriminante. También, se pueden utilizar variables diferentes a medidas de otros instrumentos, las cuales se denominan criterios, por ejemplo, en el caso de la evaluación educativa, el desempeño de un estudiante a lo largo de un periodo académico.

Las cuatro fuentes de evidencia descritas están relacionadas con procesos que se podrían ejecutar durante la construcción y pilotaje de los instrumentos. En cambio, las evidencias basadas en las consecuencias de uso están relacionadas con lo que sucede luego de que se aplica el instrumento en algún contexto. Dado que este es el objeto de este documento, a continuación, se trata con mayor amplitud.

1.2. Consecuencias de uso en las pruebas

La inclusión de las consecuencias de uso dentro de las evidencias de validez se debe, en parte, al trabajo de Samuel Messick y su preocupación por las propiedades psicométricas de los test y el impacto a nivel social y humano del uso de sus resultados (Camargo, 2017). La pregunta por el impacto llevó a que este autor desarrollara un cuerpo teórico sobre la validez que influyó en la definición que actualmente se tiene de esta, incluida en los estándares del 2014.

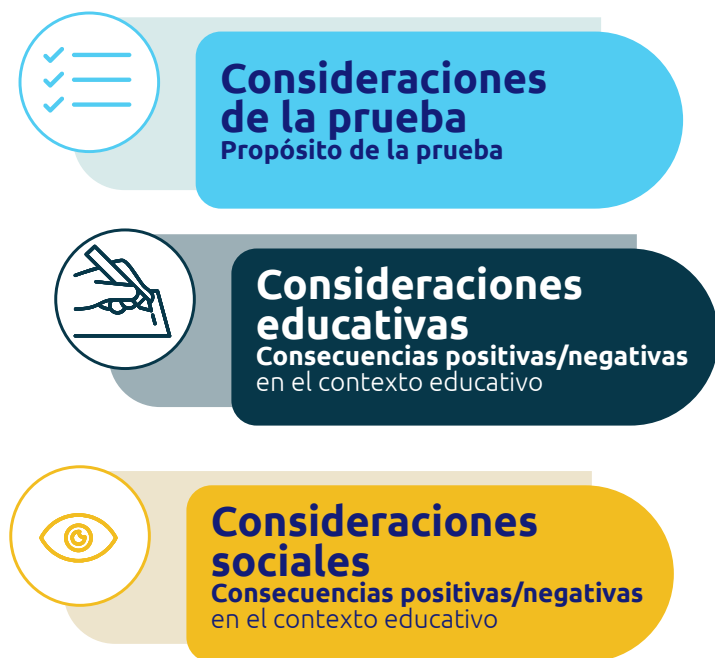
Con la evidencia basada en las consecuencias de uso, se busca “evaluar la solidez de las interpretaciones propuestas para sus usos previstos” (AERA et al., 2014, p. 20), así como los usos que nacen de interpretaciones que trascienden a las de las puntuaciones de las pruebas o para usos con consecuencias imprevistas que, por lo general, son negativas (AERA et al., 2014). En este sentido, Arias y Sireci (2021) afirman lo siguiente:

Las pruebas están diseñadas para tener consecuencias específicas, es decir, están diseñadas para propósitos específicos y para tomar decisiones. Por lo tanto, la evaluación de las consecuencias de una prueba consiste en determinar en parte si los resultados de una prueba satisfacen su propósito y, al mismo tiempo, que no tengan ningún efecto negativo. (p.7)

Dado que se evalúan las interpretaciones que formulan los constructores y las que pueden hacer los usuarios, se ha generado una gran discusión en la comunidad de la medición sobre si estas evidencias deben ser tenidas en cuenta para la validez. Con el fin de profundizar en esta discusión, se puede consultar la tesis de Camargo (2017).

Por otro lado, Pan (2009) ha propuesto un modelo que categoriza tres tipos de consideraciones al momento de analizar las consecuencias de uso de los instrumentos. En la Figura 2, se presentan las tres consideraciones mencionadas.

Figura 2. Modelo de consecuencias de una prueba propuesto por Pan (2009)



Nota: Tomado de Informes de estrategias de validación complementarias por Icfes, 2020, p. 18.

En primer lugar, las consideraciones de la prueba exploran sus objetivos y propósitos de uso. En el caso de evaluar el uso de una prueba, se debe considerar por qué se está utilizando para determinar qué se debe medir en esta. Por ejemplo, las pruebas Saber 3°, 5° y 9° fueron creadas como una herramienta que permite hacer seguimiento de los avances en la calidad de la educación con el fin de brindar información sobre la calidad de la educación primaria y secundaria.

En segundo lugar, las consideraciones educativas y sociales hacen referencia a investigar las consecuencias positivas y negativas a partir de varias fuentes sobre contextos educativos y sociales. En ambos contextos, se deben tener en cuenta los usos de las pruebas en la generación de acciones de mejora para los procesos educativos, así como el conocimiento y apropiación de estas.

Pan (2009) señala que las consideraciones educativas se refieren al potencial, ya sea positivo (normalmente intencionado) o negativo (normalmente no intencionado), en el contexto educativo e involucra a los diferentes actores en la educación. Los estudios han evidenciado diferentes consecuencias para los estudiantes tras la aplicación de las pruebas y el conocimiento de sus resultados, tales como promoción al siguiente grado o consideración de acceso a programas de educación superior (Marchant, 2004). En cuanto a las consecuencias identificadas para los docentes, están limitar su enseñanza del currículo a lo que será evaluado en la prueba y dejar de lado estrategias de enseñanza innovadoras para favorecer el desempeño de los estudiantes en las pruebas (Minarechová, 2012).

Finalmente, las consecuencias sociales, que también pueden ser de carácter positivo o negativo, hacen referencia a los usos de las pruebas que impactan a la sociedad. La evaluación educativa cumple una función social porque permite ver el estado de los aprendizajes, el nivel de preparación de niños, niñas y jóvenes y el grado de escolarización que impacta su continuidad escolar y desempeño social a futuro, así como su calidad de vida. Esta última posibilita el acceso a oportunidades de distintas índoles. Según Chapelle (2020), la función social de las pruebas que afectan las estructuras sociales más amplias y las oportunidades de los individuos son aquellas

que cumplen una “función de movilidad”, es decir, que les proporciona a los estudiantes acceso a las credenciales que necesitan para avanzar en la educación de una forma adecuada según sus habilidades.

1.3. Consecuencias de uso en el contexto de las pruebas Saber 3°, 5° y 9°

Clarke y Luna (2021) describen las evaluaciones a gran escala como aquellas que proveen información sobre el logro de los estudiantes en el sistema educativo para áreas curriculares, edades o grados de un sistema educativo. En general, con esta información, se obtiene un panorama amplio del desempeño estudiantil y se entregan datos a los hacedores de política para la toma de decisiones, aunque también pueden ser consultados por otros actores del sistema. La información se basa en el desempeño de los estudiantes, aunque se puede consultar a docentes, padres, madres y otras personas con el fin de capturar información contextual que permita realizar análisis adicionales.

La anterior descripción se ajusta a las pruebas Saber 3°, 5° y 9°, por lo que se podría decir que estas son consideradas una evaluación a gran escala. Clarke y Luna (2021) también mencionan lo siguiente sobre este tipo de evaluación:

En la mayoría de los casos, los resultados son usados para informar decisiones de bajo impacto sobre políticas docentes y desarrollo profesional, reformas curriculares y asignación equitativa de recursos a las escuelas, sin embargo, en algunos casos, los resultados son usados para decisiones de alto impacto, como imponer sanciones a escuelas que fallan en alcanzar estándares de desempeño o entregar estímulos a aquellas que alcanzan dichos estándares. (p. 3)

Las pruebas Saber 3°, 5° y 9° tienen el objetivo de “servir como línea base de una medición nacional de competencias para estudiantes [...] que le dé información al país acerca del desempeño de estos estudiantes” (Icfes, 2021, p. 15). Sumado a esto, se entrega información de las habilidades socioemocionales y otros factores que pueden impactar en el desarrollo de los aprendizajes (Icfes, 2021, p. 15).

Teniendo en cuenta lo anterior, las pruebas Saber 3°, 5° y 9° son evaluaciones con las que se toman decisiones de bajo impacto, ya que no están destinadas a evaluar la calidad de los establecimientos educativos con el fin de asignar recursos y tampoco influyen en la promoción de estudiantes a grados superiores. Por este motivo, consideramos que el análisis y justificación de las consecuencias de uso ocupan un lugar preponderante, en especial si las pruebas se utilizan con el fin de tomar decisiones críticas para personas e instituciones (Prieto & Delgado, 2010). También, como lo señala Camargo (2017, p. 28, como se cita en Thompson, 1991), “en la medida en que las prácticas de evaluación realizadas por los académicos cambian la realidad social que estudian, el estudio de las consecuencias se convierte en un aspecto esencial de la validez.

Por lo tanto, el objetivo del ejercicio propuesto en este documento es recolectar información para determinar las consecuencias de uso de los resultados de las pruebas Saber 3°, 5° y 9°. De esta manera, se espera identificar la forma en la que los usan colaboradores de las secretarías de educación, docentes y directivos docentes de básica primaria y secundaria del país.

A continuación, se presenta la propuesta metodológica para la recolección de la información que permita cumplir con el objetivo propuesto.



CAPÍTULO 2.

Propuesta metodológica

2.1. Marco metodológico

Con el propósito de proyectar categorías para la construcción de los instrumentos y el análisis de la información, se consideró el marco para la evaluación de las consecuencias de uso propuesto por Pan (2009). En este, se explicita que, desde el análisis de las consecuencias de uso, se pueden recopilar evidencias para apoyar o refutar la conveniencia de la prueba en relación con los efectos colaterales que ésta puede producir.

En atención a las características de las pruebas Saber 3°, 5° y 9°, su alcance y los usuarios de los resultados, se plantea usar dos de las categorías propuestas por Pan (2009): consideraciones de la prueba y consideraciones educativas. Por lo tanto, no se indagará sobre consideraciones sociales, debido a que no se proyecta obtener información de estudiantes, padres, madres o cuidadores. Adicional a esto, las consideraciones sociales suponen evidenciar cambios de mediano y largo plazo, tales como impactos en las oportunidades de ascenso social o acceso a oportunidades educativas, algo que no se logra estudiar en el marco temporal de esta investigación.

En la categoría de consideraciones de la prueba, se indagará por las percepciones sobre el objetivo de la prueba y la confianza en la prueba, es decir, la correspondencia con los objetivos de aprendizaje desarrollados en el proceso educativo y las expectativas de uso de los resultados por parte de docentes y estudiantes. Estas expectativas son consecuencia de la realización de la prueba y el conocimiento de sus resultados.

Por otro lado, se propone subdividir la categoría “consecuencias educativas” en dos subcategorías: en la primera, se incluye la percepción de los docentes sobre los estudiantes. Se aplicaría exclusivamente a docentes y estudiantes que participaron en la prueba, debido a que indaga por efectos de la prueba previos a la aplicación, tales como esfuerzo y motivación para prepararse; aspectos emocionales, como nerviosismo o temor; y expectativas e interés de los estudiantes por conocer los resultados. En

efecto, se parte del supuesto de que los docentes elaboran una percepción de estos temas desde la observación y conocimiento que tienen de los estudiantes.

La segunda subcategoría se refiere a posibles cambios en la dinámica de los equipos de las instituciones educativas o secretarías de educación para hacer uso de los resultados. Por esto, se propone identificar cambios en la organización de los equipos de trabajo para analizar y apropiarse de los resultados y modificaciones en la planificación para generar acciones de mejora. La Figura 3 presenta el esquema de las categorías propuestas para cada una de las consideraciones de la prueba.

Figura 3. Esquema de categorías para el estudio de consecuencias de uso



2.2. Recolección de datos

La recolección de información se hará en dos fases: en un primer momento, se aplicarán encuestas para recolectar información sobre el conocimiento y uso de los informes Saber 3°, 5° y 9°. En una segunda fase, se realizarán entrevistas a profundidad para conocer usos y consecuencias por cada uno de los participantes. A continuación, se describe cada una de las fases.

2.2.1. Fase 1: Encuesta

La encuesta es una técnica de investigación que se vale de procedimientos estandarizados para recolectar datos sobre un conjunto de personas (Casas et al., 2003). En general, es usada para la recolección de grandes cantidades de datos con preguntas cerradas con diferentes formatos de respuesta. Además, puede contener una pequeña cantidad de preguntas abiertas o combinar formatos de pregunta.

En este estudio, el objetivo de la encuesta es indagar por el conocimiento del informe de resultados y algunos usos generales de la información contenida en éste, así como acerca de las percepciones sobre las pruebas Saber 3°, 5° y 9°. Las preguntas formuladas estarán en el marco de las dos categorías de la propuesta de Pan (2009). Se construirán dos formatos de encuesta: uno para directivos docentes y docentes, y otro enfocado en los colaboradores de las secretarías de educación.

La encuesta se enviará a un grupo de establecimientos educativos, en concreto, a los directivos docentes y docentes de grado 3° de las áreas de matemáticas y lenguaje. También, se remitirá a docentes de los grados 5°, 7° y 9° de las áreas de matemáticas, lenguaje, ciencias naturales y ciencias sociales. En el caso de los establecimientos educativos que aplicaron la prueba, la encuesta incluirá preguntas sobre la percepción que los docentes tienen de los estudiantes en relación con la motivación, el esfuerzo, los aspectos emocionales y las expectativas frente a la aplicación de la

prueba. La selección de los establecimientos se hará a partir de las bases de datos de establecimientos que posee el Icfes.

La encuesta para líderes(as) de calidad, evaluación y/o mejoramiento de las secretarías de educación tendrá los mismos objetivos. En este caso, el envío se hará por correo electrónico a los contactos de colaboradores de estas entidades que están en las bases de datos del Icfes.

Las encuestas tendrán preguntas de selección múltiple y, si es necesario, se incluirán algunas preguntas abiertas en un número no mayor a cinco. También, se tendrán algunas preguntas de identificación para poder diferenciar las encuestas según región, cargo y otras variables de interés para el estudio. La Tabla 1 describe el número de encuestas esperado por tipo de población.

Tabla 1. Esquema de categorías para el estudio de consecuencias de uso

Población		Encuestas proyectadas
Instituciones de educación básica primaria y secundaria	Docentes y directivos docentes	96
Secretarías de educación	Líderes de calidad, evaluación y/o mejoramiento de las secretarías de educación	20
Total		116

El análisis se proyecta de dos formas: con respecto a las preguntas de selección múltiple, se harán análisis de frecuencias de las respuestas para cada una de las poblaciones. En el caso de las preguntas abiertas, el análisis se realizará usando “análisis de contenido”, una “técnica que permite investigar el contenido de las comunicaciones mediante la clasificación en categorías de los contenidos o elementos manifiestos de dicha comunicación o mensaje” (Aigeneren, 1999).

Las categorías de clasificación de la información se corresponden con el marco metodológico presentado y se podrán sugerir algunos temas adicionales que brinden mayor claridad al análisis y guarden coherencia con el estudio de las consecuencias de uso.

2.2.2. Fase 2: Entrevistas

La entrevista es una técnica de investigación cualitativa en la que se propone una conversación con un fin específico, más allá del simple hecho de conversar (Díaz et al., 2013). Esta puede variar entre ser muy estructurada, es decir, no se cambia el protocolo de entrevista en su aplicación y contenido, y no estructurada, por lo que no tiene un protocolo definido de antemano y está sujeta a las decisiones que tome el entrevistador.

En el presente estudio, los objetivos de las entrevistas son obtener información sobre los contextos de uso de los resultados de la prueba y ampliar la comprensión de los usos que se han dado a los resultados para identificar saberes, percepciones y prácticas no visibles, así como consecuencias de uso positivas o negativas no esperadas.

Los participantes pertenecerán a las mismas poblaciones a las que se les enviará las encuestas, por lo que se convocará a docentes, directivos docentes y líderes de calidad, evaluación y/o mejoramiento de las secretarías de educación de cada una de las regiones en las que se aplicaron las pruebas Saber 3°, 5° y 9°: Caribe, Pacífico, Llanos, Eje Cafetero y Antioquia, Centro Oriente y Centro Sur.

Los entrevistados serán seleccionados de las bases de datos del Icfes. Una vez elegidos, serán contactados y consultados por su disponibilidad para participar. En caso de aceptar, se acordará fecha y hora de la entrevista. Este instrumento se aplicará con algún programa o plataforma web que permita la conversación fluida y la grabación. Esto último será informado en el contacto inicial y antes de iniciar cada entrevista.

El formato de las entrevistas será semiestructurado, por lo que se crearán protocolos de entrevista que contengan preguntas estándares para todos los entrevistados, pero se dará oportunidad al entrevistador de formular algunas preguntas que permitan ampliar respuestas de interés, en caso de que se requiera. Las preguntas se construirán y analizarán teniendo en cuenta las categorías expuestas en la Figura 3. La Tabla 2 presenta la distribución de entrevistas por tipo de participante.

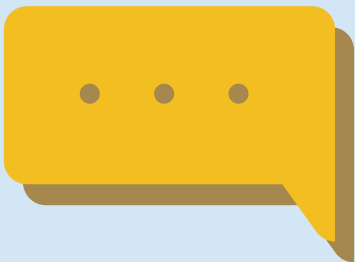
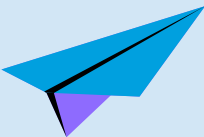
Tabla 2. Distribución de las entrevistas por tipo de participante

Población	Participación en la aplicación de la prueba Saber 3°, 5° y 9° en 2021	Número de entrevistas proyectadas
Docentes	Participante	24
	No participante	3
Directivos docentes	Participante	6
	No participante	3
Profesionales ETC		6
	Total	42

El análisis de las entrevistas se hará a partir del análisis de contenido, tal como se presentó en la sección de encuestas.

Referencias

- Aigeneren M. (1999). Análisis de contenido: una introducción. *La Sociología en sus Escenarios*, (3). <https://revistas.udea.edu.co/index.php/ceo/article/view/1550/1207>
- American Educational Research Association [AERA], American Psychological Association [APA] y National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Arias, A., & Sireci, S. G. (2021). Validez y Validación para Pruebas Educativas y Psicológicas. *Revista Iberoamericana de Psicología*, 14 (1), 11-22. <https://reviberopsicologia.ibero.edu.co/article/view/rip.14102>
- Camargo, S. (2017). *En búsqueda de consenso sobre el concepto de validez: Un estudio Delphi* [Tesis de Maestría, Universidad Nacional de Colombia]. Repositorio UN. <https://repositorio.unal.edu.co/handle/unal/60989>
- Clarke, M. y Luna, D. (2021). *Primer on Large-Scale Assessments of Educational Achievement. National Assessments of Educational Achievement*. World Bank. <https://openknowledge.worldbank.org/handle/10986/35494>
- Casas, J., Repullo, J. y Donado, J. (2003) La encuesta como técnica de investigación. Elaboración de cuestionarios y tratamiento estadístico de los datos (I). *Atención Primaria*, 31(8), 527-538. [https://doi.org/10.1016/S0212-6567\(03\)70728-8](https://doi.org/10.1016/S0212-6567(03)70728-8)
- Chapelle, C. (2020). An introduction to Language Testing's first Virtual Special Issue: Investigating consequences of language test use. *Language Testing*, 37(4), 638-645. <https://doi.org/10.1177/0265532220928533>
- Diaz, L., Torruco, U., Martínez, M. y Varela, M. (2013). La entrevista, recurso flexible y dinámico. *Investigación en Educación Médica*, 2(7), 162-167. <https://www.redalyc.org/pdf/3497/349733228009.pdf>
- Instituto Colombiano para la Evaluación de la Educación [Icfes]. (2021). *Prueba Nacional Muestral y Controlada 2021. Saber 5°*.
- Marchant, G. (2004). What is at stake with high stakes testing? A Discussion of issues and research. *The Ohio Journal of Science*, 104(2), 2-7. https://www.researchgate.net/publication/287435827_What_is_at_Stake_with_High_Stakes_Testing_A_Discussion_of_Issues_and_Research1
- Minarechová, M. (2012). Negative impacts of high-stakes testing. *Journal of Pedagogy*, 3(1), 82-100. <https://sciendo.com/article/10.2478/v10159-012-0004-x>
- Pan, Y. (2009). Evaluación de la conveniencia y de las consecuencias del uso de pruebas. *Colombian Applied Linguistics Journal*, 11, 93-105. http://www.scielo.org.co/scielo.php?script=sci_abstract&pid=S0123-46412009000100007&lng=e&nrm=iso
- Prieto, G. y Delgado, A. (2010). Fiabilidad y Validez. *Papeles del Psicólogo*, 31(1), 67-74. <https://www.papelesdelpsicologo.es/pdf/1797.pdf>



Queremos conocer tu opinión
sobre este documento, escríbenos a:

[analisisydifusion@icfes.gov.co](mailto: analisisydifusion@icfes.gov.co)

