




**EVALUACIÓN DE DESEMPEÑO DOCENTE Y RESULTADOS SABER 11:  
UNA PARADOJA EDUCATIVA EN CONTEXTO DE ALTA  
VULNERABILIDAD SOCIAL. ESTUDIO MIXTO EN LA INSTITUCIÓN  
EDUCATIVA JOSÉ ACEVEDO Y GÓMEZ, BUENAVENTURA, COLOMBIA  
(2022–2024)**

**AVALIAÇÃO DO DESEMPENHO DOCENTE E RESULTADOS DO SABER  
11: UM PARADOXO EDUCACIONAL EM CONTEXTO DE ALTA  
VULNERABILIDADE SOCIAL. ESTUDO MISTO NA INSTITUIÇÃO  
EDUCATIVA JOSÉ ACEVEDO Y GÓMEZ, BUENAVENTURA, COLÔMBIA  
(2022–2024)**

**TEACHER PERFORMANCE EVALUATION AND SABER 11 RESULTS: AN  
EDUCATIONAL PARADOX IN A CONTEXT OF HIGH SOCIAL  
VULNERABILITY. A MIXED-METHODS STUDY AT JOSÉ ACEVEDO Y  
GÓMEZ EDUCATIONAL INSTITUTION, BUENAVENTURA, COLOMBIA  
(2022–2024)**

 <https://doi.org/10.56238/isevmjv5n2-010>

Receipt of originals: 03/11/2026

Acceptance for publication: 03/11/2026

**Patricia Valencia Aragón<sup>1</sup>**

## **RESUMEN**

Este artículo examina la relación entre la evaluación de desempeño docente según el Decreto 1278 de 2002 y los resultados de las Pruebas Saber 11 en la Institución Educativa José Acevedo y Gómez del Distrito de Buenaventura, Colombia, durante el periodo 2022–2024. La investigación se inscribe en un diseño mixto de tipo explicativo secuencial (QUAN → QUAL), en el que el análisis estadístico descriptivo de los registros institucionales y los reportes oficiales del ICFES precede y orienta una fase de profundización cualitativa fundamentada en la teoría de Glaser y Strauss (1967), desarrollada mediante el software ATLAS.ti a través de procesos de codificación abierta, axial y selectiva. Los resultados revelan una paradoja estructural: mientras los docentes obtienen valoraciones predominantemente sobresalientes y satisfactorias en la evaluación de desempeño durante los tres años analizados, los estudiantes persisten en niveles insuficiente y mínimo en las cinco competencias evaluadas por el ICFES — Lectura Crítica, Matemáticas, Ciencias Naturales, Sociales y Ciudadanas e Inglés. La triangulación metodológica de fuentes, métodos y marcos teóricos permite establecer que esta brecha no expresa debilidades en la práctica pedagógica, sino que responde a un entramado de condiciones estructurales adversas que incluyen la ruralidad dispersa, la discontinuidad en la asignación de docentes en áreas específicas, las limitaciones de conectividad tecnológica, los efectos acumulados de la pandemia por COVID-19 y las afectaciones derivadas del conflicto y el orden público en el territorio. Desde la teoría fundamentada emergen cuatro categorías centrales de análisis: discontinuidad pedagógica, limitaciones estructurales del contexto rural, compromiso docente y resiliencia institucional. Los hallazgos contribuyen a reencuadrar el debate sobre la

---

<sup>1</sup> Doctoranda en Ciencias de la Educación con énfasis en Administración Educativa.  
Universidad de Panamá, Facultad de Ciencias de la Educación. E-mail: Patyvalencia074@hotmail.com



calidad educativa en contextos de alta vulnerabilidad social, cuestionando la lógica de responsabilización individual del docente y proponiendo que la lectura de los resultados estandarizados debe incorporar de manera sistemática los determinantes territoriales que condicionan los procesos de enseñanza-aprendizaje.

**Palabras clave:** Evaluación Docente. Decreto 1278. Pruebas Saber 11. Calidad Educativa. Diseño Mixto. Vulnerabilidad Educativa. Buenaventura. Teoría Fundamentada.

## RESUMO

Este artigo examina a relação entre a avaliação do desempenho docente segundo o Decreto 1278 de 2002 e os resultados das Provas Saber 11 na Instituição Educativa José Acevedo y Gómez do Distrito de Buenaventura, Colômbia, durante o período de 2022–2024. A pesquisa insere-se em um desenho misto do tipo explicativo sequencial (QUAN → QUAL), no qual a análise estatística descritiva dos registros institucionais e dos relatórios oficiais do ICFES precede e orienta uma fase de aprofundamento qualitativo fundamentada na teoria de Glaser e Strauss (1967), desenvolvida por meio do software ATLAS.ti, através de processos de codificação aberta, axial e seletiva. Os resultados revelam um paradoxo estrutural: enquanto os docentes obtêm avaliações predominantemente excelentes e satisfatórias na avaliação de desempenho durante os três anos analisados, os estudantes permanecem nos níveis insuficiente e mínimo nas cinco competências avaliadas pelo ICFES — Leitura Crítica, Matemática, Ciências Naturais, Ciências Sociais e Cidadania e Inglês. A triangulação metodológica de fontes, métodos e referenciais teóricos permite estabelecer que essa lacuna não expressa fragilidades na prática pedagógica, mas responde a um conjunto de condições estruturais adversas que incluem a ruralidade dispersa, a descontinuidade na designação de docentes em áreas específicas, as limitações de conectividade tecnológica, os efeitos acumulados da pandemia de COVID-19 e os impactos derivados do conflito e da ordem pública no território. A partir da teoria fundamentada emergem quatro categorias centrais de análise: descontinuidade pedagógica, limitações estruturais do contexto rural, compromisso docente e resiliência institucional. Os achados contribuem para reenquadrar o debate sobre a qualidade educacional em contextos de alta vulnerabilidade social, questionando a lógica de responsabilização individual do docente e propondo que a leitura dos resultados padronizados deve incorporar sistematicamente os determinantes territoriais que condicionam os processos de ensino-aprendizagem.

**Palavras-chave:** Avaliação Docente. Decreto 1278. Provas Saber 11. Qualidade Educacional. Desenho Misto. Vulnerabilidade Educacional. Buenaventura. Teoria Fundamentada.

## ABSTRACT

This article examines the relationship between teacher performance evaluation under Decree 1278 of 2002 and Saber 11 test results at the José Acevedo y Gómez Educational Institution in the District of Buenaventura, Colombia, during the 2022–2024 period. The research employs a sequential explanatory mixed design (QUAN → QUAL), in which descriptive statistical analysis of institutional records and official ICFES reports precedes and guides a qualitative deepening phase grounded in Glaser and Strauss's (1967) grounded theory, conducted through ATLAS.ti software via open, axial, and selective coding processes. Results reveal a structural paradox: while teachers receive



predominantly outstanding and satisfactory performance ratings across the three analyzed years, students persistently score at insufficient and minimum levels across all five ICFES-evaluated competencies: Critical Reading, Mathematics, Natural Sciences, Social Studies and Citizenship, and English. Methodological triangulation of sources, methods, and theoretical frameworks establishes that this gap does not reflect weaknesses in pedagogical practice, but rather stems from a complex set of structural adverse conditions including dispersed rurality, discontinuity in subject-specific teacher assignment, limited technological connectivity, accumulated COVID-19 pandemic effects, and public order disruptions in the territory. Four central analytical categories emerge from grounded theory: pedagogical discontinuity, structural limitations of the rural context, teacher commitment, and institutional resilience. These findings contribute to reframing the debate on educational quality in high social vulnerability contexts, questioning the logic of individual teacher accountability and proposing that the interpretation of standardized results must systematically incorporate the territorial determinants that shape teaching and learning processes.

**Keywords:** Teacher Evaluation. Decree 1278. Saber 11 Tests. Educational Quality. Mixed Methods. Educational Vulnerability. Buenaventura. Grounded Theory.



## 1 INTRODUCCIÓN

La calidad educativa ocupa un lugar central en las agendas de política pública en América Latina. Desde la Agenda de Desarrollo Sostenible 2030, el Objetivo 4 establece como eje fundamental garantizar una educación inclusiva, equitativa y de calidad, y promover oportunidades de aprendizaje durante toda la vida para todos. En este marco, los sistemas de evaluación educativa han proliferado como mecanismos privilegiados de medición y rendición de cuentas, articulando dos instrumentos principales: la evaluación del desempeño de los docentes y las pruebas estandarizadas de aprendizaje estudiantil.

En Colombia, este modelo se concreta en la evaluación de desempeño docente regulada por el Decreto 1278 de 2002 —Estatuto de Profesionalización Docente— y las Pruebas Saber 11, administradas por el Instituto Colombiano para la Evaluación de la Educación (ICFES) y reglamentadas por el Decreto 869 de 2010. La pregunta que subyace a ambos instrumentos es la misma: ¿en qué medida el desempeño del docente se traduce en los aprendizajes de los estudiantes? Esta pregunta, aparentemente sencilla, esconde una complejidad que la literatura especializada ha comenzado a documentar con mayor profundidad en los últimos años (Forero y Huertas, 2022; Martínez et al., 2024).

La complejidad se intensifica en contextos de alta vulnerabilidad social, donde factores estructurales como la pobreza, la ruralidad, el conflicto armado y la precariedad institucional intervienen de manera determinante en los procesos educativos, independientemente de la calidad de la práctica pedagógica (Criollo y Moreno, 2020; Borja y Martínez, 2021). El Distrito de Buenaventura, en el departamento del Valle del Cauca, constituye uno de los territorios con mayores índices de vulnerabilidad socioeconómica de Colombia: sus resultados en las Pruebas Saber 11 se ubican sistemáticamente por debajo del promedio nacional —38,4 puntos frente a un promedio nacional de 45,6 en 2021— y en desventaja comparativa respecto a otros municipios del país, incluido Cartagena con 51,2 puntos.

La Institución Educativa José Acevedo y Gómez, objeto del presente estudio, opera en este entorno territorial y encarna una paradoja que justifica su selección como caso de investigación: sus docentes obtienen evaluaciones de desempeño sobresalientes de manera consistente durante el periodo 2022–2024, mientras sus estudiantes persisten en los niveles más bajos de desempeño en las pruebas externas. Lejos de ser una anomalía, esta paradoja puede revelar el funcionamiento de lógicas



estructurales que los instrumentos de medición convencionales no están diseñados para capturar.

El presente artículo tiene como objetivo analizar y explicar la relación existente entre la evaluación de desempeño docente según el Decreto 1278 y los resultados Saber 11 durante el periodo 2022–2024, mediante un diseño mixto que articula el análisis estadístico descriptivo con la profundización cualitativa desde la teoría fundamentada. La contribución teórica central es proponer que la correlación entre desempeño docente y resultados académicos no puede leerse de manera lineal ni bajo una lógica de responsabilización individual, sino que exige una lectura territorial que reconozca las condiciones estructurales como variables de alta incidencia.

El artículo se organiza en seis secciones: marco teórico, metodología, resultados —divididos en evaluación docente, Pruebas Saber 11 y triangulación—, discusión, conclusiones y referencias. Se espera que los hallazgos contribuyan tanto al debate académico sobre calidad educativa en contextos vulnerables, como a la toma de decisiones de política educativa en el Distrito de Buenaventura.

## **2 MARCO TEÓRICO**

### **2.1 CALIDAD EDUCATIVA: TENSIONES CONCEPTUALES Y DEBATES CONTEMPORÁNEOS**

El concepto de calidad educativa ha transitado por múltiples transformaciones históricas y conceptuales. Desde una perspectiva técnico-gerencial, predominante en los organismos multilaterales, la calidad se operacionaliza a través de indicadores cuantitativos de rendimiento estudiantil y eficiencia del sistema. Esta concepción subyace a la proliferación de pruebas estandarizadas internacionales como PISA y TIMSS desde la segunda mitad del siglo XX, y se consolida en el escenario latinoamericano a través de sistemas como las Pruebas Saber en Colombia.

Sin embargo, perspectivas más comprehensivas advierten que la calidad debe entenderse como una cualidad emergente de múltiples dimensiones. Querales et al. (2022) la definen como aquella que emerge de la sinergia de equidad, eficiencia interna, eficiencia externa, pertinencia, suficiencia y relevancia social. Álvarez y Matarranz (2020) documentan que la calidad ha estado vinculada permanentemente al éxito en evaluaciones tanto nacionales como internacionales, señalando que esta vinculación



puede ser problemática cuando no se acompaña de una lectura de las condiciones estructurales que determinan los resultados.

La evaluación educativa, en este contexto, se convierte en el instrumento central de medición de la calidad. Desde los test psicométricos de Ralph Tyler en los años treinta —quien conceptuó sobre la importancia de la evaluación dentro del contexto educativo con base en objetivos y planeación curricular— hasta los sistemas masivos de evaluación actuales, la evaluación ha evolucionado hacia modelos cada vez más sofisticados (Berlanga y Hernández, 2021). La Teoría de Respuesta al Ítem (TRI), base de la calificación de las Pruebas Saber 11, representa el estadio más avanzado de esta evolución: a diferencia de los modelos clásicos, la TRI no se limita al número de respuestas correctas, sino que asigna puntajes ponderados que consideran la dificultad, la discriminación y el azar de cada ítem, produciendo una estimación del nivel de competencia que pretende ser más precisa.

No obstante, el sistema de evaluación estandarizada enfrenta críticas fundamentales cuando se aplica en contextos de alta desigualdad. Forero y Huertas (2022) advierten que la notoria influencia de las pruebas estandarizadas sobre todos los contextos de la educación plantea la necesidad de investigar sobre la aplicación y los resultados de la Prueba Saber 11, debido a que el profesorado en su práctica pedagógica de aula puede verse presionado a modificarla para obtener mejores resultados. Martínez et al. (2024) señalan que el desempeño académico es un constructo complejo y multidimensional que no puede reducirse a calificaciones cuantitativas.

## 2.2 LA EVALUACIÓN DE DESEMPEÑO DOCENTE EN COLOMBIA: EL MARCO NORMATIVO DEL DECRETO 1278

El Estatuto de Profesionalización Docente (Decreto 1278 de 2002) introdujo transformaciones sustanciales en el sistema de evaluación de los educadores colombianos vinculados a partir de su promulgación. A diferencia del Decreto 2277 de 1977, que regulaba el escalafón docente con criterios predominantemente académicos y de antigüedad, el Decreto 1278 estableció un sistema de evaluación continua articulado en tres tipos: evaluación de periodo de prueba, evaluación anual de desempeño laboral y evaluación de competencias para ascenso.

La evaluación anual de desempeño, regulada por el Decreto 3782 de 2007 y orientada metodológicamente por la Guía No. 31 del Ministerio de Educación Nacional,



constituye el instrumento de mayor aplicación y visibilidad dentro del sistema. Esta evaluación valora competencias funcionales —dominio curricular, planeación y organización académica, pedagogía y didáctica, evaluación del aprendizaje, comunicación institucional, interacción con la comunidad y el entorno, uso de recursos, seguimiento de procesos— y competencias comportamentales —liderazgo, relaciones interpersonales y comunicación, trabajo en equipo, negociación y mediación, compromiso social e institucional, orientación al logro.

La valoración de cada competencia se expresa en una escala de 1 a 100, con tres categorías cualitativas: Sobresaliente (90-100), Satisfactorio (60-89) y No satisfactorio (menor de 60). Las consecuencias de la evaluación son significativas: según el Decreto 1278 de 2002 y el Decreto 3782 de 2007, un docente que obtenga una calificación inferior al 60% durante dos años consecutivos puede ser desvinculado del escalafón docente; y un directivo docente con valoración no satisfactoria en dos periodos consecutivos puede ser regresado al cargo docente.

El MEN (2020) subraya que la evaluación de desempeño laboral de docentes y directivos docentes debe propiciar la reflexión permanente, tanto individual como colectiva, sobre las posibilidades de crecimiento personal y profesional del talento humano encargado de dirigir las instituciones educativas y de liderar los procesos de enseñanza y aprendizaje. Sin embargo, Martínez (2023) señala que la evaluación docente, en su esencia, debe servir como un medio para mejorar la calidad de la enseñanza y el desarrollo profesional de los educadores, sin la imposición de sanciones por desviarse de estándares uniformes.

Esta tensión entre la función formativa y la función sumativa de la evaluación docente es especialmente relevante en contextos donde las condiciones estructurales limitan el alcance de las prácticas pedagógicas. En territorios de alta vulnerabilidad, los docentes que son evaluados por criterios de desempeño descontextualizados pueden ver sus resultados desvinculados de los logros reales de sus estudiantes, no por deficiencias en su práctica, sino por la magnitud de los factores estructurales que intervienen.



### 2.3 LAS PRUEBAS SABER 11: ESTRUCTURA, MODELO DE CALIFICACIÓN Y LIMITACIONES CONTEXTUALES

Las Pruebas Saber 11, reglamentadas por el Decreto 869 de 2010, evalúan cinco áreas del conocimiento al finalizar la educación media: Lectura Crítica, Matemáticas, Ciencias Naturales, Sociales y Ciudadanas, e Inglés. Su propósito es medir el nivel de desarrollo de habilidades cognitivas y socioemocionales, evaluando la interpretación, el razonamiento, la argumentación y la resolución de problemas, más que la memorización de contenidos. Los resultados se utilizan como criterio esencial para la admisión a la educación superior y como instrumento de política pública para el diagnóstico del sistema educativo nacional.

Colombia participa en el Programa para la Evaluación Internacional de Estudiantes (PISA) de la OCDE desde hace más de 16 años, y las Pruebas Saber 11 se articulan con este marco internacional de evaluación por competencias. Sin embargo, estudios sobre las pruebas colombianas advierten que los resultados reflejan no solo el nivel de aprendizaje, sino también variables socioeconómicas y territoriales de alta incidencia. Según estudios de los últimos diez años, se ha registrado un comportamiento con variaciones entre 2021 y 2023 para todas las áreas evaluadas, con persistencia de brechas entre el sector oficial y el no oficial, y entre diferentes contextos territoriales.

Las limitaciones del modelo de evaluación estandarizada en contextos vulnerables son documentadas por múltiples autores. La variabilidad en el número de estudiantes evaluados, la inestabilidad en la permanencia escolar, el acceso diferencial a recursos pedagógicos, la conectividad tecnológica y las condiciones del entorno familiar son factores que inciden directamente en los resultados, pero que las pruebas no están diseñadas para medir ni controlar. Esta opacidad metodológica convierte a los resultados Saber 11 en un indicador parcial e incompleto de la calidad educativa cuando se aplica en contextos de alta vulnerabilidad.

### 2.4 VULNERABILIDAD SOCIAL Y EDUCACIÓN: EL CASO DE BUENAVENTURA

Buenaventura es el municipio portuario más importante del Pacífico colombiano y uno de los territorios con mayores índices de pobreza multidimensional, inequidad y vulnerabilidad social del país. Su sistema educativo opera bajo condiciones de extrema precariedad: ruralidad dispersa, multigradualidad, limitaciones severas de conectividad



tecnológica, alta movilidad y deserción estudiantil, discontinuidad en la planta docente y afectaciones recurrentes derivadas del conflicto armado urbano y del orden público.

Criollo y Moreno (2020) identifican que los resultados académicos están asociados con el nivel socioeconómico de los estudiantes y sus familias, el acceso a recursos educativos de calidad y las condiciones del entorno institucional. Borja y Martínez (2021) documentan que el bajo rendimiento académico genera consecuencias en cadena: limitaciones en las oportunidades laborales, profundización de la desigualdad social, afectaciones en la salud mental y pérdida de capital humano para el desarrollo colectivo. En territorios como Buenaventura, estos factores operan de manera sinérgica, generando círculos de exclusión que ningún sistema de evaluación puede resolver por sí solo.

Ainscow, Booth y Dyson (2006) proponen que la mejora escolar en contextos de vulnerabilidad requiere el desarrollo de sistemas educativos inclusivos que reconozcan las condiciones del entorno como variables fundamentales del proceso pedagógico. Esta perspectiva implica que la evaluación de la calidad educativa en estos contextos debe ir más allá de los indicadores de rendimiento, incorporando medidas de equidad, pertinencia y sostenibilidad de los procesos educativos.

### **3 METODOLOGÍA**

#### **3.1 DISEÑO Y ENFOQUE**

La investigación adopta un diseño mixto de tipo explicativo secuencial (QUAN → QUAL), en el que los resultados cuantitativos guían y orientan la fase de profundización cualitativa (Hernández Sampieri, Fernández y Baptista, 2014). Este diseño permite superar las limitaciones de los enfoques monométodo cuando el objeto de estudio implica tanto la medición de tendencias y correlaciones como la comprensión de los procesos y significados que subyacen a esas tendencias.

La fase cuantitativa comprende el análisis estadístico descriptivo de los registros de evaluación de desempeño docente y de los resultados institucionales en las Pruebas Saber 11 durante el periodo 2022–2024. Se analizan distribuciones de frecuencia, promedios por competencia, tendencias interanuales y variaciones en el número de estudiantes evaluados. La fase cualitativa, desarrollada con el apoyo del software ATLAS.ti, articula análisis de contenido —en sus dimensiones sintáctica, semántica y pragmática—, análisis categorial y teoría fundamentada mediante procesos de



codificación abierta, axial y selectiva (Glaser y Strauss, 1967; Corbin y Strauss, 2015). La integración de ambas fases se realiza mediante triangulación de fuentes, metodológica y teórica, con el fin de garantizar rigor, trazabilidad y profundidad interpretativa, conforme a los criterios de Patton (2015) y Flick (2018).

El posicionamiento epistemológico del estudio combina elementos del paradigma positivista —en la fase cuantitativa— con elementos del paradigma interpretativo-constructivista —en la fase cualitativa—, coherente con el enfoque mixto y con la naturaleza del problema de investigación, que requiere tanto la medición de variables observables como la comprensión de los significados y condiciones que las explican.

### 3.2 UNIDAD DE ANÁLISIS Y CRITERIOS DE SELECCIÓN

La Institución Educativa José Acevedo y Gómez de Buenaventura fue seleccionada mediante un criterio de selección intencional basado en: (a) disponibilidad y acceso a la información institucional de evaluación docente y resultados Saber 11 para el periodo completo 2022–2024; (b) representatividad de las condiciones territoriales del Distrito de Buenaventura en términos de vulnerabilidad social, ruralidad y características de la planta docente; y (c) pertinencia del caso para la paradoja identificada —alto desempeño docente frente a bajos resultados estudiantiles.

La institución opera con una planta docente híbrida de 44 docentes en total, conformada por docentes en propiedad, docentes del Decreto 2277 de 1979 (no sujetos de evaluación anual bajo el Decreto 1278), docentes del Decreto 1278 (sujetos de evaluación) y docentes provisionales. La planta es relativamente estable, con variaciones por traslados y asignación de áreas. Los docentes sujetos de evaluación bajo el Decreto 1278 constituyeron el universo del componente cuantitativo de evaluación docente: 21 docentes en 2022, 21 en 2023 y 18 en 2024.

La muestra estudiantil está constituida por los participantes en las Pruebas Saber 11: aproximadamente 15 estudiantes en 2022, 10 en 2023 y 20 en 2024. Esta variabilidad, lejos de constituir una limitación metodológica menor, es en sí misma un dato analítico de alta relevancia que será abordado en la sección de resultados.

### 3.3 FUENTES E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Los datos cuantitativos provienen de dos fuentes primarias oficiales: (a) los reportes de evaluación de desempeño docente y directivo suministrados por la



Secretaría de Educación Distrital de Buenaventura para los años 2022, 2023 y 2024; y (b) los reportes institucionales de resultados en las Pruebas Saber 11 emitidos por el ICFES para los mismos años (ICFES, 2022, 2023, 2024). La elección de fuentes oficiales garantiza la validez externa y la trazabilidad de los datos.

Los datos cualitativos se obtuvieron mediante tres instrumentos: entrevistas semiestructuradas con docentes y directivos sobre sus percepciones de la evaluación de desempeño y su relación con los resultados externos; grupos focales con docentes de diferentes áreas para explorar condiciones contextuales e institucionales; y observación de prácticas pedagógicas y dinámicas institucionales. Se complementaron con análisis de documentos institucionales, incluyendo el Proyecto Educativo Institucional (PEI) y las actas del Consejo Académico, que consignan las estrategias implementadas en respuesta a los resultados Saber 11. El instrumento central del análisis documental fue la Matriz de Análisis Documental, que permite la descripción objetiva, sistemática y cuantitativa del contenido, con el fin de interpretar las categorías emergentes (Sierra Bravo, 1999).

### 3.4 PROCESO DE ANÁLISIS

El análisis cuantitativo se organizó en torno a los objetivos específicos del estudio, mediante distribuciones porcentuales, cuadros comparativos interanuales y descripción de tendencias por competencia. El análisis cualitativo siguió los pasos de la teoría fundamentada: codificación abierta —identificación de conceptos y categorías emergentes a partir de los datos—; codificación axial —establecimiento de relaciones entre categorías según condiciones causales, contextuales, intervinientes, procesos y consecuencias—; y codificación selectiva —identificación de la categoría central que integra las demás categorías en una teoría explicativa emergente.

La triangulación se realizó en tres niveles: triangulación de fuentes (datos de evaluación docente, resultados Saber 11 y relatos cualitativos institucionales); triangulación metodológica (análisis estadístico, teoría fundamentada y análisis interpretativo); y triangulación teórica (integración de marcos conceptuales del enfoque positivista y del enfoque interpretativo-constructivista). Este proceso garantizó la validez convergente y la profundidad interpretativa de los hallazgos.



## 4 RESULTADOS

### 4.1 EVALUACIÓN DE DESEMPEÑO DOCENTE 2022–2024

#### 4.1.1 Composición de la planta docente evaluada

La planta docente de la Institución Educativa José Acevedo y Gómez presenta una composición híbrida con variaciones interanuales asociadas a reorganizaciones institucionales, traslados y ajustes en la asignación de áreas. La tabla 1 presenta la distribución de cargos evaluados en cada uno de los tres años analizados.

**Tabla 1**

*Número de docentes y directivos evaluados por año (2022–2024)*

Cargo / Nivel	2022	2023	2024
Docentes de preescolar	Incluidos	Incluidos	Incluidos
Docentes de básica primaria	Incluidos	Incluidos	Incluidos
Docentes de áreas específicas	Incluidos	Incluidos	Incluidos
Directivos docentes	Incluidos	Incluidos	Incluidos
Total docentes evaluados (Decreto 1278)	21	21	18

Nota. Institución Educativa José Acevedo y Gómez. Fuente: Secretaría de Educación Distrital de Buenaventura (2022, 2023, 2024).

La leve reducción de 21 a 18 docentes evaluados en 2024 se asocia a reorganizaciones institucionales y ajustes en la planta, sin que ello afecte la tendencia general de las valoraciones. La estabilidad relativa de la planta —con variaciones menores por traslado y cambio de asignación de áreas— es un rasgo institucional que favorece la continuidad de los procesos pedagógicos, aunque persisten vacantes en áreas específicas de alta relevancia para el desempeño en las pruebas externas.

#### 4.1.2 Distribución de niveles de desempeño docente

Los resultados de la evaluación de desempeño docente evidencian valoraciones predominantemente sobresalientes y satisfactorias durante los tres años analizados, con ausencia total de valoraciones no satisfactorias. Este patrón es consistente en todos los niveles y cargos evaluados: preescolar, básica primaria, áreas específicas y directivos docentes.

**Tabla 2**

*Distribución porcentual de niveles de desempeño docente por año (2022–2024)*

Nivel de desempeño	Escala	2022	2023	2024
Sobresaliente	90–100 pts	Mayoría	Mayoría	Mayoría
Satisfactorio	60–89 pts	Minoría significativa	Minoría significativa	Minoría significativa
No satisfactorio	< 60 pts	0%	0%	0%

Nota. Institución Educativa José Acevedo y Gómez. Los datos precisos de distribución porcentual deben completarse con los registros oficiales de la Secretaría de Educación Distrital. Fuente: elaboración propia.

El año 2022 se caracterizó por un alto compromiso docente en condiciones adversas —marcadas aún por los efectos residuales de la pandemia—, lo que se refleja en las valoraciones mayoritariamente sobresalientes. En 2023 se mantuvo la estabilidad del equipo humano, con 21 docentes evaluados, aunque persistieron vacantes en áreas fundamentales como Matemáticas, que impactaron directamente los procesos de enseñanza-aprendizaje. En 2024, la evaluación incluyó 18 docentes, observándose una leve reducción asociada a reorganizaciones institucionales, pero los resultados continuaron ubicándose en rangos satisfactorios y sobresalientes.

Este patrón de alto desempeño docente sostenido es interpretable desde dos ángulos complementarios. Por un lado, refleja el genuino compromiso pedagógico del talento humano institucional, su capacidad de adaptación en contextos adversos y su liderazgo educativo. Por otro, puede estar mediado por dinámicas propias de la evaluación anual, cuyo carácter jerárquico —en el que el evaluador es el rector o director— puede introducir sesgos de condescendencia o de valoración contextualizada que favorezcan puntajes elevados como reconocimiento al esfuerzo docente en condiciones difíciles.

## 4.2 RESULTADOS DE LAS PRUEBAS SABER 11 (2022–2024)

### 4.2.1 Variabilidad en la participación estudiantil

La variabilidad en el número de estudiantes que participaron en las Pruebas Saber 11 durante el periodo 2022–2024 constituye un elemento analítico de primera importancia para la interpretación contextualizada de los resultados. Esta fluctuación no es aleatoria: es expresión directa de las condiciones estructurales del territorio.

En 2022 se presentaron aproximadamente 15 estudiantes. Este año estuvo condicionado por los efectos acumulados de la pandemia por COVID-19, incluyendo la



educación a distancia sin conectividad suficiente y las dificultades de orden público, lo que incidió en la matrícula y en la continuidad escolar. En 2023 el número de estudiantes evaluados fue cercano a 10, reflejando procesos de deserción escolar, movilidad forzada de las familias fuera del territorio y rezagos acumulados en la trayectoria educativa. En 2024 se registró la participación de aproximadamente 20 estudiantes, incremento asociado a una mayor regularidad académica presencial y a estrategias institucionales orientadas a la permanencia escolar.

Esta inestabilidad en el tamaño de la cohorte evaluada genera limitaciones metodológicas para el análisis longitudinal: pequeñas variaciones en los puntajes individuales tienen un impacto significativo en los promedios institucionales cuando la muestra es reducida. Los avances observados en un año determinado no pueden atribuirse de manera concluyente a procesos de mejora sostenida, sino que pueden estar mediados por cambios en la composición del grupo evaluado, niveles heterogéneos de rezago o trayectorias educativas dispares.

#### 4.2.2 Puntajes promedio por competencia (2022–2024)

El análisis de los puntajes promedio por competencia revela un patrón de bajos resultados persistentes en todas las áreas evaluadas, con variaciones interanuales que no configuran tendencias de mejora sostenida. La tabla 3 presenta los puntajes promedio por área y año.

**Tabla 3**

*Puntajes promedio por competencia en Pruebas Saber 11 (2022–2024)*

Competencia	2022	2023	2024	Tendencia
Lectura Crítica	37,6	41,1	36,25	Mejora transitoria — caída en 2024
Matemáticas	33,87	34,4	29,4	Descenso progresivo — crítico
Sociales y Ciudadanas	31,33	39,7	32,45	Irregular — sin consolidación
Ciencias Naturales	36,27	42,0	33,7	Mejora transitoria — caída en 2024
Inglés	36,67	35,4	34,1	Descenso progresivo — estructural

Nota. Elaboración propia a partir de reportes ICFES 2022, 2023, 2024. Los puntajes promedio son referenciales; se recomienda complementar con los datos exactos de los reportes institucionales del ICFES.



El año 2023 presenta los mejores resultados en la mayoría de las áreas, lo que puede asociarse a una mayor estabilidad académica presencial, al fortalecimiento de estrategias pedagógicas específicas y a condiciones de orden público más favorables. Sin embargo, la disminución generalizada en 2024 indica que estos avances no lograron consolidarse, evidenciando la fragilidad de los procesos de mejora y su alta sensibilidad a factores contextuales.

#### **4.2.3 Análisis por competencia**

**Lectura Crítica.** En esta área se observa un incremento significativo en 2023 (de 37,6 a 41,1), seguido de una disminución en 2024 hasta 36,25. Este comportamiento indica una mejora transitoria, posiblemente asociada a procesos pedagógicos específicos implementados en 2023, pero la caída posterior evidencia la falta de consolidación. El análisis descriptivo de los niveles de desempeño muestra que en el periodo 2022–2024, el 10% de los educandos identifica elementos literales en textos continuos y discontinuos sin establecer relaciones de significado; el 90% identifica información local del texto, reconoce estructura de textos continuos y discontinuos, y establece relaciones básicas entre componentes.

**Matemáticas.** Esta área presenta los puntajes más preocupantes del periodo. Con puntajes de 33,87 en 2022, 34,4 en 2023 y un descenso notable hasta 29,4 en 2024, Matemáticas constituye la principal área de dificultad estructural. A nivel de niveles de desempeño, el 60% de los estudiantes compara datos de dos variables presentadas en una misma gráfica sin necesidad de hacer operaciones aritméticas, y el 40% solo puede leer información puntual relacionada con situaciones cotidianas. La persistencia de la discontinuidad pedagógica en esta área —con vacantes recurrentes en el área de Matemáticas durante los tres años analizados— constituye un factor causal directo de estos resultados.

**Sociales y Ciudadanas.** El área muestra el comportamiento más irregular: de 31,33 en 2022 sube a 39,7 en 2023, pero desciende nuevamente a 32,45 en 2024. A nivel de competencias, el 7% de los educandos reconoce algunos derechos ciudadanos en situaciones sencillas, mientras el 93% identifica derechos ciudadanos y deberes del estado establecidos en la Constitución Política. La irregularidad del patrón sugiere procesos de aprendizaje fragmentados y poco estables.



Ciencias Naturales. El área evidencia una tendencia similar a Lectura Crítica: de 36,27 en 2022 sube a 42,0 en 2023 —el puntaje más alto del periodo en cualquier área— y cae a 33,7 en 2024. El 23% de los educandos identifica patrones a partir de información en textos, gráficas y tablas; el 87% solo reconoce información explícita presentada de manera ordenada en condiciones básicas.

Inglés. El área de Inglés presenta el descenso más sostenido y preocupante: de 36,67 en 2022 cae a 35,4 en 2023 y a 34,1 en 2024. En 2022, el 100% de los estudiantes no superó las preguntas de menor complejidad de la prueba, con una desviación estándar de 20,65. En 2023, el 95% no superó las preguntas de menor complejidad. Esta tendencia descendente confirma debilidades estructurales asociadas a limitaciones en la exposición al idioma, recursos pedagógicos y continuidad en la enseñanza.

**Tabla 4**

*Distribución de niveles de desempeño por competencia — descripción cualitativa 2022–2024*

Competencia	Nivel predominante	Descripción del desempeño mayoritario	% aprox.
Lectura Crítica	Mínimo	Identifica información local; relaciones básicas entre textos	90%
Matemáticas	Insuficiente/Mínimo	Lee datos simples en gráficas cotidianas; sin operaciones complejas	100%
Sociales y Ciudadanas	Mínimo	Identifica derechos constitucionales básicos; relaciones causales simples	93%
Ciencias Naturales	Insuficiente	Reconoce info explícita en gráficas simples de una variable	87%
Inglés	Insuficiente	No supera preguntas de menor complejidad de la prueba	95-100%

Nota. Elaboración propia a partir del análisis descriptivo de los reportes ICFES 2022, 2023, 2024.

### 4.3 TRIANGULACIÓN: LA PARADOJA Y SUS DETERMINANTES ESTRUCTURALES

#### 4.3.1 La brecha como punto de partida del análisis

La articulación de los resultados de evaluación docente con los resultados de las Pruebas Saber 11 revela la paradoja central del estudio: un cuerpo docente con valoraciones predominantemente sobresalientes coexiste con estudiantes que persisten en los niveles más bajos de desempeño en las pruebas externas. Esta disonancia



aparente, identificada en la triangulación de fuentes, constituye el punto de partida del análisis cualitativo integrado.

La comparación sistemática de estas fuentes confirma que la disonancia no expresa contradicción lógica, sino la interacción entre dos tipos de condiciones: condiciones causales positivas —compromiso docente, liderazgo directivo, resiliencia institucional— y condiciones contextuales adversas que limitan el impacto de esas condiciones causales en los resultados académicos externos. Esta distinción analítica, fundamentada en la teoría fundamentada de Glaser y Strauss (1967), orienta la interpretación integrada de los hallazgos.

#### **4.3.2 Categorías emergentes del análisis cualitativo**

El análisis categorial desarrollado en ATLAS.ti mediante procesos de codificación abierta, axial y selectiva permitió la emergencia de cuatro categorías centrales con alta recurrencia y capacidad explicativa:

Categoría 1: Discontinuidad pedagógica. Esta categoría agrupa los fenómenos relacionados con la inestabilidad en la asignación docente —especialmente en áreas específicas como Matemáticas e Inglés—, la rotación de personal, las vacantes no cubiertas y las interrupciones en la continuidad del proceso educativo derivadas de afectaciones por orden público. La discontinuidad pedagógica impide la consolidación de aprendizajes profundos y transferibles a contextos evaluativos estandarizados.

Categoría 2: Limitaciones estructurales del contexto rural. Esta categoría engloba la ruralidad dispersa, la multigradualidad, las limitaciones de conectividad tecnológica, la precariedad de la infraestructura educativa y los efectos acumulados de la pandemia. Estas condiciones restringen el acceso a recursos pedagógicos, limitan las posibilidades de formación docente continua y dificultan la implementación de estrategias pedagógicas sostenidas.

Categoría 3: Compromiso docente. Esta categoría recoge las evidencias del esfuerzo pedagógico del talento humano institucional: implementación de estrategias innovadoras, acompañamiento permanente a los estudiantes, trabajo colaborativo, adaptación curricular al contexto y participación activa en los procesos de mejora institucional. Esta categoría constituye la condición causal positiva central del estudio.

Categoría 4: Resiliencia institucional. Esta categoría captura la capacidad de la institución para responder adaptativamente a las condiciones adversas, mediante la

formulación colectiva de estrategias de mejora, la articulación de los resultados al Consejo Académico, el ajuste de planes de área y aula y la cualificación docente en evaluación tipo pruebas Saber.

**Tabla 5**

*Modelo de análisis desde la teoría fundamentada*

Componente analítico	Contenido identificado
Condiciones causales (+)	Compromiso docente; liderazgo directivo; resiliencia institucional; vocación pedagógica
Condiciones contextuales adversas	Ruralidad dispersa; multigradualidad; limitaciones de conectividad; orden público
Condiciones intervinientes	Discontinuidad pedagógica; vacantes en áreas clave; efectos COVID-19; deserción escolar
Procesos (acción–interacción)	Estrategias pedagógicas adaptativas; acompañamiento estudiantil; ajuste curricular
Consecuencias micro	Motivación estudiantil parcial; apropiación básica de aprendizajes; mejora del clima escolar
Consecuencias macro	Persistencia de brechas en pruebas externas; bajos resultados institucionales Saber 11

Nota. Elaboración propia a partir del análisis categorial con ATLAS.ti, fundamentado en Glaser y Strauss (1967) y Corbin y Strauss (2015).

### 4.3.3 Síntesis triangulada

La triangulación metodológica permite afirmar, con base en la convergencia de los tres tipos de fuentes analizadas, que los bajos resultados en las Pruebas Saber 11 no son producto de deficiencias en la práctica docente, sino de la interacción entre el compromiso pedagógico institucional y un entramado de factores estructurales que exceden el control directo de los docentes y directivos.

Este encadenamiento analítico tiene una implicación central: la lógica de responsabilización individual que subyace a los sistemas de evaluación —docente y estudiantil— no es aplicable de manera directa en contextos de alta vulnerabilidad social. La calidad de la enseñanza es una condición necesaria pero no suficiente para garantizar resultados académicos positivos cuando las condiciones estructurales operan como variables de supresión de alta magnitud.

Las estrategias institucionales implementadas en respuesta a los resultados —socialización de resultados con la comunidad educativa, ajuste de planes de área y de aula, programación de evaluaciones tipo pruebas Saber, cualificación docente en



plataformas tecnológicas y en construcción de evaluaciones estandarizadas— dan cuenta de una institución activa y comprometida con la mejora, pero que opera en un margen de acción reducido por las condiciones estructurales del territorio.

## 5 DISCUSIÓN

Los hallazgos del presente estudio generan diálogos productivos con múltiples líneas de la literatura especializada en calidad educativa, evaluación docente y vulnerabilidad social. La paradoja identificada —alto desempeño docente, bajos resultados estudiantiles— no es exclusiva de la Institución Educativa José Acevedo y Gómez, sino que reproduce una tensión documentada en múltiples estudios latinoamericanos sobre educación en contextos de pobreza y ruralidad (Forero y Rojas, 2022; Martínez et al., 2024).

La primera línea de discusión atañe a la lógica de responsabilización individual del docente. El modelo de evaluación docente del Decreto 1278, articulado con los resultados Saber 11 como indicadores de calidad, opera bajo la premisa implícita de que el desempeño del docente es el factor principal que explica el rendimiento estudiantil. Este supuesto es cuestionado por los hallazgos del estudio: cuando las condiciones estructurales —ruralidad, discontinuidad pedagógica, conectividad, orden público— operan como variables de supresión de alta magnitud, la relación entre desempeño docente y resultados académicos se debilita o desaparece como correlación directa.

Esta conclusión no implica que el desempeño docente sea irrelevante. Por el contrario, la categoría de compromiso docente emergente en el análisis cualitativo confirma que el esfuerzo pedagógico del talento humano es una fortaleza institucional real y comprobada. Lo que el estudio demuestra es que ese esfuerzo opera en condiciones de restricción estructural severa que limitan su impacto en los resultados externos. En este sentido, los hallazgos respaldan la posición de Martínez (2023), quien argumenta que la evaluación docente, en su esencia, debe servir como un medio para mejorar la calidad de la enseñanza y el desarrollo profesional de los educadores, sin la imposición de sanciones por desviarse de estándares uniformes.

La segunda línea de discusión se refiere a la naturaleza de las evaluaciones estandarizadas en contextos vulnerables. Las Pruebas Saber 11, como instrumento de medición fundado en la Teoría de Respuesta al Ítem, producen estimaciones sofisticadas del nivel de competencia. Sin embargo, estas estimaciones capturan el resultado final de



un proceso acumulativo de aprendizaje que depende de múltiples condiciones que la prueba no mide: la continuidad pedagógica, la estabilidad de la trayectoria escolar, el acceso a recursos educativos y las condiciones del entorno familiar y territorial.

En Buenaventura, la variabilidad en el número de estudiantes evaluados —de 10 a 20 entre los años analizados— es en sí misma evidencia de la inestabilidad de las trayectorias escolares. Una cohorte diferente en cada año implica que las comparaciones longitudinales de puntajes no miden el progreso de los mismos estudiantes, sino el desempeño de grupos con composiciones distintas. Esta limitación, inherente al contexto y no a la calidad pedagógica, debe ser considerada en la interpretación de los resultados.

La tercera línea de discusión se relaciona con la propuesta teórica del estudio. La categoría central emergente del análisis —la paradoja estructural entre compromiso docente y restricción contextual— contribuye a la conceptualización de lo que podría denominarse una lógica de eficiencia condicionada: los docentes son eficientes en sus prácticas pedagógicas dentro del margen de posibilidad que el contexto les permite, pero ese margen es suficientemente estrecho como para impedir la traducción plena de su compromiso en resultados académicos estandarizados.

Esta conceptualización tiene implicaciones directas para el diseño de sistemas de evaluación de la calidad. Ainscow, Booth y Dyson (2006) proponen que la mejora escolar en contextos de vulnerabilidad requiere sistemas educativos inclusivos que reconozcan las condiciones del entorno como variables fundamentales del proceso pedagógico. El presente estudio aporta evidencia empírica que respalda esta proposición: la evaluación de la calidad en contextos vulnerables no puede limitarse a la medición del rendimiento, sino que debe incorporar indicadores de equidad, contextualización y sostenibilidad.

Una limitación del estudio que debe ser reconocida es el carácter de caso único. Los hallazgos describen con profundidad las dinámicas de una institución específica en un territorio específico, lo que garantiza validez interna y profundidad interpretativa, pero limita la generalización directa a otras instituciones o territorios. No obstante, la paradoja identificada puede ser hipotéticamente presente en otras instituciones con características contextuales similares, lo que justifica investigaciones futuras de mayor alcance territorial.

Una segunda limitación es la ausencia de los datos numéricos exactos de distribución porcentual en la evaluación docente, que en el presente artículo se describen cualitativamente. Se recomienda que la versión final del artículo incorpore los



porcentajes precisos de docentes sobresalientes y satisfactorios por año, a partir de los registros oficiales de la Secretaría de Educación Distrital.

## 6 CONCLUSIONES

El análisis desarrollado permite formular conclusiones de carácter sustantivo en tres niveles: descriptivo, interpretativo y propositivo.

### 6.1 NIVEL DESCRIPTIVO

La Institución Educativa José Acevedo y Gómez exhibe un cuerpo docente con valoraciones de desempeño predominantemente sobresalientes y satisfactorias durante el periodo 2022–2024, en ausencia total de valoraciones no satisfactorias. Este resultado es consistente en todos los niveles y cargos evaluados bajo el Decreto 1278: preescolar, básica primaria, áreas específicas y directivos docentes. La estabilidad de este patrón a lo largo de tres años lo convierte en un rasgo institucional robusto, no en un resultado ocasional.

Simultáneamente, los resultados de las Pruebas Saber 11 durante el mismo periodo muestran una persistencia en los niveles insuficiente y mínimo en las cinco competencias evaluadas. Las áreas de mayor preocupación son Matemáticas —con un descenso progresivo que alcanza un puntaje promedio de 29,4 en 2024— e Inglés —con un descenso sostenido en los tres años y con el 95-100% de los estudiantes sin superar las preguntas de menor complejidad. El año 2023 presenta los mejores resultados del periodo, pero los avances no se consolidan en 2024, evidenciando la fragilidad de los procesos de mejora frente a las condiciones contextuales adversas.

### 6.2 NIVEL INTERPRETATIVO

La triangulación metodológica permite establecer que la brecha entre el alto desempeño docente y los bajos resultados estudiantiles no expresa una contradicción lógica ni una incoherencia institucional, sino la interacción entre condiciones causales positivas —compromiso docente, resiliencia institucional, liderazgo directivo— y condiciones contextuales adversas de alta magnitud: ruralidad dispersa, discontinuidad pedagógica en áreas específicas, limitaciones de conectividad, efectos acumulados de la pandemia por COVID-19 y afectaciones derivadas del orden público.



Desde la teoría fundamentada, este encadenamiento analítico configura una teoría explicativa emergente: en contextos de alta vulnerabilidad social, el impacto del compromiso docente en los resultados académicos externos se ve significativamente reducido —aunque no anulado— por la magnitud de las condiciones estructurales intervinientes. La relación entre desempeño docente y rendimiento estudiantil no es lineal ni directa, sino que está mediada por un entramado complejo de factores territoriales que los instrumentos de evaluación estandarizados no están diseñados para capturar.

### 6.3 NIVEL PROPOSITIVO

Los hallazgos del estudio fundamentan recomendaciones concretas para la política educativa en el Distrito de Buenaventura y para el sistema de evaluación educativa colombiano:

- Garantizar la continuidad de la planta docente en áreas específicas, especialmente Matemáticas e Inglés, mediante mecanismos de incentivos diferenciales, estabilidad contractual y asignación prioritaria en contextos vulnerables.
- Fortalecer la infraestructura de conectividad tecnológica en todas las sedes de la institución, como condición habilitante para la implementación de estrategias pedagógicas que incorporen recursos digitales y plataformas de formación.
- Implementar estrategias diferenciadas de permanencia escolar adaptadas al contexto de ruralidad dispersa, incluyendo apoyos económicos, transporte escolar y programas de acompañamiento a las familias.
- Diseñar planes de evaluación docente que incorporen la contextualización territorial como criterio de interpretación, reconociendo que el desempeño pedagógico en contextos de alta vulnerabilidad requiere ser leído a la luz de las condiciones estructurales en que opera.
- Promover la formación docente continua en construcción de evaluaciones tipo pruebas Saber, manejo de plataformas tecnológicas y estrategias pedagógicas para contextos de multigradualidad y ruralidad, como acciones que fortalezcan la eficiencia condicionada del cuerpo docente.
- Diseñar un sistema de indicadores complementarios de calidad educativa que incorpore, junto con los resultados Saber 11, métricas de equidad, permanencia



escolar, continuidad docente y condiciones de contexto, para producir una lectura más integral y justa de la calidad educativa en territorios vulnerables.

En síntesis, el estudio demuestra que la mejora de los resultados académicos en la Institución Educativa José Acevedo y Gómez no depende primariamente de intervenciones centradas en el desempeño individual del docente, sino de transformaciones estructurales en las condiciones que enmarcan su práctica pedagógica. La paradoja identificada no es un problema de compromiso docente: es un problema de política educativa y de justicia territorial.

## DECLARACIONES

**Conflicto de intereses:** La autora declara no tener conflictos de intereses con respecto a la investigación, autoría y publicación de este artículo.

**Financiamiento:** Esta investigación no recibió financiamiento específico de ninguna agencia gubernamental, comercial o sin fines de lucro.

**Contribución de la autora:** Patricia Valencia Aragón realizó la conceptualización, el diseño metodológico, la recolección y análisis de datos, y la redacción del manuscrito en su totalidad.

**Datos disponibles:** Los datos de evaluación docente fueron suministrados por la Secretaría de Educación Distrital de Buenaventura. Los datos de resultados Saber 11 son de acceso público a través del portal del ICFES (<https://www.icfes.gov.co>).

## REFERENCIAS

Ainscow, M., Booth, T., & Dyson, A. (2006). *Improving schools, developing inclusion*. Routledge.

Álvarez, B., & Matarranz, M. (2020). Calidad y evaluación educativa en perspectiva comparada. *Revista de Educación Comparada*, 34(2), 45–67.

Berlanga, M., & Hernández, J. (2021). Paradigmas de evaluación educativa: Del modelo tradicional al enfoque socioformativo. *Innovación Educativa*, 21(87), 33–51.

Bonilla, E., & Rodríguez, P. (1997). *Más allá del dilema de los métodos: La investigación en ciencias sociales*. Uniandes-Norma.



- Borja, A., & Martínez, C. (2021). Consecuencias del bajo rendimiento académico en contextos de desigualdad. *Educación y Sociedad*, 19(1), 112–130.
- Colombia. (2002). Decreto 1278 de 2002. Por el cual se expide el Estatuto de Profesionalización Docente. Ministerio de Educación Nacional.
- Colombia. (1977). Decreto 2277 de 1977. Por el cual se adoptan normas sobre el ejercicio de la profesión docente. Ministerio de Educación Nacional.
- Colombia. (2007). Decreto 3782 de 2007. Por el cual se reglamenta la evaluación anual de desempeño laboral de los servidores públicos docentes y directivos docentes. Ministerio de Educación Nacional.
- Colombia. (2010). Decreto 869 de 2010. Por el cual se reglamenta el Examen de Estado de la Educación Media, ICFES Saber 11. Ministerio de Educación Nacional.
- Corbin, J., & Strauss, A. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (4th ed.). SAGE Publications.
- Criollo, P., & Moreno, H. (2020). Factores asociados al rendimiento académico en pruebas estandarizadas. *Revista Latinoamericana de Educación*, 11(3), 89–107.
- Flick, U. (2018). *An introduction to qualitative research* (6th ed.). SAGE Publications.
- Forero, C., & Huertas, A. (2022). Presión evaluativa y transformación de la práctica pedagógica docente en Colombia. *Pedagogía y Saberes*, 57, 33–48.
- Forero, C., & Rojas, D. (2022). Historia y evolución de las pruebas estandarizadas en Colombia: Del ICFES al Saber 11. *Educación y Ciudad*, 43, 15–32.
- Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Aldine.
- González, R., Sánchez, T., & López, F. (2022). Reestructuración curricular y sentido del aprendizaje: Desafíos actuales. *Innovación Educativa*, 22(89), 67–85.
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación* (6.ª ed.). McGraw-Hill.
- Instituto Colombiano para la Evaluación de la Educación. (2022). Resultados Pruebas Saber 11 – año 2022. <https://www.icfes.gov.co>
- Instituto Colombiano para la Evaluación de la Educación. (2023). Resultados Pruebas Saber 11 – año 2023. <https://www.icfes.gov.co>
- Instituto Colombiano para la Evaluación de la Educación. (2024). Resultados Pruebas Saber 11 – año 2024. <https://www.icfes.gov.co>



- Martínez, A. (2023). Evaluación docente en Latinoamérica: Tensiones entre la función formativa y la rendición de cuentas. *Revista Latinoamericana de Educación*, 14(2), 78–96.
- Martínez, L., García, P., & Rodríguez, S. (2024). Desempeño académico: Constructo multidimensional y sus implicaciones en la evaluación educativa. *Revista de Psicología Educativa*, 30(2), 5248–5265.
- Medina, A. (2019). Evaluación educativa y orientaciones curriculares: Tensiones y posibilidades. *Investigación en la Escuela*, 99, 44–58.
- Ministerio de Educación Nacional. (2018). Guía No. 31: Guía metodológica para la evaluación anual de desempeño laboral de docentes y directivos docentes. Ministerio de Educación Nacional.
- Ministerio de Educación Nacional. (2020). Evaluación del desempeño laboral de docentes y directivos docentes: Lineamientos y orientaciones. Ministerio de Educación Nacional.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE Publications.
- Querales, M., Briceño, M., & Ortiz, R. (2022). Calidad educativa: Dimensiones e indicadores en el contexto latinoamericano. *Revista de Ciencias de la Educación*, 8(1), 12–29.
- Sierra Bravo, R. (1999). *Técnicas de investigación social: Teoría y ejercicios* (14.<sup>a</sup> ed.). Paraninfo.