



**DISEÑO DE UN INSTRUMENTO DE EVALUACIÓN PARA
EL DIAGNÓSTICO DE LAS COMPETENCIAS
LINGÜÍSTICAS EN INGLÉS EN AMBIENTES DE
APRENDIZAJE DE BÁSICA PRIMARIA**

Myrian Yolanda Camacho Rodríguez,

Aida Vanessa González Pineda y

Carmen Natalia Guzmán Hernández

Facultad de Educación, Universidad Santo Tomás

Maestría en Ambientes Bilingües de Aprendizaje

Directora

Jenny Catalina Loaiza Fuquen

myrian.rodriquezcam@usantotomas.edu.co

<https://orcid.org/0009-0005-4570-7960>

aida.pinedagon@usantotomas.edu.co

<https://orcid.org/0009-0006-6649-0331>

carmen.hernandezguz@usantotomas.edu.co

<https://orcid.org/0009-0008-9056-8906>

We have no known conflict of interest to disclose.

ABSTRACT

Resumen

El presente estudio aborda la ausencia de instrumentos diagnósticos contextualizados para identificar la competencia lingüística en inglés de estudiantes de básica primaria en instituciones educativas públicas colombianas. A pesar de las políticas de bilingüismo promovidas por el Ministerio de Educación Nacional (MEN), persiste la falta de mecanismos de seguimiento que permitan evaluar su impacto real. Asimismo, pruebas estandarizadas como las Saber (ICFES) no incluyen la evaluación del inglés en este nivel educativo.

El objetivo de la investigación fue diseñar un instrumento de evaluación contextualizado, accesible y pertinente, alineado con los Derechos Básicos de Aprendizaje (DBA), los Estándares Básicos de Competencias (EBC) y los descriptores del nivel A1 del Marco Común Europeo de Referencia para las Lenguas (MCER). Metodológicamente, se adoptó un enfoque cualitativo descriptivo mediante análisis documental, diseño del instrumento, prueba piloto y validación por juicio de expertos.

Los resultados evidencian que la herramienta posee validez cultural, claridad funcional y pertinencia respecto al nivel A1. Además, se identificaron oportunidades de mejora relacionadas con la simplificación del léxico, el fortalecimiento de la habilidad de escritura y la incorporación de un mayor componente cultural colombiano. También se observó que los estudiantes obtienen mejores desempeños en tareas de reconocimiento que en actividades de uso contextual del idioma, así como diferencias significativas entre contextos rurales y urbanos.

Se concluye que el instrumento constituye una herramienta válida y confiable para el diagnóstico formativo en básica primaria, favoreciendo procesos de evaluación más equitativos y contextualizados. Asimismo, su implementación puede contribuir a reducir brechas educativas, fortalecer la labor docente y promover el bilingüismo en diversos contextos educativos colombianos.

Palabras clave: evaluación diagnóstica, competencia comunicativa en inglés, educación básica primaria, educación pública colombiana, enseñanza contextualizada

Abstract

The present study addresses the lack of contextualized diagnostic instruments to identify the English language proficiency of primary school students in Colombian public educational institutions. Despite the bilingualism policies promoted by the Ministry of National Education (MEN), there is a persistent lack of monitoring mechanisms to evaluate their real impact. Furthermore, standardized tests, such as the Saber (ICFES) exams, do not include English assessment at this educational level.

The objective of the research was to design a contextualized, accessible, and pertinent assessment instrument, aligned with the Basic Learning Rights (DBA), the Basic Competence Standards (EBC), and the A1 level descriptors of the Common European Framework of Reference (CEFR). Methodologically, a descriptive qualitative approach was adopted through documentary analysis, instrument design, pilot testing, and expert judgment validation.

The results show that the tool possesses cultural validity, functional clarity, and relevance regarding the A1 level. Furthermore, areas for improvement were identified related to the simplification of vocabulary, the strengthening of writing skills, and the incorporation of a greater Colombian cultural component. It was also observed that students perform better in recognition tasks than in activities requiring the contextual use of the language, as well as significant differences between rural and urban contexts.

It is concluded that the instrument constitutes a valid and reliable tool for formative diagnosis in primary education, favoring more equitable and contextualized assessment processes. Likewise, its implementation can contribute to reducing educational gaps, strengthening teaching practices, and promoting bilingualism in diverse Colombian educational contexts.

Keywords: diagnostic assessment, English communicative competence, primary education, Colombian public education, contextualized teaching



TABLA DE CONTENIDO

ABSTRACT	2
INTRODUCCIÓN.....	8
Pregunta problematizadora:	12
OBJETIVOS.....	13
Objetivo general.....	13
Objetivos específicos	13
MARCO CONCEPTUAL	14
Conceptualización.....	21
METODOLOGÍA.....	27
Enfoque y diseño de investigación	28
Contexto y participantes	29
Instrumentos.....	31
Método de recolección de datos.....	37
ANÁLISIS DE RESULTADOS.....	42
Análisis descriptivo de la prueba piloto.....	43
CONCLUSIONES.....	59
CONTRIBUCIONES A LOS AMBIENTES BILINGÜES DE APRENDIZAJE.....	66
RECOMENDACIONES	80
APÉNDICES	83
Apéndice 1	83
Apéndice 2	83



Apéndice 3	83
Apéndice 4	83
Apéndice 5	83
ANEXOS	84
Anexo 1	84
Anexo 2	84
Anexo 3	84
Anexo 4	84
REFERENCIAS	85



INTRODUCCIÓN

El presente estudio se enmarca en el eje de Gestión Educativa de la Maestría en Ambientes Bilingües de Aprendizaje, entendido como un proceso de planeación, organización y liderazgo orientado a garantizar la calidad y pertinencia de los programas académicos. La gestión educativa constituye un proceso estratégico orientado al fortalecimiento de las dinámicas institucionales, pedagógicas y administrativas en función del mejoramiento de la calidad educativa. En este sentido, Álvarez Huertas (2024) señala que la gerencia educativa mantiene una relación directa con la implementación de políticas pedagógicas y con el rendimiento académico, debido a que favorece la organización institucional, la toma de decisiones y el desarrollo de ambientes de aprendizaje más efectivos.

Desde esta perspectiva, la gestión educativa trasciende las funciones exclusivamente administrativas y se orienta hacia la construcción de procesos participativos, reflexivos y contextualizados que contribuyan al fortalecimiento de las prácticas pedagógicas. Lo anterior resulta especialmente pertinente en programas como la Maestría en Ambientes Bilingües de Aprendizaje, donde se busca potenciar competencias profesionales y sociales en diversos contextos educativos. Asimismo, Álvarez Huertas (2024) destaca que una adecuada articulación entre la gestión institucional y las políticas pedagógicas favorece procesos de innovación, calidad educativa y transformación de las prácticas de enseñanza.

De manera que, uno de los principales desafíos en la enseñanza del inglés como lengua extranjera es la falta de información precisa y actualizada sobre el nivel de proficiencia de los estudiantes de básica primaria, especialmente de grado quinto, en instituciones educativas públicas de países latinoamericanos como Colombia. Esta situación

dificulta identificar el nivel de desempeño alcanzado al finalizar este ciclo educativo y realizar una transición adecuada hacia la básica secundaria. Aunque el MEN ha implementado políticas públicas orientadas al fortalecimiento del proceso de enseñanza-aprendizaje del inglés en el país, la evaluación de la eficacia de dichas iniciativas continúa siendo limitada debido a la ausencia de informes consistentes sobre su implementación y resultados.

Investigaciones recientes sobre políticas lingüísticas y bilingüismo en Colombia han evidenciado que, aunque estas iniciativas buscan fortalecer la calidad educativa y la competitividad global, aún persisten dificultades relacionadas con su aplicación y seguimiento en los contextos locales. En este sentido, Morales Llano (2022) señala que las políticas lingüísticas no solo influyen en el uso y adquisición de una lengua, sino que también responden a perspectivas ideológicas y socioculturales que impactan directamente las prácticas educativas. Asimismo, el autor destaca que en Colombia predominan enfoques críticos y socioculturales que evidencian tensiones entre los lineamientos oficiales de bilingüismo y las realidades presentes en las aulas. De manera complementaria, diversos estudios han señalado que la ausencia de instrumentos diagnósticos contextualizados y de mecanismos sistemáticos de seguimiento dificulta valorar el impacto real de dichas políticas en los diferentes contextos educativos del país.

Una de las herramientas utilizadas para reportar el nivel de conocimiento de los estudiantes es la prueba Saber 3°, 5°, 7° y 9°, la cual es una evaluación estandarizada de carácter externo, aplicada periódicamente por el ICFES a los estudiantes de educación básica primaria y secundaria. Esta evaluación busca identificar el desarrollo de las competencias en áreas fundamentales y obligatorias. Específicamente, esta prueba mide las competencias de

los estudiantes en las áreas de Matemáticas, Competencias Comunicativas en Lenguaje: Lectura y Escritura, Competencias Ciudadanas: Pensamiento Ciudadano, Acciones y Actitudes, Ciencias Naturales y Educación Ambiental (ICFES, 2023). Sin embargo, no existen registros oficiales que den cuenta de los resultados y desempeño en el área de inglés de los estudiantes de básica primaria.

En contraste con algunas instituciones privadas, que suelen contratar pruebas internacionales para medir el nivel de inglés de sus estudiantes, son muy pocas las instituciones públicas que cuentan con los recursos necesarios para destinar parte de su presupuesto a este tipo de evaluaciones. Esta limitación impide conocer de manera precisa el nivel de competencia lingüística de los estudiantes y dificulta la toma de decisiones pedagógicas informadas, la implementación de estrategias pedagógicas tempranas, adaptadas a las fortalezas y debilidades de los alumnos.

Este es el caso de las instituciones educativas públicas Adolfo María Jiménez del sector rural, ubicada en Sotaquirá, Boyacá y la Lucrecio Jaramillo Vélez – Sede Agrupación Colombia, ubicada en Medellín, Antioquia. Los estudiantes de grado quinto inician su formación en lengua extranjera, inglés, desde primer grado y transición, respectivamente. Al finalizar quinto grado, concluyen su etapa de básica primaria e ingresan al grado sexto, correspondiente a la secundaria. Sin embargo, durante este tránsito entre ciclos no existe una herramienta institucional que permita determinar con claridad el nivel de competencia lingüística con el que los estudiantes culminan la básica primaria. Como señala el Consejo de Europa (s.f.), la proficiencia lingüística debe entenderse como la capacidad de usar el lenguaje en contextos diversos y específicos, lo que implica contar con instrumentos de

evaluación adecuados que permitan medir y garantizar un desarrollo progresivo en las diferentes etapas escolares.

Por lo tanto, es fundamental considerar propuestas de instrumentos de evaluación flexibles y adaptables que permitan subsanar esta brecha. Este tipo de herramientas minimiza la subjetividad en la evaluación y contribuye con el seguimiento del progreso de los estudiantes a largo plazo, con el fin de fortalecer el proceso de enseñanza-aprendizaje del inglés en el ciclo de básica primaria. Según Fulcher (2015), los instrumentos de evaluación de lengua deben diseñarse para ser válidos, confiables y sensibles al contexto educativo, de manera que no solo midan resultados estáticos, sino que también sirvan como apoyo al aprendizaje continuo y así ajustar eficazmente los planes de estudio, tomar decisiones pedagógicas contextualizadas, y al mismo tiempo, reducir la brecha entre lo público y lo privado, democratizando el acceso a una evaluación de calidad para todos los estudiantes de las instituciones públicas del país (Ordoñez & Rodríguez, 2017).

En consecuencia, es necesario plantear un instrumento contextualizado, asequible y pertinente para el sector público; asimismo, que esté alineado con las particularidades del contexto colombiano, las características cognitivas y lingüísticas de los estudiantes de básica primaria, teniendo en cuenta la intensidad horaria de la lengua extranjera y el nivel esperado de inglés; es decir, que se articule con los DBA, que proporcionan lineamientos claros sobre lo que los estudiantes deben aprender por grado (MEN, 2016), los EBC, que buscan orientar la enseñanza del inglés con base en niveles internacionales como el MCER (2020) y los referentes nacionales como los propuestos por el ICFES en las pruebas Saber (ICFES, 2023).

Pregunta problematizadora:

¿Cuáles son los elementos necesarios de un instrumento de evaluación contextualizado, asequible y pertinente para diagnosticar las competencias lingüísticas en inglés en ambientes de aprendizaje de básica primaria?



OBJETIVOS

Objetivo general

Diseñar un instrumento de evaluación contextualizado, asequible y pertinente que permita el diagnóstico de las competencias lingüísticas en inglés en ambientes de aprendizaje de básica primaria.

Objetivos específicos

1. Analizar referentes teóricos, pedagógicos y normativos nacionales e internacionales que orientan la evaluación de competencias en inglés en ambientes de aprendizaje de básica primaria.
2. Identificar los elementos para la construcción del instrumento de evaluación contextualizado, asequible y pertinente que contribuya al diagnóstico de las competencias lingüísticas en inglés en ambientes de aprendizaje de básica primaria.
3. Analizar la pertinencia, claridad y coherencia del instrumento con el contexto escolar a partir de la aplicación de una prueba piloto.
4. Validar la coherencia y adecuación del instrumento mediante la valoración cualitativa de un experto, garantizando su correspondencia con los estándares nacionales e internacionales.

MARCO CONCEPTUAL

La enseñanza del inglés en las instituciones educativas colombianas ha estado marcada por una serie de desafíos que, a pesar de los esfuerzos por mejorar el aprendizaje, siguen siendo evidentes en diversos niveles. En este contexto, diversos estudios han analizado desde diferentes perspectivas la metodología de enseñanza y los instrumentos de evaluación utilizados en el país.

El primer estudio, titulado Análisis del Ava Accessit y su enfoque metodológico para la enseñanza del inglés en nivel B1 (Gaitán *et al.*, 2023), destaca la importancia de adaptar las metodologías de enseñanza y evaluación a las necesidades específicas de los estudiantes, con énfasis en el contexto colombiano. A su vez, el segundo estudio, Políticas educativas estandarizadas en la enseñanza del inglés en básica primaria (Tique-Estrada, 2025), realiza un análisis de las políticas públicas y su implementación en el ámbito educativo, identificando limitaciones en la formación docente y en la aplicación de los estándares nacionales.

Por otro lado, Desde el diseño de evaluaciones y hacia la literacidad en evaluación de lenguas (Giraldo *et al.*, 2023), propone una reflexión crítica sobre la evaluación, abogando por una comprensión contextualizada y equitativa de los procesos evaluativos. A su vez, el estudio de Ordoñez y Rodríguez (2017), Análisis de las estrategias e instrumentos de evaluación en el aprendizaje del inglés como lengua extranjera, diferencia entre evaluación formativa y calificación, y subraya la importancia de la evaluación como un proceso integral y adaptado a los ritmos de aprendizaje de los estudiantes.

Finalmente, el estudio de Fulcher (2012), *Assessment Literacy for the Language Classroom*, amplía el concepto de “alfabetización en evaluación”, sugiriendo que los docentes deben ser capaces no solo de diseñar y aplicar pruebas, sino también de reflexionar sobre su impacto en el aprendizaje.

Todos estos estudios coinciden en la necesidad de un enfoque más contextualizado y reflexivo en la evaluación del aprendizaje del inglés, especialmente en el contexto colombiano, donde existen brechas significativas en la infraestructura educativa, la capacitación docente y la implementación de políticas públicas.

Por consiguiente, el presente trabajo se basa en estas discusiones previas y propone un instrumento diagnóstico diseñado específicamente para evaluar el nivel de inglés de los estudiantes de grado quinto en las instituciones educativas públicas de Colombia. Este instrumento busca no solo ofrecer datos objetivos sobre el dominio del idioma, sino también suscitar prácticas evaluativas que respondan a las realidades sociales y educativas del país, fomentando la literacidad en evaluación y mejorando el proceso de enseñanza-aprendizaje en el aula.

En este sentido, el primer estudio destaca la importancia de la comunicación efectiva en la segunda lengua y la evaluación formativa para fortalecer las habilidades lingüísticas y colaborativas. Los resultados principales fueron la importancia de adaptar las metodologías de enseñanza y evaluación a las necesidades específicas de los estudiantes, lo cual es crucial para el diseño de instrumentos efectivos en el contexto colombiano y se conecta con nuestro estudio dado que coincide en la necesidad de herramientas de evaluación contextualizadas y adaptadas al entorno educativo colombiano.

El segundo estudio examina cómo se han implementado las políticas estandarizadas de enseñanza del inglés en la educación básica primaria en Colombia. Igualmente, se analizaron 15 estudios a través del protocolo PRISMA y se identificaron limitaciones en la formación docente, deficiencias en la aplicación de los EBC y una falta de contextualización de las políticas nacionales en los entornos locales. Los resultados principales fueron identificar que los instrumentos de evaluación estén alineados no solo con los estándares nacionales, sino también con las realidades locales, como lo evidencia esta revisión. Reforzar esta alineación mejorará la eficacia de cualquier política de enseñanza del inglés en Colombia. Se diferencia con nuestro estudio porque el artículo aborda un enfoque macro (análisis de políticas públicas), mientras que la propuesta se centra en un instrumento específico de evaluación contextualizado.

El tercer estudio explora la importancia de integrar la literacidad en evaluación dentro del diseño de evaluaciones de lenguas. Los autores argumentan que, más allá de la elaboración técnica de pruebas, es fundamental que los docentes desarrollen una comprensión crítica y contextualizada de los procesos evaluativos. El texto destaca cómo la literacidad en evaluación permite a los educadores reflexionar sobre las prácticas evaluativas, considerando factores como la equidad, la validez y el impacto en el aprendizaje de los estudiantes. Se enfatiza la necesidad de formar a los docentes no solo en la creación de instrumentos de evaluación, sino también en la interpretación y uso de los resultados para mejorar la enseñanza y el aprendizaje. Los resultados principales plantean una mirada crítica frente al proceso de evaluación al introducir el concepto de literacidad en evaluación.

Desde este punto de vista, se fomenta la importancia de ir más allá del diseño técnico de pruebas, para comprender la evaluación como una práctica integral. Se conecta con

nuestro estudio porque se plantea que la evaluación debe responder a contextos reales, culturales y sociales, especialmente en contextos educativos diversos como el colombiano. Además, propone que los docentes desarrollen literacidad en evaluación, es decir, una comprensión crítica y reflexiva sobre cómo evaluar, con qué fines y en qué condiciones. En último término, argumenta que la evaluación no debe ser solo un proceso de medición, sino una herramienta formativa que oriente y mejore el aprendizaje. Hace referencia a la necesidad de alinear la evaluación con políticas educativas, estándares y necesidades locales.

El cuarto estudio tiene como finalidad analizar los instrumentos y estrategias de evaluación que utilizaron las maestras en el curso de inglés como lengua extranjera en la primera infancia, a través de una diferenciación entre lo que es la evaluación y la calificación. Lo anterior permite ver cómo cada una de ellas contribuye a un aprendizaje perdurable (evaluación formativa), que trasciende en el aprendizaje de los niños, dando sentido al conocimiento y a lo que reconocen. La calificación busca la memorización de un tema, con el fin de obtener una nota; examina, mide y certifica lo que está bien o está mal; ello puede llevar a que el estudiante pierda el interés por el aprendizaje; por esta razón surgió la preocupación por saber si los maestros califican o evalúan, teniendo en cuenta que la evaluación formativa es aquella que tiene en consideración los ritmos de aprendizaje de los niños, y permite utilizar diferentes instrumentos ajenos a la calificación que propician la enseñanza de la lengua extranjera.

Los resultados principales fueron que la evaluación del aprendizaje del inglés en básica primaria se reconoce como un proceso que debería ser formativo e integral; sin embargo, en la práctica se evidencian diferencias según la institución. Mientras algunas logran integrar la evaluación como un proceso continuo que favorece la reflexión y el

aprendizaje significativo a través de diversos instrumentos, en otras predomina un enfoque cuantitativo centrado en la calificación y la memorización, lo cual limita la motivación y el verdadero desarrollo de competencias comunicativas en los estudiantes. se conecta con nuestro estudio ya que ambos coinciden en que la evaluación del inglés debe ser un proceso formativo y contextualizado, no reducido a la simple calificación. Reconocen que los instrumentos de evaluación deben responder a las necesidades reales de los estudiantes y al contexto educativo, de manera que favorezcan un aprendizaje significativo y el desarrollo progresivo de competencias comunicativas.

Sin embargo, se diferencian porque el cuarto estudio analiza las prácticas de evaluación en la primera infancia, comparando cómo diferentes instituciones aplican instrumentos cualitativos y cuantitativos; tu propuesta, en cambio, no se limita a describir prácticas existentes, sino que diseña un instrumento diagnóstico específico para grado quinto en instituciones públicas, con un enfoque aplicado y orientado a generar datos objetivos sobre el nivel de inglés de los estudiantes al finalizar la básica primaria.

El quinto estudio muestra como las pruebas de idiomas han experimentado una expansión sin precedentes durante la primera parte del siglo XXI. Como resultado, existe una necesidad cada vez mayor de que los profesionales de la evaluación lingüística consideren con mayor precisión lo que significa «alfabetización en evaluación» y articulen su papel en la creación de nuevos materiales y programas pedagógicos en la evaluación y el examen lingüístico para satisfacer las necesidades cambiantes de los profesores y otras partes interesadas en una nueva era.

Este artículo describe un proyecto de investigación en el que se desarrolló, puso a prueba y distribuyó en Internet un instrumento de encuesta para determinar las necesidades de formación en evaluación de los profesores de idiomas. Los resultados se utilizaron para diseñar nuevos materiales didácticos y seguir desarrollando recursos en línea que pudieran utilizarse para apoyar la impartición de programas. El proyecto aporta dos contribuciones significativas; en primer lugar, proporciona nuevos contenidos derivados empíricamente para el concepto de alfabetización en evaluación, en cuyo marco se enmarca el desarrollo de materiales y la enseñanza.

En segundo lugar, ha puesto de manifiesto problemas metodológicos con las técnicas de encuesta existentes que pueden haber influido en estudios anteriores, y se sugieren soluciones a estos problemas. Los resultados principales fueron descubrir que muchos profesores de idiomas reconocen la importancia de ser competentes en evaluación (*assessment literacy*), pero tienen claras necesidades de formación, especialmente en aspectos conceptuales y estadísticos del diseño de pruebas.

En este sentido, Fulcher (2012) señala que los docentes requieren fortalecer su *assessment literacy* mediante una comprensión crítica de la evaluación, que incluya el diseño de instrumentos válidos y confiables, la interpretación de resultados, la diferenciación entre evaluación formativa y pruebas estandarizadas, y la consideración de dimensiones éticas, sociales y contextuales. En consonancia con estos planteamientos, el presente estudio propone un instrumento diagnóstico contextualizado para estudiantes de grado quinto en instituciones públicas colombianas, orientado no solo a identificar el nivel de inglés de los estudiantes, sino también a apoyar la práctica docente y promover procesos evaluativos pertinentes al contexto educativo nacional.

Los referentes analizados muestran la preocupación inicial de este proyecto basado en la poca información sobre la proficiencia del inglés en las instituciones educativas públicas del país, especialmente en el análisis de los estudiantes al terminar su ciclo de básica primaria en grado quinto. Estas brechas se clasifican en tres frentes que son el programa de bilingüismo nacional estandarizado, que no tiene en cuenta las desigualdades de educación frente a la infraestructura, la capacitación docente y los intereses de los estudiantes; se siguen manteniendo las mismas prácticas pedagógicas en su mayoría sin tener una actualización didáctica y que cada día los estudiantes tienen acceso a mayor diversidad lingüística gracias al constante contacto con la tecnología.

Como lo plantea Tique-Estrada (2025), el Programa Nacional de Bilingüismo está basado en estándares internacionales lo cual permite hacer un análisis crítico de las brechas mencionadas anteriormente. Las políticas al ser públicas deben ir de la mano de la realidad del país y no a los estándares de países del primer mundo con modelos educativos muy diferentes a los nuestros. Dichas políticas condicionan el aprendizaje significativo y la forma de aprender de los estudiantes.

En efecto, el diagnóstico enfocado en las necesidades reales de la comunidad es un instrumento clave para apoyar la literacidad en la evaluación, desarrollada por Giraldo *et al.* (2023), donde dicho instrumento no busca tener buenos o malos resultados sino indagar en el uso real del idioma.

Conceptualización

Instrumentos para la evaluación de los aprendizajes

En el presente estudio se definen los instrumentos de evaluación como herramientas para el seguimiento, control y regulación de los conocimientos, habilidades y destrezas del estudiante. De acuerdo con la Universidad Nacional de Educación a Distancia (UNED), son recursos de apoyo en la labor educativa que permiten monitorear y regular el aprendizaje a lo largo de una estrategia o técnica de evaluación. Se interpretan los instrumentos como elementos técnicos esenciales dentro del proceso educativo, pues complementan las técnicas o estrategias de evaluación: no son equivalentes a estas, sino que las ponen en funcionamiento y hacen posible su aplicación práctica.

En consonancia con Torres-Lara *et al.* (2021), los instrumentos de evaluación deben responder a principios técnicos de validez, confiabilidad, pertinencia, factibilidad, responsabilidad y rigor para garantizar que los resultados reflejen con precisión el desempeño real de los estudiantes. La validez implica que el instrumento mida efectivamente lo que pretende medir; la confiabilidad, que arroje resultados consistentes y estables en diferentes aplicaciones; la pertinencia, que se relacione directamente con los objetivos de aprendizaje y el contexto educativo; y la factibilidad, que sea viable de aplicar considerando los recursos y condiciones del entorno escolar. Asimismo, la responsabilidad ética del evaluador se manifiesta en la selección y aplicación justa, inclusiva y transparente de los instrumentos, evitando sesgos o interpretaciones arbitrarias.

Estos instrumentos deben guardar una relación directa con los objetivos del curso, las instrucciones dadas a los estudiantes y las técnicas usadas, evitando diseñarse de forma aislada. En consecuencia, se conciben como medios que favorecen la retroalimentación

continua, promueven la objetividad y la autorregulación del estudiante, y fortalecen la transparencia y equidad en los procesos evaluativos. Los instrumentos de evaluación son, por tanto, el puente entre la teoría y la práctica: transforman los productos académicos en evidencias medibles y verificables, lo cual aporta validez y legitimidad al proceso pedagógico.

Elementos de los instrumentos

El diseño de los instrumentos de evaluación exige comprender sus componentes esenciales. Palomares *et al.* (2023) destacan que su validez, confiabilidad, pertinencia y factibilidad son condiciones indispensables para garantizar que los resultados reflejen de forma precisa el nivel de desempeño de los estudiantes. Además, se añade el principio de rigurosidad técnica, que implica que los ítems estén redactados de forma clara, coherente y alineada con los objetivos de aprendizaje y los estándares curriculares.

En el marco de este trabajo, se procurará que el instrumento diagnóstico sea válido respecto a los EBC y DBA; confiable en su aplicación en contextos educativos públicos; pertinente al nivel cognitivo y lingüístico de quinto grado; y ético y responsable en cuanto a su implementación y uso de resultados. Se busca, así, una herramienta que no solo mida, sino que oriente y promueva la mejora continua del proceso educativo.

Instrumentos de evaluación en lengua extranjera

En el ámbito de la enseñanza de lenguas extranjeras, los instrumentos de evaluación constituyen medios fundamentales para evidenciar las competencias comunicativas que desarrollan los estudiantes. Según Díaz *et al.* (2021), estos instrumentos incluyen pruebas

escritas, orales, rúbricas y observaciones, y deben trascender el simple proceso de calificación para valorar el uso funcional del idioma.

De acuerdo con Torres-Lara *et al.* (2021), la aplicación de instrumentos válidos y confiables en la enseñanza de idiomas garantiza que la evaluación se base en criterios observables y medibles, y no en percepciones subjetivas del docente. Por tanto, el diseño debe considerar la validez de contenido (alineación con los componentes del idioma y las habilidades comunicativas), la validez de constructo (coherencia con la competencia lingüística que se desea evaluar), y la confiabilidad de la evaluación, que se fortalece mediante rúbricas y guías claras de observación.

En este proyecto, la comprensión de los instrumentos de evaluación en lengua extranjera permitirá orientar la construcción de un dispositivo que mida las habilidades de lectura y escritura en inglés de manera significativa, evitando que la evaluación se limite a la reproducción mecánica de estructuras gramaticales. Además, se procurará que los resultados sean interpretables, consistentes y transferibles, de manera que sirvan como evidencia útil para la mejora del proceso de enseñanza-aprendizaje.

Pruebas diagnósticas

Las pruebas diagnósticas se entienden como evaluaciones diseñadas para identificar las fortalezas y debilidades de los estudiantes, con el propósito de orientar la enseñanza. Alderson *et al.* (2015) explica que estas pruebas, a diferencia de las estandarizadas, buscan información detallada sobre un grupo concreto y proporcionan retroalimentación inmediata para la toma de decisiones pedagógicas.

De acuerdo con Torres-Lara *et al.* (2021), la validez y confiabilidad en este tipo de pruebas se sustentan en la correspondencia entre los objetivos de aprendizaje, los contenidos evaluados y los criterios de logro establecidos. Un instrumento diagnóstico será válido si sus ítems representan adecuadamente las habilidades que se desea observar, y confiable si los resultados son consistentes y replicables ante condiciones similares. Además, la responsabilidad profesional del docente implica analizar los resultados con sentido formativo, evitando el uso punitivo o comparativo de los datos obtenidos.

Por tanto, el presente proyecto asume que el instrumento a diseñar debe cumplir un papel eminentemente diagnóstico, formativo y no sancionador, de manera que los resultados sirvan como punto de partida para ajustar las prácticas de aula, personalizar la enseñanza y promover el aprendizaje autónomo.

Pruebas diagnósticas en lengua extranjera

En la enseñanza del inglés como lengua extranjera, las pruebas diagnósticas adquieren un valor adicional. Leontjev *et al.* (2023) subrayan que este tipo de evaluaciones permiten detectar con precisión los obstáculos específicos que enfrentan los aprendices, ya sea en vocabulario, comprensión lectora, gramática o producción escrita.

Desde una perspectiva técnica, se busca garantizar su validez de criterio (relación con estándares internacionales como el MCER y su confiabilidad interna, mediante la revisión y pilotaje de los ítems antes de su aplicación. Asimismo, la responsabilidad ética y pedagógica demanda que las pruebas diagnósticas respeten la diversidad lingüística y cultural del estudiantado, evitando sesgos que puedan afectar la equidad.

En este sentido, el instrumento propuesto no solo clasificará a los estudiantes en niveles generales de desempeño, sino que también permitirá identificar aspectos lingüísticos puntuales que requieren fortalecimiento. De este modo, se favorecerá una evaluación válida, confiable, justa y contextualizada, que aporte información útil para la planificación docente y la mejora del aprendizaje del inglés en contextos escolares diversos.

Ambientes de aprendizaje

El concepto de ambientes de aprendizaje también resulta fundamental para este estudio. Para Ye (2024), los ambientes de aprendizaje son el conjunto de condiciones físicas, sociales y pedagógicas que influyen en el aprendizaje, ya sean presenciales o mediados por tecnologías emergentes. Un ambiente estimulante promueve la motivación y la interacción significativa con el conocimiento. Asimismo, otras investigaciones han mostrado el potencial de los entornos virtuales y de la realidad aumentada en el aprendizaje de lenguas (Schorr *et al.* 2024).

En relación con el proyecto, esto implica considerar que el instrumento debe ser aplicable, factible y confiable tanto en aulas rurales con recursos limitados como en entornos digitales más avanzados. La validez del instrumento, en este caso, dependerá también de su capacidad de adaptación contextual y de su pertinencia cultural y tecnológica.

Ambientes de aprendizaje en básica primaria

Los ambientes de aprendizaje en básica primaria requieren un abordaje particular. Como señala León Auris (2024), en esta etapa el aprendizaje debe articularse con experiencias lúdicas y significativas, acordes al desarrollo cognitivo de los niños entre 6 y 12

años. En coherencia con ello, el instrumento diagnóstico propuesto tendrá en cuenta la claridad del lenguaje, la adecuación del nivel de dificultad, la equidad en la presentación de las tareas y la coherencia con los intereses y contextos cotidianos de los estudiantes de quinto grado.

De esta manera, la validez del instrumento no se limitará a su precisión técnica, sino que incluirá su pertinencia pedagógica y social, consolidando un proceso evaluativo confiable, justo y significativo que promueva la mejora del aprendizaje y el desarrollo integral del estudiante.



METODOLOGÍA

El presente estudio se desarrolla bajo un enfoque cualitativo, dado que busca comprender, analizar e interpretar los significados que subyacen en los referentes teóricos, pedagógicos y normativos relacionados con la evaluación de las competencias lingüísticas en inglés en la educación en básica primaria, así como diseñar y validar un instrumento pertinente para este fin.

De acuerdo con Hernández Sampieri *et al.* (2022), la investigación cualitativa se orienta a comprender los significados de las experiencias y procesos educativos en sus contextos naturales. Por tanto, este estudio no pretende generalizar resultados, sino interpretar y aplicar conocimientos que contribuyan al mejoramiento de la evaluación del inglés en básica primaria mediante la construcción de un instrumento diagnóstico contextualizado.

El enfoque cualitativo se caracteriza por su naturaleza exploratoria, descriptiva y aplicada, al permitir examinar datos no numéricos, como documentos, lineamientos y opiniones expertas, para identificar categorías, significados y relaciones entre los elementos analizados. Para la fase documental se adopta el análisis de contenido como técnica principal, entendido como un proceso sistemático de interpretación de comunicaciones y textos en sus contextos sociales, orientado a construir significados y establecer inferencias válidas a partir de los mensajes analizados (Krippendorff, 2018).

Enfoque y diseño de investigación

El presente estudio se desarrolló bajo un enfoque cualitativo de alcance descriptivo, ya que busca caracterizar y sistematizar la información teórica y normativa existente, y utilizar ese conocimiento para elaborar un instrumento que responda a las necesidades del contexto educativo. Este enfoque se compone de cuatro fases descritas por Charry y Navarro (2018) así:

1. **Preparatoria.** Revisión y análisis documental, se identifican y analizan los referentes teóricos, pedagógicos y normativos nacionales e internacionales sobre la evaluación de competencias lingüísticas en inglés en básica primaria.
2. **Trabajo de campo.** Diseño del instrumento, a partir de los hallazgos del análisis documental, se determinan los componentes, criterios y descriptores que conforman el instrumento de diagnóstico.
3. **Analítica.** Prueba piloto, se aplica el instrumento en un contexto escolar controlado con el fin de evaluar su pertinencia, claridad y aplicabilidad; se registran observaciones y se realizan ajustes en la redacción, formato o nivel de dificultad de los ítems según los resultados obtenidos.
4. **Informática.** Validación por juicio de experto, el instrumento ajustado será sometido a un proceso de validación de contenido realizado por un experto designado por la universidad, con formación en enseñanza del inglés y evaluación educativa. El experto revisará la coherencia del instrumento con los EBC y con los descriptores del nivel A1 del MCER, valorando aspectos de claridad, pertinencia, congruencia, representatividad y adecuación al contexto escolar. Sus observaciones se recogerán mediante una ficha de valoración cualitativa y se analizarán con el fin de identificar

coincidencias y ajustes necesarios antes de la versión final del instrumento. Este proceso garantiza la validez de contenido y la consistencia teórica y práctica del diseño propuesto.

Este enfoque permite vincular la teoría con la práctica educativa, garantizando la coherencia entre los referentes conceptuales, la construcción del instrumento y su validación contextual.

Contexto y participantes

La población objeto de estudio está conformada por 144 estudiantes de grado quinto de dos instituciones educativas públicas, con edades entre 9 y 12 años, pertenecientes a contextos rurales y urbanos y con niveles socioeconómicos estratificados entre los niveles 1 y 5.

La Institución Educativa Adolfo María Jiménez de Sotaquirá, Boyacá, cuenta con un total de 31 estudiantes matriculados en el grado quinto en la sede central, conformado por 16 niños y 15 niñas. La mayoría de los estudiantes provienen de veredas aledañas al colegio. La población pertenece principalmente a los estratos socioeconómicos 1 y 2, por lo que muchas familias enfrentan condiciones económicas desfavorables que inciden en el acompañamiento escolar y en la disponibilidad de materiales de apoyo para el aprendizaje.

En términos generales, se observa que varios estudiantes provienen de núcleos familiares con dinámicas complejas, lo cual puede repercutir en su rendimiento académico y en su comportamiento dentro del aula.

Por su parte, la Institución Educativa Lucrecio Jaramillo Vélez cuenta con un total de 113 estudiantes en el grado quinto en la sede Agrupación Colombia, conformado por 64 niños y 49 niñas. La mayoría de los estudiantes provienen de barrios aledaños a la institución. La

población pertenece principalmente a los estratos socioeconómicos 3 y 4, con estudiantes también en los estratos 1, 2 y 5.

De acuerdo con el Proyecto Educativo Institucional (PEI, 2019) se destacan diversas fortalezas institucionales. La mayoría de las familias posee un buen capital cultural, están integradas en su mayoría por padres y acudientes profesionales, con un alto nivel educativo y las habilidades necesarias para guiar académicamente a sus hijos; que muestran interés y compromiso en el acompañamiento escolar de los mismos, lo que favorece el desarrollo académico y personal de los estudiantes. Además, la mayoría de los estudiantes cuenta con los recursos básicos para satisfacer sus necesidades educativas y participar activamente en las actividades pedagógicas.

Los docentes se caracterizan por ser profesionales comprometidos, cercanos y atentos a las necesidades individuales de los estudiantes, promoviendo ambientes de aprendizaje seguros, inclusivos y estimulantes. La institución también ofrece procesos pedagógicos claros y estructurados, lo que facilita el seguimiento académico y el fortalecimiento de hábitos de estudio desde edades tempranas.

El entorno físico y urbano de la sede Agrupación Colombia brinda condiciones adecuadas de accesibilidad, seguridad y bienestar, lo que contribuye al buen desarrollo de las jornadas escolares. Asimismo, la imagen positiva de la institución en la comunidad y el apoyo de programas gubernamentales y sociales fortalecen la formación integral de los niños en esta etapa fundamental de su educación.

La población total de las dos instituciones es de ciento cuarenta y cuatro estudiantes de grado quinto de básica primaria, la muestra fue de treinta estudiantes, equivalente a

aproximadamente el 20 % de la población total de ambas instituciones, distribuida en 18 niños y 12 niñas.

Instrumentos

A continuación, presentamos los instrumentos del presente estudio: matriz de análisis documental, prueba diagnóstica, encuesta de pertinencia y rúbrica de validación.

1. Matriz de análisis

Esta matriz (Anexo 1) se diseñó en tres tablas con cinco columnas que abordan los referentes teóricos, pedagógicos y normativos nacionales en la evaluación de competencias en inglés en ambientes de aprendizaje de básica primaria. Cada tabla presenta los campos Descripción, Aportes, Elementos a tener en cuenta y Limitaciones o retos, lo que permite realizar un examen sistemático y comparativo de los documentos revisados.

En el caso de los referentes normativos, la matriz integra políticas y lineamientos del MEN como el Programa Nacional de Bilingüismo (2006), los EBC (2006), los DBA (2016), las Mallas de Aprendizaje (2014), las Pruebas Saber 3°, 5° y 9° (ICFES, 2023) y el MCER (2020). El análisis de estos documentos permitió identificar los principios de validez, confiabilidad y pertinencia requeridos para el diseño de instrumentos de evaluación coherentes con el contexto colombiano.

La segunda tabla, correspondiente a los referentes teóricos, recoge los aportes de autores clave en el campo de la evaluación lingüística como Fulcher (2012, 2015), Alderson (2015), Giraldo *et al.* (2023), Ordoñez y Rodríguez (2017), Torres-Lara *et al.* (2021), y Ye (2024), entre otros. Desde estas perspectivas se destacan los conceptos de literacidad en

evaluación, diagnóstico pedagógico, validez y confiabilidad técnica, y la importancia de contextualizar la evaluación a los ambientes de aprendizaje en básica primaria.

La tercera tabla, centrada en los referentes pedagógicos, articula los enfoques de evaluación formativa, contextualización y equidad, participación y literacidad docente. Este componente permitió definir criterios pedagógicos para el diseño de un instrumento que promueva la retroalimentación, la autorregulación del aprendizaje y la toma de decisiones pedagógicas fundamentadas.

En conjunto, la matriz permitió sistematizar y triangular los referentes que sustentan la construcción del instrumento de evaluación, garantizando que su diseño esté alineado con los estándares nacionales e internacionales, los principios pedagógicos contemporáneos y las realidades del aula de básica primaria colombiana.

2. Prueba diagnóstica

El segundo instrumento fue una prueba diagnóstica (Anexo 2) en inglés dirigida a estudiantes de grado quinto de educación básica primaria, con el propósito de evaluar sus habilidades en comprensión auditiva, comprensión de lectura, vocabulario, gramática y producción oral, en coherencia con los EBC y los descriptores del nivel A1 del MCER.

El instrumento está conformado por siete partes, cada una orientada a valorar un componente específico de la competencia comunicativa en inglés:

- ***Part 1: Lexical Knowledge (Conocimiento lexical)***

Evalúa el reconocimiento y la comprensión de vocabulario relacionado con las partes del cuerpo. Incluye cinco ítems de selección múltiple con única respuesta, en los que el estudiante debe escoger la palabra que corresponde a una descripción dada. Este



tipo de pregunta permite identificar el nivel de conocimiento lexical y su relación con contextos cotidianos.

- ***Part 2: Pragmatic and Sociolinguistic Knowledge (Conocimiento pragmático y sociolingüístico)***

Valora la comprensión y el uso funcional del lenguaje en situaciones comunicativas comunes. Contiene cinco preguntas de opción múltiple con única respuesta, en las que los estudiantes deben seleccionar la respuesta más adecuada para completar diálogos breves. Este formato permite medir la competencia pragmática y la pertinencia sociolingüística de las expresiones utilizadas.

- ***Part 3: Grammar and Language Use (Gramática y uso del lenguaje)***

Evalúa la aplicación de estructuras gramaticales y vocabulario en oraciones simples. Está compuesta por nueve ítems de selección múltiple con única respuesta, orientados a identificar el dominio de tiempos verbales (presente simple, pasado simple y presente continuo), el uso de auxiliares y de preposiciones. Se seleccionó este formato por su eficacia para medir precisión gramatical y consistencia en el uso de estructuras básicas del idioma.

- ***Part 4: Reading Comprehension (Comprensión lectora)***

Incluye un texto narrativo corto titulado “*Tom’s Pet Turtle*”, seguido de cuatro preguntas de selección múltiple con única respuesta que evalúan la comprensión literal de información explícita (personajes, acciones y entorno). Este tipo de texto y formato se consideran adecuados al nivel A1 del MCER y al desarrollo lector esperado para estudiantes de grado quinto.

- **Part 5: Lexical and grammar knowledge (Conocimiento gramatical y lexical)**

Evalúa fluidez, pronunciación y adecuación léxica en preguntas personales. Permite observar la competencia comunicativa real del estudiante. Combina gramática y vocabulario contextual. Requiere lectura detallada para seleccionar palabras coherentes y correctas gramaticalmente.

- **Part 6: Listening Comprehension (Comprensión auditiva)**

Presenta un texto oral titulado “*A Day at the Farmer’s Market*”, en el que Olivia y su hermano visitan el mercado local los sábados. A partir del audio, los estudiantes responden seis preguntas de selección múltiple con única respuesta que indagan sobre información general y específica del relato (lugar, actividades, personajes y acciones). Esta sección evalúa la comprensión auditiva global y literal de mensajes orales breves relacionados con experiencias cotidianas.

- **Part 7: Speaking (Producción oral)**

Evalúa la producción oral controlada mediante diez preguntas cortas que el evaluador formula individualmente a cada estudiante. Las preguntas abordan información personal, gustos, rutinas y descripciones simples. Esta parte permite valorar la pronunciación, la fluidez y la capacidad de responder a preguntas simples utilizando vocabulario y estructuras propias del nivel A1 del MCER.

En total, la prueba consta de 50 ítems distribuidos en siete secciones, predominantemente de selección múltiple con única respuesta, complementadas con una sección de producción oral que permite valorar la interacción comunicativa básica. El formato facilita la recolección y el análisis cuantitativo de los datos, al tiempo que promueve la evaluación integral de las habilidades lingüísticas receptivas y productivas.

El instrumento se diseñó con criterios de claridad, validez de contenido y adecuación al contexto escolar colombiano, priorizando instrucciones simples, vocabulario conocido y un equilibrio entre las dimensiones lingüísticas, pragmáticas y textuales de la competencia comunicativa en inglés.

3. Encuesta de pertinencia

La encuesta de percepción (Anexo 3) constituye el tercer instrumento del estudio y tiene como propósito recoger la opinión de los estudiantes de grado quinto acerca de su experiencia al presentar la prueba diagnóstica de inglés. Este instrumento busca identificar la claridad, pertinencia y accesibilidad de la prueba desde la perspectiva de los participantes, lo que permite complementar la información obtenida en la fase piloto y realizar ajustes al diseño del instrumento evaluativo.

La encuesta está conformada por siete ítems cerrados y una pregunta abierta. Los ítems se estructuran con una escala de tipo Likert de cuatro niveles (Muy de acuerdo, De acuerdo, En desacuerdo, Muy en desacuerdo), orientada a valorar aspectos como:

- La claridad de las instrucciones en cada parte de la prueba.
- La adecuación del número de preguntas.
- La suficiencia del tiempo asignado para responder.
- La comprensibilidad de los enunciados.
- La calidad del audio en la sección de comprensión auditiva.
- La dificultad percibida en las preguntas de escucha y de producción oral.

Adicionalmente, se incluye un espacio para comentarios abiertos, donde los estudiantes pueden expresar observaciones, dificultades o sugerencias respecto al contenido, formato o desarrollo de la prueba.

El diseño de esta encuesta responde a criterios de validez de contenido y pertinencia contextual, al utilizar un lenguaje claro y accesible para niños de 10 a 12 años, y al enfocarse en dimensiones relevantes para evaluar la experiencia del usuario frente al instrumento aplicado. La información obtenida mediante esta encuesta se analiza de manera descriptiva para identificar patrones de percepción y oportunidades de mejora en la versión final del instrumento de evaluación.

4. Rúbrica de validación

La validación del instrumento constituye la fase final del proceso metodológico, cuyo propósito es verificar la validez de contenido, coherencia interna y pertinencia contextual de la prueba diagnóstica diseñada para evaluar las competencias lingüísticas en inglés en grado quinto.

En esta etapa, el instrumento conformado por siete secciones que valoran vocabulario, uso pragmático del lenguaje, gramática, comprensión lectora, comprensión auditiva y producción oral fue sometido a la revisión de un experto con formación en enseñanza del inglés y evaluación educativa.

El especialista analiza la correspondencia del instrumento con los EBC y con los descriptores del nivel A1 del MCER.

La validación se desarrolla mediante una rúbrica de valoración (Anexo 4) cualitativa y cuantitativa, la cual permite determinar el grado de cumplimiento de criterios como claridad, pertinencia, congruencia, representatividad y adecuación al contexto escolar colombiano. Las observaciones obtenidas se analizan e incorporan para realizar los ajustes finales al instrumento antes de su aplicación definitiva, garantizando su consistencia teórica y aplicabilidad práctica.

Método de recolección de datos

Es importante contextualizar que los instrumentos de evaluación fueron aplicados en dos instituciones educativas con características significativamente distintas. La primera correspondió a una institución rural del departamento de Boyacá, donde las dinámicas pedagógicas, los recursos disponibles y las experiencias previas de los estudiantes con el inglés reflejaban un contexto educativo de zonas rurales. La segunda institución se ubicó en la ciudad de Medellín, en un entorno urbano con condiciones organizativas y tecnológicas diferentes.

En primer lugar, la prueba diagnóstica en la Institución Educativa Adolfo María Jiménez se aplicó en formato impreso el 14 de noviembre de 2025. Inicialmente, se había previsto realizarla en la sala de informática; sin embargo, esto no fue posible debido a la falta de disponibilidad y a problemas de conectividad. Por tal motivo, y con la autorización del rector y del director de grado quinto, se llevó a cabo en la biblioteca, espacio adecuado para trabajar con el grupo seleccionado como muestra.

La jornada inició a las 7:00 a. m., con la bienvenida a los diez estudiantes participantes (cinco niñas y cinco niños). La docente agradeció su asistencia y reiteró el

propósito de la actividad, explicando que esta hacía parte del proyecto de investigación y no tendría una calificación. Este recordatorio tuvo como intención disminuir la ansiedad y promover respuestas honestas de acuerdo con el nivel real de competencia lingüística.

Cuando los estudiantes ya habían comenzado la prueba, el docente titular de inglés de primaria se acercó a la biblioteca solicitando la presencia del grupo en el aula, dado que se encontraba cerrando actividades del área. Esta solicitud no fue atendida, pues los estudiantes contaban con las autorizaciones institucionales necesarias. Se espera que esta situación no haya influido de manera negativa en el desempeño de los participantes.

Los estudiantes comenzaron desarrollando las primeras cinco partes de la prueba: el plan consistía en continuar luego con la sección de escucha; no obstante, dado que no todos finalizaron al mismo tiempo, la docente decidió avanzar paralelamente con la prueba de producción oral. Para ello, fue llamando individualmente a los estudiantes que ya habían completado la parte escrita, mientras los demás continuaban trabajando en las preguntas restantes.

Al finalizar la entrevista oral, cada estudiante diligenció el Instrumento 3: Encuesta de percepción de la prueba. Cuando se habían entrevistado cuatro participantes, el resto del grupo informó haber concluido la parte escrita. En ese momento, se realizó una pausa en las entrevistas para avanzar con la sección de escucha (*listening*), la cual se aplicó empleando un teléfono celular y un parlante.

Para esta sección, se siguieron las fases recomendadas para el desarrollo de actividades de comprensión auditiva en inglés:

1. Pre-listening (Antes de la escucha)

En esta fase, la docente preparó a los estudiantes para la actividad auditiva y redujo posibles barreras de comprensión.

- **Introducción al tema:** Se contextualizó brevemente el contenido del audio mediante preguntas generales relacionadas con la situación presentada, con el fin de activar la familiaridad con el tópico.
- **Activación de conocimientos previos:** A través de preguntas guía, los estudiantes anticiparon la temática del audio, favoreciendo la conexión con experiencias previas.
- **Pre-enseñanza del vocabulario esencial:** Se presentaron únicamente las palabras que podrían afectar la comprensión global del fragmento auditivo, evitando una carga excesiva de términos nuevos.
- **Explicación de la tarea:** Se indicó de forma clara qué información debían identificar durante la escucha, especificando si debían captar la idea principal o atender a detalles específicos.

2. *While-listening* (Durante la escucha)

En esta fase, los estudiantes interactuaron directamente con el audio para completar la tarea propuesta.

- **Primera escucha: Comprensión general:** Se reprodujo el audio de manera continua, con el objetivo de que los estudiantes identificaran el sentido global del mensaje.

- **Segunda escucha: Comprensión específica:** El audio se repitió para permitir la búsqueda de información puntual, como respuestas a preguntas o verificación de predicciones iniciales.

Hacia las 8:30 a. m., los estudiantes tuvieron un descanso de 30 minutos y, posteriormente, se retomó la prueba de producción oral con los seis participantes restantes. Cada uno completó la encuesta de percepción al finalizar su entrevista. Durante toda la jornada, se observó que algunos estudiantes manifestaron nerviosismo o ansiedad. En tales casos, se les brindó acompañamiento cercano, recordándoles que la prueba buscaba orientar futuras decisiones pedagógicas dirigidas a fortalecer el aprendizaje del inglés, más que evaluar con fines calificativos. La aplicación de la prueba finalizó a las 9:25 a.m.

Por otro lado, para la aplicación de la prueba diagnóstica y la encuesta de percepción en la Institución Educativa Lucrecio Jaramillo Vélez, sede Agrupación Colombia, la prueba diagnóstica fue cargada y aplicada a través del programa Master 2000, una plataforma institucional utilizada por su facilidad en la gestión académica, su agilidad para asignar y calificar actividades, el ahorro de papel que permite, y porque los estudiantes ya estaban familiarizados con su uso. Desde este entorno se configuró el instrumento para cada estudiante, además, de la encuesta de pertinencia. Una vez cargadas la prueba y la encuesta fueron compartidas con el profesor de inglés de grado quinto, quien recibió toda la información sobre la prueba, sus objetivos y su carácter diagnóstico. Una vez revisado el instrumento, lo asignó a los tres grupos de grado quinto de la jornada de la tarde en la sede Agrupación Colombia de la Institución Educativa Lucrecio Jaramillo Vélez, en los horarios que consideró pertinentes. Durante la aplicación se desarrollaron las cinco primeras secciones de la prueba, junto con la encuesta diagnóstica complementaria.

La sección de comprensión auditiva se realizó posteriormente de manera presencial en un salón con televisor, lo que permitió reproducir el audio desde un computador con suficiente claridad. El sonido fue adecuado y la actividad se llevó a cabo sin inconvenientes técnicos. Sin embargo, se evidenció la poca práctica previa de los estudiantes en la comprensión auditiva, por lo que el docente decidió disminuir la velocidad del audio y reproducirlo cuatro veces para favorecer la comprensión.

Por su parte, la sección de competencia comunicativa no se aplicó mediante Master 2000 por la naturaleza de la prueba oral. Además, el docente manifestó que no se sentía con la idoneidad para evaluar este componente, por lo que la actividad fue realizada directamente por la investigadora, (es importante mencionar que el docente es Licenciado en Educación Física y tiene a su cargo las áreas de Educación Física, Tecnología e Inglés). El profesor presentó a la investigadora ante los estudiantes, explicó en qué consistiría la prueba y organizó la salida de los participantes para la entrevista individual, creando un ambiente adecuado para el desarrollo de la tarea.

En cuanto a la actitud de los estudiantes durante la prueba de competencia comunicativa, la mayoría llegó a la entrevista con una mezcla de nerviosismo y curiosidad, propia de una actividad poco frecuente en su experiencia escolar. Algunos mostraron signos iniciales de inseguridad, como hablar en voz baja, evitar el contacto visual o dudar antes de responder; mientras que otros se acercaron con una actitud más expectante y receptiva, interesados en la dinámica de interacción individual. No obstante, conforme avanzaba la entrevista, la mayoría de los estudiantes evidenció una mayor relajación y disposición y se obtuvieron respuestas más espontáneas y representativas.

ANÁLISIS DE RESULTADOS

El análisis de los resultados se desarrolló en tres momentos complementarios que permitieron valorar la pertinencia, claridad y coherencia del instrumento. En primer lugar, se realizó un análisis descriptivo de la prueba piloto, a partir de los resultados obtenidos por los estudiantes y de la información recopilada mediante la encuesta aplicada posteriormente. Este análisis permitió identificar el grado de comprensión de las instrucciones, la claridad de los ítems y la pertinencia de las tareas propuestas en relación con el nivel de los participantes. Los resultados fueron examinados de manera descriptiva, considerando patrones de respuesta, dificultades recurrentes y percepciones reportadas por los estudiantes. De esta manera, se valoró si las actividades evaluaban efectivamente las competencias lingüísticas previstas y si las características del instrumento resultaban adecuadas para el contexto de educación básica primaria. Este procedimiento se alineó con los planteamientos de Bachman and Palmer (2010), quienes señalan que el análisis de resultados de una evaluación debe considerar tanto el desempeño observado como la interpretación contextualizada de las evidencias obtenidas, con el fin de justificar el uso del instrumento dentro de un proceso de evaluación educativa.

En segundo lugar, se desarrolló el análisis de la validación por expertos, en el cual se examinaron los juicios emitidos por especialistas en enseñanza y evaluación del inglés respecto a la claridad, pertinencia y coherencia del instrumento. Sus observaciones permitieron identificar ajustes necesarios en la redacción de ítems, en la organización de las tareas y en la correspondencia entre los objetivos evaluativos y las actividades propuestas.

Análisis descriptivo de la prueba piloto

A partir del procesamiento de los resultados de la prueba diagnóstica, en contextos rurales y urbanos, y de la encuesta de percepción aplicada a los estudiantes, las investigadoras llevaron a cabo un análisis orientado a describir la pertinencia y la claridad del instrumento. Para ello, se realizó una revisión de los desempeños obtenidos en los diferentes componentes de la prueba, identificando patrones de respuesta y comportamientos diferenciados entre los contextos evaluados. Por ejemplo, en el componente lexical se identificó un patrón diferenciado entre el reconocimiento y el uso del vocabulario. Así, los estudiantes tendieron a desempeñarse mejor en tareas de reconocimiento (asociación palabra–imagen o selección de significado en contexto), mientras que presentaron mayores dificultades en ítems que implicaban el uso contextualizado del léxico. En cuanto a los formatos de ítems, se observó una mayor consistencia en los patrones de respuesta, evidenciada en la selección recurrente de opciones similares frente a ítems del mismo tipo, lo cual sugiere familiaridad con el formato de evaluación y menor dispersión en las elecciones.

Finalmente, las interpretaciones derivadas del análisis se contrastaron con la información proveniente de la encuesta, lo que permitió incorporar la percepción de los estudiantes sobre la comprensión del instrumento. Este enfoque se enmarca en perspectivas contemporáneas de validación, que sostienen que la validez se construye a partir de la convergencia de diversas evidencias. En este sentido, como lo proponen Messick (1990) y Kane (2006), la validez debe entenderse como un argumento integrado que combina distintas fuentes de evidencia y no como una propiedad estática del instrumento.

En este proceso, las investigadoras identificaron variaciones en el desempeño entre los contextos evaluados, lo cual fue interpretado a partir de la relación entre las características de los estudiantes y las demandas del instrumento. Este análisis no se limitó a la observación de resultados globales, sino que implicó la revisión de los diferentes componentes evaluados, permitiendo reconocer cómo se comportaban las habilidades lingüísticas en función del tipo de tarea propuesta. De esta manera, las conclusiones no se derivaron únicamente de los resultados cuantitativos, sino de su articulación con el análisis cualitativo de la estructura del instrumento y la naturaleza de los ítems.

En cuanto a la pertinencia, las investigadoras analizaron la correspondencia entre los resultados obtenidos y los constructos que el instrumento pretendía evaluar. Este análisis se sustentó en la observación de que las variaciones en el desempeño no se presentaron de manera aleatoria, sino que evidenciaron una relación con las características de los contextos educativos y con las habilidades lingüísticas implicadas en cada componente. A partir de esta relación, se interpretó que el instrumento logró captar diferencias en los niveles de dominio de los estudiantes, lo cual se asocia con los principios de validez de contenido y sensibilidad evaluativa.

Esta sensibilidad del instrumento se hizo particularmente evidente al contrastar los resultados de estudiantes de contextos rurales y urbanos. No se trató simplemente de registrar una diferencia de desempeños, sino de observar cómo esa diferencia se manifestaba de manera lógica en componentes específicos. Por ejemplo, si el contexto rural ofrecía menor exposición a ciertos tipos de textos o prácticas comunicativas, era de esperarse que las diferencias fueran más marcadas en los componentes que evaluaban esas habilidades, mientras que, en otros, como el reconocimiento lexical básico, los desempeños podían ser

más homogéneos. Esta capacidad de la prueba para reflejar diferencias contextualmente explicables constituye una evidencia sólida de su pertinencia, ya que demuestra que mide el constructo de competencia comunicativa de una manera que no es ajena a las realidades de los estudiantes.

La importancia de esta adecuación contextual se ve reforzada por los hallazgos de Negrete Raura *et al.* (2025), quienes, en su análisis sobre la validación de instrumentos en educación básica, identificaron que una de las principales limitaciones en este campo es la “falta de adaptación cultural, lingüística y cognitiva”, lo que “compromete la validez cultural de los instrumentos, especialmente en poblaciones vulnerables”. Esta preocupación por la validez cultural no es exclusiva del diseño de pruebas, sino que también permea el ámbito de la formación inicial del profesorado. García-Manzanares *et al.* (2025), en su estudio sobre futuros docentes de Educación Física para contextos bilingües, concluyen que existe un claro 'desajuste entre la oferta curricular actual y las exigencias derivadas de un contexto educativo cada vez más bilingüe'. Este paralelismo refuerza la noción de que tanto un instrumento de evaluación como un plan de estudios demuestran su pertinencia en la medida en que logran captar y responder a las demandas específicas de los contextos para los que fueron diseñados, ya sea el aula de clase o la realidad profesional futura del estudiante.

En coherencia con estos planteamientos, el comportamiento diferencial pero lógico observado en la prueba piloto sugiere que el instrumento no solo es técnicamente sólido, sino que también posee la validez cultural necesaria para ser aplicado en la diversidad de contextos para los que fue diseñado, evitando así los sesgos que surgen de instrumentos descontextualizados.

En relación con la claridad del instrumento, las investigadoras partieron del análisis de los resultados de la encuesta, en la cual los estudiantes reportaron su percepción frente a aspectos como la comprensión de instrucciones, la redacción de los ítems y la organización general de la prueba. Este análisis se complementó con la revisión de los patrones de respuesta en la prueba diagnóstica, con el fin de identificar posibles inconsistencias que pudieran atribuirse a dificultades de comprensión. A partir de este análisis, se estableció que los estudiantes lograron interpretar adecuadamente las demandas de las tareas propuestas.

El análisis de la información proveniente de la encuesta de percepción y los patrones de respuesta fue clave para esta conclusión. Si bien los estudiantes manifestaron comprender las instrucciones, fue el análisis de sus respuestas lo que confirmó esta percepción. Se observó, por ejemplo, una baja tasa de omisión de ítems y una alta consistencia en las respuestas a preguntas de un mismo formato, lo que sugiere que los estudiantes no se enfrentaron a barreras de comprensión que los llevaran a adivinar o a dejar de responder.

Se observó, por ejemplo, que, en el bloque de preguntas de comprensión lectora del contexto urbano, la gran mayoría de los estudiantes respondió todos los ítems sin dejarlos en blanco, registrándose una tasa de omisión prácticamente nula en ese segmento. Además, dentro de ese mismo bloque, las respuestas mostraron patrones consistentes, es decir, los estudiantes tendían a mantener un nivel de acierto similar a lo largo de ítems con la misma estructura (por ejemplo, preguntas de identificación de idea principal o detalle específico), en lugar de presentar respuestas aleatorias o altamente variables.

De manera similar, en el contexto rural, aunque el nivel de acierto fue más bajo, los estudiantes también completaron la mayoría de los ítems, incluso en secciones de mayor dificultad como *listening*. Esto sugiere que, aunque no dominaron completamente la

competencia evaluada, sí comprendieron lo que se les pedía hacer, ya que no recurrieron sistemáticamente a omitir respuestas ni se evidenciaron patrones típicos de adivinación desorganizada. Estos últimos, entendidos como respuestas aleatorias no relacionadas con la habilidad del estudiante, son inconsistencias entre ítems de la misma tipología o distribuciones de respuesta sin coherencia interna. Este tipo de comportamiento ha sido descrito en la literatura reciente como *rapid guessing*, caracterizado por una baja implicación cognitiva y una probabilidad de acierto cercana al azar (Jin *et al.*, 2023; Ríos & Deng, 2022). Asimismo, estudios actuales en psicometría señalan que este tipo de respuestas tiende a ser independiente del nivel de competencia del estudiante y puede afectar la validez de las interpretaciones de los resultados (Wang & Xu, 2020). En este sentido, la ausencia de dichos patrones en los resultados analizados sugiere que los estudiantes interactuaron con la prueba desde un procesamiento efectivo de las tareas, lo cual fortalece la validez interpretativa de los resultados.

La claridad fue entendida, entonces, no solo como una característica formal del instrumento, sino como un elemento que se evidencia en la interacción efectiva entre el estudiante y la prueba. En este marco, la literatura reciente señala que la claridad de los ítems constituye un componente fundamental de la validez, en la medida en que reduce la ambigüedad y favorece que las respuestas reflejen el constructo evaluado y no dificultades de interpretación. En coherencia con estos planteamientos, las investigadoras interpretaron que la comprensión reportada por los estudiantes, junto con la ausencia de patrones erráticos atribuibles a problemas de redacción, constituyó evidencia de claridad en el instrumento.

No obstante, al contrastar esta información con los resultados de desempeño, se identificó que la comprensión de las instrucciones y de los ítems no implicó necesariamente un dominio equivalente de las competencias evaluadas. Este hallazgo se interpretó a partir de la distinción entre la claridad formal del instrumento y las demandas cognitivas de las tareas, entendiendo que un instrumento puede ser comprensible en su formulación, pero exigir procesos complejos que inciden en el desempeño de los estudiantes. Esta distinción es crucial, ya que permitió a las investigadoras validar que los ítems eran formalmente claros y que las dificultades observadas se debían, por lo tanto, a la complejidad inherente de la tarea y no a problemas de redacción o diseño.

Para profundizar en esta distinción, resulta determinante el trabajo de Tárraga-Mínguez y Gómez-Marí (2023) sobre accesibilidad cognitiva. Los autores definen este concepto como la adaptación de la información del entorno para que todos puedan comprenderla y destacan que va más allá de la simple claridad formal. Un entorno (o un instrumento de evaluación) puede tener instrucciones claras, pero si su diseño no considera las diversas capacidades cognitivas de los usuarios, puede seguir siendo inaccesible. En el contexto de la prueba piloto, el hecho de que los estudiantes comprendieran las instrucciones (claridad formal) pero tuvieran dificultades con tareas de “uso contextualizado del léxico” no implica una falla en la claridad, sino que evidencia la complejidad cognitiva de la tarea, un aspecto central de la accesibilidad cognitiva explicada previamente. Así, este instrumento no solo es claro, sino que también logra discriminar entre diferentes niveles de demanda cognitiva, lo que es una fortaleza desde la perspectiva de la evaluación.

Esta capacidad de discriminar puede ser entendida con mayor profundidad a la luz de los hallazgos de García-Manzanares *et al.* (2025), quienes, mediante un análisis de clústeres, identificaron diferentes perfiles de estudiantes con necesidades formativas específicas en el contexto AICLE (Aprendizaje Integrado de Contenidos y Lenguas Extranjeras). En una línea complementaria, estudios recientes en evaluación de lenguas han demostrado que el uso de técnicas de perfilamiento, como el análisis de clústeres o *latent profile analysis*, permite identificar patrones diferenciados de desempeño entre estudiantes, evidenciando la existencia de subgrupos con características lingüísticas y niveles de competencia distintos dentro de una misma población (Gao, 2022; Jin *et al.*, 2021).

De manera análoga, el hecho de que la prueba piloto revele patrones de respuesta diferenciados (por ejemplo, mejores resultados en reconocimiento léxico que en uso contextualizado) no implica una falta de claridad, sino que evidencia la existencia de perfiles de competencia entre los estudiantes. Se evidencia entonces que el instrumento, por tanto, no solo es claro en su formulación, sino que posee la sensibilidad necesaria para visibilizar distintos niveles de dominio del constructo, una cualidad esencial para cualquier herramienta de diagnóstico educativo.

En cuanto a la estructura del instrumento, se evidenció que este se organizó en torno a múltiples componentes de la competencia comunicativa, lo cual es coherente con enfoques contemporáneos de evaluación en lengua extranjera. La inclusión de diversas habilidades responde a la necesidad de evaluar el desempeño de manera integral, permitiendo una representación más precisa del constructo lingüístico, tal como lo plantean Chapelle (2020) y Bachman y Palmer (2010), quienes destacan que la evaluación de lenguas debe considerar



múltiples dimensiones de la competencia comunicativa para garantizar la validez de las interpretaciones.

Asimismo, el uso de terminología técnica en la categorización de competencias (por ejemplo, *lexical knowledge* o *pragmatic competence*) no pareció afectar negativamente a los estudiantes, lo cual indicó que estos elementos se gestionaron principalmente a nivel de diseño del instrumento y no interfirieron en la interacción directa con la prueba. Este hallazgo es consistente con los planteamientos de Chapelle (2020), quien destaca que la validez de una prueba depende de la adecuada articulación entre el constructo teórico y las tareas evaluativas, así como con los aportes de American Educational Research Association, American Psychological Association and National Council on Measurement in Education (2018), quienes subrayan la importancia de que el lenguaje del instrumento sea accesible y no introduzca dificultad innecesaria para el examinado. Igualmente, enfatizan la necesidad de una definición clara y operativa de los constructos en la evaluación del desempeño lingüístico, lo que implica que estos se materialicen en tareas comprensibles para los estudiantes. De manera complementaria, el enfoque de uso del test propuesto por Bachman y Palmer (2010) refuerza esta distinción al plantear que los argumentos de validez deben sustentarse en el diseño técnico del instrumento, mientras que la interacción con el examinado debe caracterizarse por la claridad, accesibilidad y pertinencia del lenguaje empleado.

El análisis permitió establecer que la claridad del instrumento se sustentó en un proceso de análisis entre resultados de desempeño y percepción estudiantil, así como en su articulación con referentes teóricos actuales. Estas características no se asumieron de manera

previa, sino que fueron inferidas a partir de la evidencia recogida, en coherencia con enfoques contemporáneos de validación en evaluación educativa.

En un segundo momento del proceso, se realizó el análisis en paralelo de las tres rúbricas de validación diligenciadas por expertos en evaluación de instrumentos diagnósticos en lengua extranjera. El presente análisis se fundamentó en el método de validación por juicio de expertos, el cual, como señalan Robles Garrote y Rojas (2015), constituye un “método de validación útil para verificar la fiabilidad de una investigación” al solicitar la opinión informada de personas con trayectoria en el tema.

Los expertos que participaron en este proceso fueron:

Experto 1: Mg. Sandra Rodríguez. Profesora de educación superior con más de 20 años de experiencia en la formación de docentes de idiomas. Cuenta con trayectoria en coordinación de grupos de investigación estudiantil y en procesos de formación en bilingüismo en modalidades presencial, virtual y a distancia. Fue coautora del proyecto que permitió la creación del Máster en Entornos de Aprendizaje Bilingües en la Universidad Santo Tomás. Posee experiencia en investigación formativa centrada en docentes en formación y en ejercicio, así como en el análisis del bilingüismo en contextos rurales y sus desafíos socioculturales.

Experto 2: Docente Catalina Herrera. Profesora e investigadora en educación y enseñanza de lenguas extranjeras. Cuenta con amplia experiencia en formación docente, investigación en bilingüismo y análisis de prácticas pedagógicas en contextos diversos, incluyendo educación rural y entornos virtuales. Ha participado en proyectos académicos y publicaciones sobre formación docente, educación en casa, competencias digitales y aprendizaje en contextos mediados por tecnología.

Experto 3: Docente Eduardo García. Profesional en educación con formación en Enseñanza de Lenguas Extranjeras, egresado de la Universidad de Antioquia, institución reconocida por su alta calidad académica y trayectoria en investigación y formación docente en Colombia. Cuenta, además, con estudios de Maestría en Educación en la Universidad Católica de Oriente. Desde 2005, se desempeña como docente en la IE Técnico Industrial Santiago de Arma, reconocida como una de las instituciones públicas más destacadas del departamento de Antioquia y del país, donde ha consolidado una trayectoria sólida y comprometida con la calidad educativa. Orientó cursos de comprensión lectora en inglés dirigidos a estudiantes de distintas carreras de la Universidad de Antioquia, espacio académico de carácter obligatorio para los procesos de graduación, lo que le permitió fortalecer competencias en formación universitaria y acompañamiento académico. Adicionalmente, se ha desempeñado como docente en el bachillerato de adultos (CLEI), ampliando su experiencia en educación flexible e inclusiva. Su amplia trayectoria en distintos niveles y contextos educativos, junto con su sólida formación académica, evidencian un alto nivel de criterio pedagógico, rigor académico y capacidad de análisis crítico.

Este análisis tuvo como propósito valorar la calidad del instrumento en función de seis criterios: claridad, pertinencia, congruencia interna, representatividad, adecuación al contexto y validez técnica. A continuación, se presenta el análisis detallado por criterio, respetando la información aportada por cada experto y estableciendo relaciones interpretativas fundamentadas.

En relación con la claridad del instrumento, la Experta 1 señaló que este cumple parcialmente con dicho criterio, indicando que *“las instrucciones deben optimizarse aclarando los nombres precisos de los componentes de la prueba diagnóstica”* y que



“algunas aclaraciones respecto al cómo se desarrollarán los componentes de listening y speaking no son claros para quienes desarrollarán la prueba”. Por su parte, la Experta 2 identificó una dificultad asociada al nivel léxico, al afirmar que *“hay vocabulario difícil de entender para los niños”*, se alinea con la literatura especializada. Autores como Gavilema Vaca y Tovar Viera (2025) identifican el “vocabulario limitado” como un obstáculo significativo para el desarrollo de competencias lingüísticas en niveles básicos, lo que subraya la necesidad de simplificar la carga léxica del instrumento. En contraste, el Experto 3 no realizó observaciones directas sobre la redacción de instrucciones o ítems, pero señaló que *“se necesita un buen proceso durante toda la primaria con docentes capacitados para el área”*. Esta afirmación, aunque no evalúa directamente el instrumento, introduce un elemento contextual que incide indirectamente en la comprensión de la prueba.

En conjunto, se evidencia que el instrumento presenta una claridad funcional, pero requiere ajustes específicos en la precisión de las instrucciones y en la simplificación del vocabulario para garantizar su plena accesibilidad al nivel de los estudiantes.

A propósito de la pertinencia, la Experta 1 destacó que *“las preguntas responden a intereses potenciales e información relacionada con ámbitos escolares”*, lo cual evidencia una adecuada alineación con el contexto educativo inmediato de los estudiantes. El Experto 3, por su parte, afirmó que el instrumento *“se encuentra muy ajustado a lo esperado en un nivel A1”*, lo que valida su coherencia con los descriptores del MCER.

En el caso de la Experta 2, no se registran observaciones cualitativas específicas para este criterio; sin embargo, tampoco se evidencian críticas explícitas, con una valoración general favorable dentro de los parámetros establecidos.

En síntesis, la pertinencia del instrumento se ve respaldada por el conjunto de valoraciones de los expertos, quienes coinciden en que las preguntas se alinean con los intereses y el contexto educativo de los estudiantes, sin evidenciarse observaciones críticas frente a este criterio. De manera particular, la apreciación del Experto 3, quien lo considera “muy ajustado a lo esperado en un nivel A1”, refuerza esta valoración global. Este señalamiento resulta clave, ya que, tal como lo establece el MCER, un usuario de nivel A1 es “capaz de comprender y utilizar expresiones cotidianas de uso muy frecuente [...] para satisfacer necesidades de tipo inmediato”. En este sentido, la concordancia del instrumento con dicho estándar internacional fortalece la validez del constructo evaluado.

Respecto a la congruencia interna, la Experta 1 sugirió “*evaluar si sería pertinente incluir una lectura con preguntas adicionales para el componente de Reading Comprehension*”, con el fin de lograr un mejor balance entre habilidades evaluadas. El Experto 3 complementó esta observación al indicar que “*no se evidencia mucho la habilidad de escritura (writing)*”, lo cual señala una posible limitación en la cobertura de las habilidades productivas. Por su parte, la Experta 2 no reportó observaciones explícitas en este criterio, y lo calificó con la puntuación positiva más alta en la escala dada.

En conjunto, se puede concluir que el instrumento presenta una estructura coherente entre objetivos, habilidades y tipos de ítems, pero requiere fortalecer el equilibrio entre las habilidades lingüísticas, especialmente en lo relacionado con la comprensión lectora y la producción escrita, para consolidar plenamente su congruencia interna.

En cuanto a la representatividad, la Experta 1 señaló que las pruebas diagnósticas suelen incluir “*preguntas relacionadas con imágenes en contexto*”, lo que evidencia una oportunidad de mejora en la diversidad de formatos de ítems. La Experta 2 indicó de manera

directa que “*en la sección de vocabulario solo se evalúan partes del cuerpo*”, lo cual pone de manifiesto una cobertura limitada del componente léxico. Por su parte, el Experto 3 sugirió que “*se podrían incluir lecturas que incluyan aspectos culturales de nuestro país*”, ampliando la discusión hacia la dimensión sociocultural de la competencia comunicativa.

De manera articulada, las observaciones de los expertos permiten concluir que el instrumento presenta una representatividad parcial, ya que, si bien aborda distintas habilidades, requiere ampliar la diversidad de contenidos, incluir apoyos visuales y fortalecer la dimensión cultural para lograr una cobertura más integral del constructo evaluado. En este sentido, la sugerencia en el criterio anterior del Experto 3 sobre la poca evidencia de la habilidad de escritura (*writing*) y la observación de la Experta 2 respecto a la limitación del vocabulario a “partes del cuerpo” evidencian debilidades en dicha representatividad. A la luz de estos hallazgos, la investigación de Gavilema Vaca y Tovar Viera (2025) subraya que el dominio de las habilidades productivas (habla y escritura) constituye el objetivo central de la competencia comunicativa, por lo que su Infrarrepresentación en una prueba diagnóstica limita su capacidad para ofrecer un panorama completo del nivel del estudiante.

Respecto del criterio de adecuación al contexto, la Experta 1 señaló que “*el contexto escolar es claro, pero el contexto auténticamente colombiano está ausente*”, lo que evidencia una limitación en la contextualización cultural del instrumento. Esta limitación refleja un problema estructural en la educación del país, donde a menudo existen “brechas entre el currículo oficial y la realidad educativa” (BBE Languages, 2025). Para que la prueba sea equitativa y significativa, es imperativo incorporar referentes culturales locales que conecten con la realidad del estudiante.

En esta tensión entre lo local y lo global, autores como Canagarajah (2005) ofrecen un marco teórico fundamental. Su concepto de glocalización en la enseñanza de lenguas postula que, en lugar de imponer metodologías o contenidos foráneos que ignoran el contexto del alumno, es necesario partir de lo local para llegar a lo global. De igual forma, aboga por una práctica pedagógica que valore los saberes y las realidades de las comunidades periféricas (como el contexto colombiano) como base para la construcción del conocimiento en una lengua extranjera. Esto implica que el instrumento de evaluación diagnóstica no solo debe medir competencias en inglés, sino también conectar con aspectos locales y regionales de la comunidad escolar para ser realmente significativo y culturalmente sensible.

Por su parte, el Experto 3 valoró positivamente el nivel de dificultad, afirmando que *“el contenido y nivel de dificultad es adecuado para el último grado de básica primaria”*, lo cual valida su pertinencia en términos de grado escolar. En el caso de la Experta 2, no se registran observaciones específicas en este criterio.

Adicionalmente, la Experta 1 recomendó incorporar elementos formales relacionados con el *habeas data* y aclarar las implicaciones de la prueba diagnóstica, lo cual introduce una dimensión ética y procedimental relevante.

En síntesis, el instrumento presenta una adecuación parcial al contexto, ya que, si bien es valorado como pertinente en términos de nivel de dificultad y grado escolar, evidencia limitaciones en su contextualización cultural al no incorporar referentes propios del entorno colombiano. Las observaciones destacan la necesidad de integrar elementos locales que hagan la prueba más significativa y equitativa para los estudiantes, en consonancia con enfoques como la glocalización, que proponen articular lo local y lo global en la enseñanza de lenguas. Asimismo, se identifican aspectos de mejora en la dimensión ética y

procedimental, como la inclusión de consideraciones sobre el habeas data y la claridad en los propósitos de la prueba, lo que contribuiría a fortalecer su pertinencia y validez en el contexto educativo.

En el último criterio, validez técnica, la Experta 1 afirmó que *“las preguntas están dirigidas a diagnosticar el nivel de comprensión, según las habilidades incluidas”*, lo que respalda la correspondencia entre los ítems y los constructos evaluados. Por su parte, el Experto 3 advirtió que *“la entrevista (speaking) puede afectar el desempeño, incluso de estudiantes con buen nivel, si se hace face-to-face”*, encuentra un sólido respaldo en la investigación contemporánea. Factores psicológicos como el miedo a la crítica, la ansiedad y la falta de confianza son barreras emocionales significativas que influyen en el desempeño oral, independientemente del conocimiento lingüístico del estudiante (Gavilema Vaca & Tovar Viera, 2025). La Experta 2 no reportó observaciones específicas en este criterio y, nuevamente, lo calificó con la puntuación positiva más alta en la escala dada.

En conjunto, se puede afirmar que el instrumento presenta una validez técnica adecuada, sustentada en el uso pertinente de formatos de ítems; no obstante, se identifican posibles factores de sesgo en la evaluación de producción oral que podrían afectar la confiabilidad de los resultados.

En conclusión, el análisis conjunto de las tres rúbricas de validación permite establecer que el instrumento presenta una base sólida en términos de diseño técnico, coherencia estructural y alineación con estándares internacionales (MCER) y nacionales (MEN). Los expertos coinciden en reconocer su pertinencia para el nivel A1 y su adecuación general al contexto escolar de básica primaria.

No obstante, el proceso de validación también permitió identificar aspectos de mejora, entre los que se destacan:

- La necesidad de optimizar la claridad de las instrucciones y ajustar el nivel léxico.
- El fortalecimiento de la representatividad del componente léxico y de las habilidades lingüísticas, especialmente *reading* y *writing*.
- La incorporación de elementos visuales y diversidad de tareas.
- La mejora en la contextualización cultural, integrando referentes propios del contexto colombiano.
- La revisión de posibles sesgos en la aplicación del componente oral (*speaking*).
- La inclusión de elementos formales y éticos en la presentación del instrumento.

En síntesis, el proceso de validación por juicio de expertos no solo ha permitido identificar fortalezas y debilidades específicas del instrumento, sino que también ha puesto de manifiesto la necesidad de atender factores que la literatura especializada considera fundamentales para una evaluación lingüística auténtica. La pertinencia con el MCER queda demostrada, pero los hallazgos evidencian que la validez técnica (Robles Garrote & Rojas, 2015) se verá fortalecida en la medida en que se aborden las limitaciones en claridad y representatividad, se mitiguen los sesgos emocionales en la evaluación de producción oral (Gavilema Vaca & Tovar Viera, 2025) y se logre una contextualización cultural más cercana a la realidad colombiana (BBE Languages, 2025). Por lo tanto, los ajustes propuestos no son meramente formales, sino que son esenciales para garantizar que el instrumento cumpla con los más altos estándares de equidad, precisión y significatividad para la población objetivo.



CONCLUSIONES

El presente capítulo tiene como propósito sintetizar los principales hallazgos de la investigación en relación con los objetivos planteados, así como discutir sus aportes teóricos y prácticos en el campo de la evaluación de lenguas extranjeras en educación básica primaria en contextos colombianos. A partir del análisis de tres fuentes de evidencia, desempeño y percepción de los estudiantes y validación por juicio de expertos, se construye un argumento integrado de validez que permite emitir un juicio objetivo sobre la calidad del instrumento diseñado. Asimismo, se abordan las limitaciones del estudio y se proponen mejoras para el mismo que permitan profundizar y ampliar los resultados obtenidos.

De acuerdo con el objetivo general de la investigación de diseñar un instrumento de evaluación contextualizado, asequible y pertinente que permita el diagnóstico de las competencias lingüísticas en inglés en ambientes de aprendizaje de la básica primaria. Los resultados obtenidos permiten afirmar que dicho objetivo fue alcanzado satisfactoriamente, en la medida en que se logró construir un instrumento técnicamente sólido, pertinente para el nivel A1 del MCER y funcional para los propósitos diagnósticos planteados. El análisis de las evidencias mencionadas anteriormente constituye un argumento robusto a favor de la validez del instrumento, en coherencia con los planteamientos de Messick (1990) y Kane (2006), quienes conciben la validez no como una propiedad estática, sino como un proceso argumentativo sustentado en múltiples fuentes de evidencia.

Con respecto al primer objetivo específico, orientado a analizar referentes teóricos, pedagógicos y normativos, los resultados evidencian que el instrumento se encuentra alineado con los principios del MCER y con los EBC. Esta alineación fue confirmada por los

tres expertos, particularmente por el Experto 3, quien señaló textualmente que el instrumento “*se encuentra muy ajustado a lo esperado en un nivel A1*”. Asimismo, su diseño responde a enfoques contemporáneos de evaluación, especialmente aquellos que conciben la validez como un proceso argumentativo (Chapelle, 2020; Bachman & Palmer, 2010). Por lo tanto, se evidencia la necesidad de optimizar la claridad de las instrucciones y ajustar el nivel léxico. Aunque los resultados de la prueba piloto muestran que los estudiantes lograron interactuar con el instrumento, las dificultades asociadas al vocabulario sugieren la presencia de una carga cognitiva adicional que puede interferir con la medición del constructo.

En cuanto al segundo objetivo específico, relacionado con la identificación de elementos clave para la evaluación de competencias en inglés, el análisis permitió establecer que la inclusión de múltiples componentes (léxico, gramática, comprensión lectora, comprensión auditiva y producción oral) favorece una aproximación integral al constructo de competencia comunicativa. Los patrones de respuesta observados en la prueba piloto, mejor desempeño en tareas de reconocimiento léxico y mayores dificultades en tareas de uso contextualizado, evidencian que el instrumento logra discriminar entre distintos niveles de demanda cognitiva, lo cual constituye una fortaleza desde la perspectiva de la sensibilidad evaluativa. Teniendo en cuenta las observaciones hechas por los expertos 2 y 3, las cuales señalan que se requiere el fortalecimiento de la representatividad del componente léxico y de las habilidades lingüísticas, especialmente en lo relacionado con *reading* y *writing*, la limitada inclusión de estas habilidades restringe el alcance del instrumento. Desde la perspectiva de Messick (1990), esta situación afecta la validez de contenido, al no representar de manera suficiente el constructo evaluado. En consecuencia, se propone ampliar los campos semánticos del componente léxico e incorporar tareas de producción escrita (esta

última, condicionada a un ente evaluador idóneo) acordes al nivel de los estudiantes, garantizando una evaluación más integral de la competencia comunicativa.

Respecto al tercer objetivo específico, centrado en el diseño del instrumento, los resultados evidencian que este presenta una estructura coherente, con una secuencia lógica de componentes (léxico, pragmática, gramática, comprensión lectora, escucha y producción oral) y una adecuada correspondencia entre objetivos, habilidades e ítems. La congruencia interna fue valorada positivamente por los tres expertos, especialmente por el Experto 2, quien otorgó la máxima puntuación. Los patrones de respuesta observados en la prueba piloto confirmaron esta coherencia, al mostrar consistencia en el comportamiento de los estudiantes frente a ítems de características similares, así como una baja tasa de omisión y ausencia de patrones asociados al *rapid guessing* (Jin *et al.*, 2023; Ríos & Deng, 2022). En cuanto al componente léxico, se plantea la incorporación de elementos visuales y una mayor diversidad de tareas, con el fin de mejorar la accesibilidad cognitiva y la autenticidad de la evaluación. De acuerdo con Brown (2004) y Mayer (2009), el uso de apoyos visuales favorece la comprensión, especialmente en aprendices principiantes, y permite construir tareas más cercanas a situaciones reales de uso del lenguaje. En este sentido, se sugiere integrar imágenes, ítems multimodales y formatos variados que faciliten la interacción significativa de los estudiantes con el instrumento.

Con referencia al cuarto objetivo específico, encauzado a validar el instrumento, el análisis de los resultados de desempeño, la percepción de los estudiantes y la validación por expertos permitió construir un argumento de validez sólido con oportunidades de mejora. Las principales fortalezas identificadas fueron: alta congruencia interna, pertinencia general para el nivel A1 y validez técnica adecuada en el uso de formatos de ítem. Las principales

debilidades identificadas fueron: baja representatividad del componente léxico, falta de validez cultural por ausencia de contextualización colombiana auténtica, dependencia excesiva del vocabulario escrito en algunos ítems (ausencia de imágenes), calidad del audio mejorable y viabilidad cuestionable de la sección de *speaking*. Esta última oportunidad de mejora, el *speaking*, en ciertos contextos (especialmente el modelo Escuela Nueva, señalado por el Experto 3) se ve condicionada a la presencia de docentes capacitados para una adecuada administración. Además, como advierte Fulcher (2003), la evaluación de la producción oral (interacción evaluador-evaluado, condiciones de aplicación) puede verse afectada por la subjetividad si no se establecen criterios claros y estandarizados.

En otro aspecto, la Experta 1 mencionó la ausencia de elementos formales (*habeas data*); sin embargo, es importante aclarar que los representantes legales de los estudiantes firmaron el asentimiento oficial de la Universidad Santo Tomás.

Finalmente, los hallazgos de la investigación permiten señalar que el instrumento diseñado no solo cumple con criterios técnicos de evaluación, sino que también presenta una capacidad significativa para captar diferencias en los niveles de competencia de los estudiantes. Esta capacidad se evidenció en los patrones diferenciados de desempeño entre contextos rurales y urbanos, los cuales se interpretaron como una manifestación de la sensibilidad evaluativa del instrumento. Desde una perspectiva teórica, estos resultados respaldan la noción de validez como un proceso integrador, en el que convergen diferentes tipos de evidencia para sustentar las interpretaciones derivadas de los resultados (Messick, 1990; Kane, 2006). En este sentido, la coherencia entre el desempeño observado, la percepción de los estudiantes y la valoración de los expertos constituye un argumento sólido a favor de la validez del instrumento.

No obstante, la discusión también permite problematizar algunos aspectos relevantes. En especial, la tensión entre la pertinencia técnica del instrumento y su limitada contextualización cultural, señalada por el Experto 1 al afirmar que “*el contexto auténticamente colombiano está ausente*”, manifestó la importancia de la validez cultural en procesos de evaluación educativa. Aunque el instrumento logra discriminar niveles de competencia, lo hace desde referentes que no siempre reflejan la realidad sociocultural de los estudiantes (nombres genéricos como “Tom”, contextos como “Farmer’s Market”), lo cual puede afectar la interpretación de los resultados en contextos diversos. Esta limitación refleja un problema estructural en la educación del país, donde a menudo existen brechas entre el currículo oficial y la realidad educativa (BBE Languages, 2025). En esta tensión entre lo local y lo global, el concepto de glocalización propuesto por Canagarajah (2005) ofrece un marco teórico fundamental: en lugar de imponer metodologías o contenidos foráneos que ignoran el contexto del alumno, es necesario partir de lo local para llegar a lo global.

Asimismo, los resultados evidencian la importancia de distinguir entre claridad formal y demanda cognitiva (Tárraga-Mínguez & Gómez-Marí, 2023). Si bien los estudiantes comprendieron las instrucciones y los ítems, como lo evidencia la baja tasa de omisión y la alta consistencia en las respuestas, las dificultades observadas en tareas de mayor complejidad (por ejemplo, uso contextualizado del léxico) no pueden atribuirse a problemas de redacción, sino a las exigencias cognitivas propias del constructo evaluado. Esta distinción resulta fundamental para evitar interpretaciones erróneas sobre la calidad del instrumento y para orientar adecuadamente los procesos de enseñanza.

Por otro lado, la ausencia de patrones de respuesta asociados a comportamientos como el *rapid guessing* (Jin *et al.*, 2023; Wang & Xu, 2020) refuerza la validez interpretativa

de los resultados, al indicar que los estudiantes interactuaron con la prueba de manera comprometida y significativa. Este hallazgo es especialmente relevante en contextos educativos donde las condiciones de aplicación pueden influir en el desempeño de los estudiantes.

En conjunto, la discusión permite afirmar que el instrumento constituye una herramienta válida y pertinente para la evaluación diagnóstica, aunque su fortalecimiento requiere ajustes orientados a mejorar su representatividad, su adecuación contextual y su viabilidad administrativa en contextos diversos.

En síntesis, la presente investigación permitió diseñar y validar un instrumento de evaluación diagnóstica en inglés que responde a criterios técnicos, pedagógicos y contextuales relevantes para la educación básica primaria en Colombia. A través del análisis de evidencias se logró construir un argumento de validez sólido, que sustenta la pertinencia y funcionalidad del instrumento.

Si bien se identificaron aspectos susceptibles de mejora, especialmente en la representatividad del componente léxico, la contextualización cultural y la viabilidad de la sección de producción oral en ciertos contextos, los resultados evidencian que el instrumento constituye una herramienta valiosa para la evaluación del aprendizaje en contextos educativos diversos, contribuyendo tanto al campo de la investigación como a la práctica pedagógica. Los ajustes propuestos no son meramente formales, sino que son esenciales para garantizar que el instrumento cumpla con los más altos estándares de pertinencia, claridad y coherencia para la población objetivo.

La validez, entendida como un argumento en construcción, exige la incorporación continua de mejoras. En este sentido, este estudio no es un punto de llegada, sino un punto

de partida para futuras investigaciones que profundicen en la evaluación de lenguas extranjeras en contextos diversos, siempre con la mirada puesta en la equidad, la pertinencia cultural y la mejora de los aprendizajes de los estudiantes colombianos.

CONTRIBUCIONES A LOS AMBIENTES BILINGÜES DE APRENDIZAJE

ARQUITECTURA DE UN NUEVO BILINGÜISMO

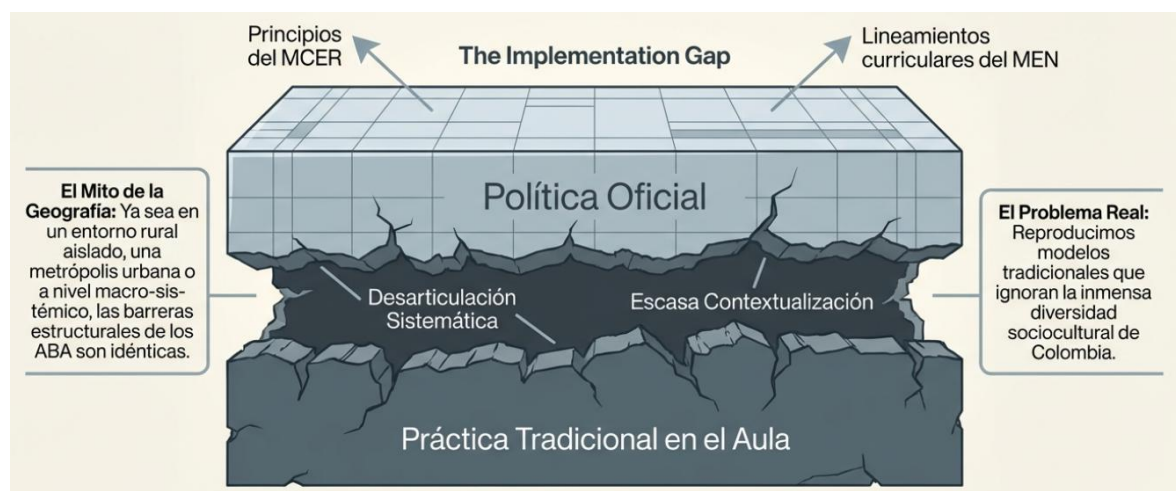
Transición de los Ambientes Bilingües de Aprendizaje (ABA) en Colombia: Del diagnóstico a la visión 2050.

La brecha de implementación es un desafío universal.

Uno de los mayores desafíos en el bilingüismo colombiano es la distancia entre la teoría y la realidad del aula. Existe una fractura denominada *The Implementation Gap* o brecha de implementación. Tradicionalmente se cree que los problemas varían según el territorio; sin embargo, los datos demuestran el “Mito de la Geografía”: ya sea en la ruralidad dispersa o en las grandes metrópolis, las barreras estructurales son idénticas. La política oficial y los lineamientos del MEN chocan contra prácticas tradicionales debido a una desarticulación sistemática y a una escasa contextualización que ignora nuestra diversidad.

Figura 1

La brecha de implementación



La estandarización invisibiliza la diversidad del estudiante

Los ABA actuales presentan avances, pero se estancan al operar bajo lógicas de fábrica que no responden al contexto vital del aprendiz. Para transformar el panorama, primero debemos asumir una postura sumamente autocrítica. Los ambientes bilingües actuales han logrado avances, pero continúan estancados bajo una “lógica de fábrica” que homogeniza los procesos. Esta rigidez se manifiesta en tres dimensiones críticas: evaluaciones punitivas que solo miden resultados uniformes, entornos escolares desarticulados de sus realidades inmediatas y una pedagogía unidireccional que anula la riqueza sociocultural de las regiones.

Figura 2

Postura Crítica



El diagnóstico estratégico fundamenta la transformación educativa

Frente a la inercia tradicional, la propuesta introduce el diagnóstico como el motor principal de cambio. El diagnóstico no debe verse como un simple trámite administrativo, sino como

una herramienta que aporta tres virtudes indispensables: da visibilidad a los niveles reales de desempeño de los estudiantes, otorga precisión para mapear las brechas de aprendizaje y brinda la pertinencia necesaria para edificar ambientes de aprendizaje que sean sostenibles en el tiempo.

Figura 3

Diagnóstico Basado en Evidencias



El diagnóstico como lente refractivo

Para que el diagnóstico funcione, necesitamos empoderar al cuerpo docente a través de la Literacidad en Evaluación o *Assessment Literacy*. Esta competencia pedagógica actúa como un lente refractivo: cambia la mentalidad del maestro para que deje de limitarse a “calificar” y comience a “decodificar” al estudiante. Al entender los datos e identificar las brechas y necesidades socioculturales reales, el diagnóstico se convierte en el puente directo hacia acciones pedagógicas verdaderamente transformadoras.

Figura 4

Lente refractivo

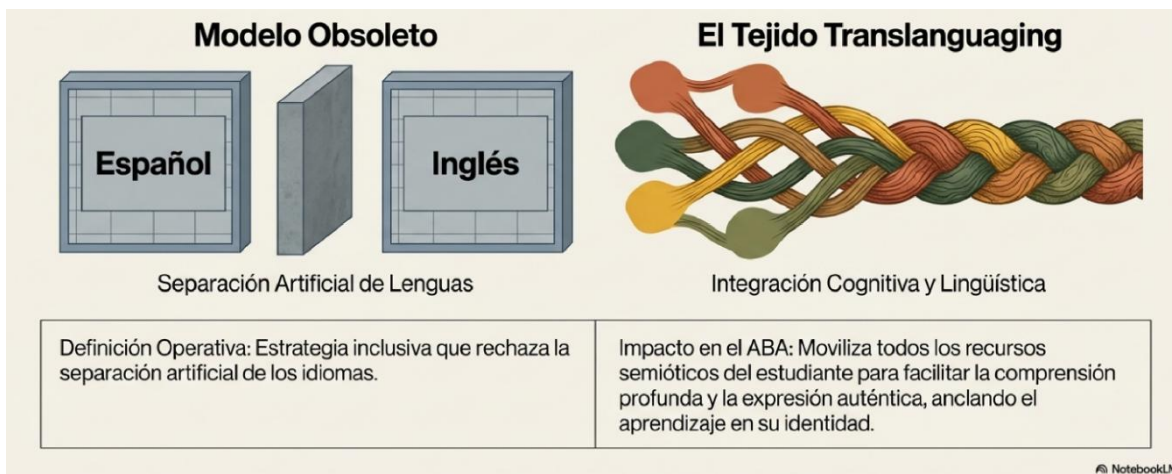


Translanguaging: Tejiendo el repertorio cognitivo completo

El enfoque lingüístico también requiere una revolución conceptual. Debemos dejar atrás el modelo obsoleto de la separación estricta y artificial de lenguas, que suele ver el uso de la lengua materna como un déficit. En su lugar, se propone el *translanguaging*. Esta es una estrategia inclusiva que reconoce que el cerebro del estudiante bilingüe no está dividido en compartimentos estancos; al contrario, integra de forma cognitiva y lingüística todos sus recursos semióticos para lograr una comprensión profunda, auténtica y respetuosa de su identidad.

Figura 5

Revolución conceptual



El Engranaje de la Evaluación Formativa Continua

¿Cómo se materializan estas ideas en el día a día escolar? A través de un sistema de acciones interconectadas que funcionan como un engranaje vivo. Este ciclo transformador se compone de cuatro pasos donde cada acción alimenta a la siguiente de manera orgánica: primero, implementando evaluaciones diagnósticas periódicas; segundo, alineando los instrumentos al MCER pero adaptados localmente; tercero, capacitando al docente en *assessment literacy*; y cuarto, integrando el *translanguaging* de forma cotidiana.

Figura 6

Engranaje vivo



Transición, de la evaluación punitiva a la formativa

Para sintetizar el cambio de paradigma, observemos esta matriz comparativa que contrasta el modelo tradicional frente al modelo transformado que proyectamos. Pasamos de una evaluación finalista y punitiva a una formativa y continua; de una práctica docente mecánica a un liderazgo ético; de la separación lingüística estricta al aprovechamiento del *translanguaging*; y de entornos rígidos a ecosistemas de aprendizaje flexibles y centrados en el bienestar.

Figura 7

Evaluación continua

Matriz de Transición ABA: De la rigidez a la adaptabilidad		
Dimensión	ABA Tradicional (Hoy)	ABA Transformado (2050)
Evaluación	Estandarizada, punitiva, finalista.	Diagnóstica, formativa, continua.
Práctica Docente	Ejecución mecánica de políticas desarticuladas.	Liderazgo ético guiado por Assessment Literacy.
Enfoque Lingüístico	Separación estricta de idiomas (Déficit).	Translanguaging como recurso integral.
Diseño del Entorno	Rígido, descontextualizado.	Flexible, híbrido, centrado en el bienestar.

Proyección 2050: Ecosistemas Híbridos

Mirando hacia el futuro, la prospectiva educativa nos exige abandonar de inmediato la inercia del pasado. Hacia el año 2050, los ABA no pueden limitarse a ser salones donde se instruye una lengua extranjera. Deben consolidarse como ecosistemas híbridos y flexibles en los que convergen armónicamente cuatro dimensiones esenciales para el estudiante: las competencias lingüísticas tradicionales, las competencias digitales, el desarrollo socioemocional y, como eje transversal, un enfoque de aprendizaje sostenible y bienestar integral.



Figura 8

Ecosistemas híbridos y flexibles

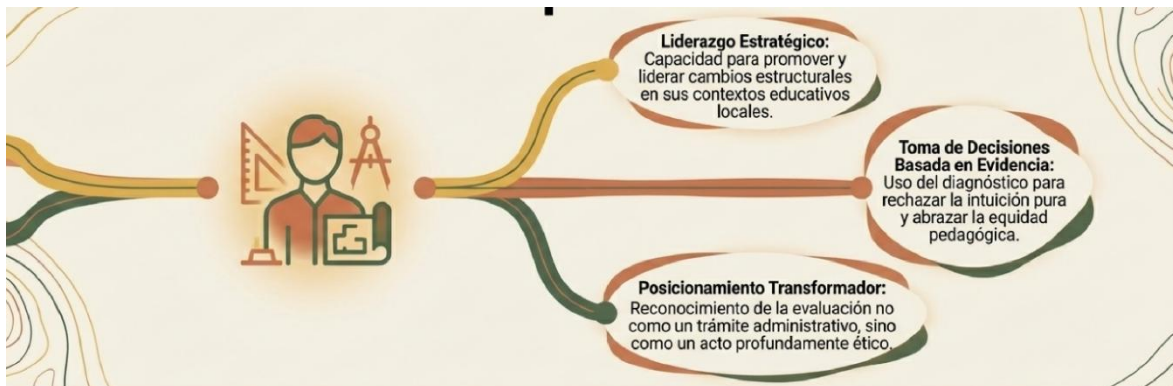


El Profesional Bilingüe como Arquitecto Ético

Esta profunda reestructuración requiere de un agente de cambio crucial: el educador. El profesional del área de inglés debe asumir una identidad renovada, transformándose en un arquitecto ético del aprendizaje. Esto demanda tres compromisos: ejercer un liderazgo estratégico para promover transformaciones estructurales en su entorno local, tomar decisiones estrictamente basadas en la evidencia recolectada mediante diagnósticos y adoptar una postura crítica que asuma la evaluación pedagógica como un acto de justicia y equidad social.

Figura 9

Innovación pedagógica



El Ecosistema ABA Equitativo

Para concluir, debemos interiorizar que la verdadera calidad educativa bilingüe no se alcanza simplemente incorporando herramientas tecnológicas de vanguardia sobre modelos pedagógicos obsoletos. La transformación integral requiere un equilibrio sistémico apoyado en una tríada indisoluble: la innovación pedagógica a través de modelos flexibles, un diagnóstico integral sustentado en la evidencia y una profunda contextualización que reconozca y respete activamente la diversidad sociocultural del territorio colombiano.

Figura 10

Transformación estructural



Punto de vista de las investigadoras

Myrian Yolanda Camacho Rodríguez. Contexto Rural.

1. Postura crítica

En los contextos rurales, los ABA evidencian una brecha estructural marcada por desigualdades en acceso, recursos y exposición al idioma. Aunque las políticas nacionales promueven el bilingüismo, en la práctica se implementan modelos estandarizados que desconocen las realidades territoriales. Esto genera una evaluación descontextualizada que mide resultados sin considerar condiciones de aprendizaje, lo cual contradice principios de equidad educativa.

2. Contribución

En este contexto, el diagnóstico adquiere un valor central como herramienta de justicia educativa. Permite identificar niveles reales de desempeño, condiciones sociolingüísticas y barreras de acceso. Desde esta perspectiva, el diagnóstico no solo mide, sino que orienta decisiones pedagógicas contextualizadas.

3. Acciones concretas

- Diseño de evaluaciones contextualizadas al entorno rural
- Uso de recursos offline (guías físicas, audios locales)
- Integración del *translanguaging* como puente lingüístico
- Evaluación formativa basada en procesos y no solo resultados

4. Proyección

Los ABA rurales hacia el 2050 deben evolucionar a modelos flexibles que integren tecnología accesible, saberes locales y aprendizaje situado. La evaluación deberá centrarse en el desarrollo integral del estudiante, incorporando dimensiones sociales, culturales y emocionales.

5. Posicionamiento profesional

El rol docente en contextos rurales implica una postura ética y crítica orientada a la equidad. El profesional debe actuar como mediador entre política educativa y realidad contextual, promoviendo prácticas inclusivas.

6. Conclusiones

La evaluación en los ABA rurales debe transformarse profundamente, priorizando la contextualización, la inclusión y la equidad como ejes fundamentales.

Carmen Natalia Guzmán Hernández. Contexto Urbano.

1. Postura crítica

En los contextos urbanos, aunque existe mayor acceso a tecnología, los ABA enfrentan el problema de la superficialidad pedagógica. La evaluación se centra en plataformas y resultados cuantitativos, sin garantizar aprendizajes significativos ni desarrollo comunicativo real.

2. Contribución

El diagnóstico en contextos urbanos permite evidenciar brechas ocultas, especialmente en habilidades como *speaking* y *listening*. Asimismo, facilita la toma de decisiones basadas en datos, optimizando el uso pedagógico de la tecnología.

3. Acciones concretas

- Implementación de evaluaciones multimodales
- Uso crítico de plataformas digitales
- Fortalecimiento de habilidades orales
- Integración de datos diagnósticos en la planeación

4. Proyección

Los ABA urbanos hacia 2050 deberán consolidar modelos híbridos donde la tecnología no sea un fin, sino un medio para el aprendizaje significativo. La evaluación se proyecta como un sistema dinámico basado en analítica de aprendizaje.

5. Posicionamiento profesional

El docente urbano debe asumir un rol innovador, capaz de transformar el uso de la tecnología en experiencias de aprendizaje significativas.

6. Conclusiones

El reto en contextos urbanos no es el acceso, sino el uso pedagógico efectivo de los recursos, lo cual exige una evaluación más crítica e integral.

Aida Vanessa González Pineda– Perspectiva Sistémica

1. Postura crítica

A nivel sistémico, los ABA evidencian una desconexión entre política educativa, currículo y práctica. La ausencia de sistemas de evaluación articulados limita la toma de decisiones estratégicas.

2. Contribución

El diagnóstico, desde una perspectiva macro, permite generar datos clave para la formulación de políticas educativas. Se convierte en un insumo fundamental para la mejora continua del sistema.

3. Acciones concretas

- Implementación de sistemas institucionales de diagnóstico
- Formación docente en evaluación
- Articulación entre editorial, escuela y política pública
- Uso de datos para la toma de decisiones

4. Proyección

Los ABA hacia 2050 deberán consolidarse como sistemas integrados donde la evaluación, la pedagogía y la gestión educativa funcionen de manera articulada.

5. Posicionamiento profesional

Desde la consultoría ELT, se asume un rol estratégico orientado a la toma de decisiones basadas en evidencia y la transformación del sistema educativo.

6. Conclusiones

La transformación de los ABA requiere una visión sistémica donde la evaluación sea el eje articulador del cambio educativo.

RECOMENDACIONES

A partir de los hallazgos obtenidos, las fortalezas identificadas y las oportunidades de mejora evidenciadas durante el proceso de diseño y validación del instrumento, se plantean una serie de recomendaciones orientadas tanto al perfeccionamiento del instrumento como al desarrollo de futuras investigaciones en el campo de la evaluación de lenguas extranjeras. Estas recomendaciones buscan contribuir al fortalecimiento de la validez, la pertinencia contextual y la viabilidad de la evaluación diagnóstica en inglés en educación básica primaria, atendiendo a las particularidades de los contextos educativos colombianos y a las exigencias contemporáneas de la evaluación del aprendizaje.

En primer lugar, se recomienda fortalecer la representatividad del constructo evaluado mediante la ampliación del componente léxico y la incorporación progresiva de habilidades lingüísticas actualmente subrepresentadas, particularmente *reading* y *writing*. Para ello, resulta pertinente diversificar los campos semánticos, equilibrar la distribución de habilidades y diseñar tareas de producción escrita acordes con el nivel A1 y con las características del estudiantado de básica primaria. Esta ampliación contribuiría a consolidar la validez de contenido del instrumento, favoreciendo una evaluación más integral de la competencia comunicativa.

En segundo lugar, se sugiere profundizar en la contextualización cultural del instrumento, integrando referentes auténticamente colombianos y cercanos a la experiencia sociocultural de los estudiantes. La incorporación de nombres, situaciones comunicativas, escenarios cotidianos y prácticas culturales propias del contexto educativo nacional permitiría reducir posibles sesgos interpretativos y fortalecer la validez cultural de la

evaluación. En este sentido, se recomienda adoptar perspectivas globales de diseño evaluativo que articulen los estándares internacionales de aprendizaje del inglés con las realidades locales del alumnado colombiano.

Asimismo, se recomienda optimizar la accesibilidad cognitiva y multimodalidad del instrumento, especialmente en aquellos componentes con alta dependencia del vocabulario escrito. La incorporación de apoyos visuales, imágenes, formatos de ítems variados y recursos multimodales puede favorecer la comprensión de las tareas sin alterar el constructo evaluado, particularmente en aprendices principiantes. Del mismo modo, conviene revisar y ajustar la complejidad léxica de instrucciones e ítems para disminuir cargas cognitivas innecesarias que puedan interferir con la medición efectiva de las competencias lingüísticas.

Por otra parte, se recomienda revisar la viabilidad administrativa y técnica de la sección de producción oral, especialmente en contextos educativos con limitaciones estructurales, como algunos modelos rurales o multigrado. Su implementación debería contemplar protocolos claros de aplicación, rúbricas estandarizadas y procesos de formación docente que reduzcan riesgos asociados a la subjetividad evaluativa. Igualmente, futuras versiones del instrumento podrían explorar alternativas tecnológicas o formatos adaptados que permitan evaluar la producción oral de manera más sostenible, objetiva y contextualizada.

Finalmente, se recomienda desarrollar investigaciones futuras orientadas a ampliar y robustecer el argumento de validez del instrumento. Entre las posibles líneas de investigación se encuentran la aplicación del instrumento en muestras más amplias y diversas, el análisis psicométrico de sus propiedades, estudios comparativos entre contextos rurales y urbanos, y la exploración longitudinal de su impacto en los procesos de enseñanza y aprendizaje del

inglés. Estos estudios permitirían consolidar la utilidad diagnóstica del instrumento y continuar avanzando hacia prácticas de evaluación más equitativas, pertinentes y sensibles a la diversidad educativa colombiana.



APÉNDICES

[Apéndice 1](#)

[Apéndice 2](#)

[Apéndice 3](#)

[Apéndice 4](#)

[Apéndice 5](#)



ANEXOS

[Anexo 1](#)

[Anexo 2](#)

[Anexo 3](#)

[Anexo 4](#)





REFERENCIAS

- Alderson, J. C., Brunfaut, T., & Harding, L. (2015). Towards a theory of diagnosis in second and foreign language assessment: Insights from professional practice across diverse fields. *Applied Linguistics*, 36(2), 236–260. <https://doi.org/10.1093/applin/amt046>
- Álvarez Huertas, L. R. (2024). Gerencia Educativa y Políticas Pedagógicas: Relación con el Rendimiento Académico. *Revista Publicando*, 11(43), 31-45. <https://doi.org/10.51528/rp.vol11.id2422>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2018). *Estándares para pruebas educativas y psicológicas* (M. Lieve, Trad.). Washington, DC: American Educational Research Association. (Trabajo original publicado en 2014)
- Bachman, L. F., & Palmer, A. S. (2010). *Language assessment in practice: Developing language assessments and justifying their use in the real world*. Oxford University Press.
- BBE Languages. (26 de junio de 2025). *Nivel de inglés en Colombia: Situación actual, demanda de profesionales bilingües y oportunidades de mejora*. <https://www.bbelanguages.com/nivel-de-ingles-en-colombia/>
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. Pearson Education.
- Canagarajah, A. S. (2005). *Reclaiming the Local in Language Policy and Practice*. Lawrence Erlbaum.



- Chapelle, C. A. (2020). *Argument-based validation in testing and assessment*. SAGE Publications.
- Charry, J. D., & Navarro Parra, S. (Eds.). (2018). *Manual de metodología de la investigación y principios en publicación científica* (1.^a ed.). Editorial Uninavarra. https://arbabpublishing.com/wp-content/uploads/2025/01/metodologia_de-investigacion.pdf
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment companion volume*. Council of Europe Publishing. <https://www.coe.int/en/web/common-european-framework-reference-languages>
- Díaz, C., Jara, A., Rosales, Y., & Sanhueza, M. (2021). Characterizing English Assessment Instruments: An Overview of Their Design. *Chakiñan, Revista De Ciencias Sociales Y Humanidades*, (15), 80-96. <https://doi.org/10.37135/chk.002.15.05>
- Fulcher, G. (2003). *Testing Second Language Speaking*. Pearson Education.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Testing*, 29(2), 113–132. <https://doi.org/10.1177/0265532211430341>
- Fulcher, G. (2015). *Re-examining language testing: A philosophical and social inquiry*. Routledge.
- Gaitán Medina, A. Y., Vargas Vargas, J. H., & Vargas Robles, J. G. (2023). *Análisis del Ava Accessit y su enfoque metodológico, para enseñanza del inglés en nivel B1*. Universidad Santo Tomás. <http://hdl.handle.net/11634/53673>
- Gao, J. (2022). An investigation of high-proficiency L2 English speakers' oral test performance: A profiling approach. *Frontiers in Communication*, 7, 926409. <https://doi.org/10.3389/fcomm.2022.926409>



- Gavilema Vaca, A. G., & Tovar Viera, V. R. (2025). Factores Psicológicos que Afectan el Mejoramiento de la Competencia Oral en Inglés. *Revista Scientific*, 10(35), 26-47. <https://doi.org/10.29394/Scientific.issn.2542-2987.2025.10.35.1.26-47>
- García-Manzanares, N., Luelmo del Castillo, M. J., Vinuesa-Benítez, V., & Izquierdo Sánchez-Migallón, E. (2025). Calidad formativa en AICLE: evaluación lingüística y metodológica de futuros docentes de Educación Física en contextos bilingües. *Retos*, 69, 1240-1252. <https://doi.org/10.47197/retos.v69.115996>
- Giraldo, F., Naranjo-Trujillo, D. E., & Ariza-Villa, J. A. (2023). *Desde el diseño de evaluaciones y hacia la literacidad en evaluación de lenguas*. *Folios*, (58), 126–139. <https://doi.org/10.17227/folios.58-16385>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista, P. (2022). *Metodología de la investigación* (7.ª ed.). McGraw-Hill.
- Institución Educativa Lucrecio Jaramillo Vélez. (2019). *Proyecto Educativo Institucional*. Medellín, Colombia. https://www.master2000.net/recursos/menu/45/9661/mper_185413_PEI%20LUCR ECIO%20JARAMILLO_compressed.pdf
- Instituto Colombiano para la Evaluación de la Educación – ICFES. (2023). *Pruebas Saber 3º, 5º, 7º y 9º*. Recuperado de <https://www.icfes.gov.co/evaluaciones-icfes/pruebas-saber-3-5-7-y-9/>
- Jin, Y., Zhang, L. J., & MacIntyre, P. D. (2021). A latent profile analysis of EFL learners' self-efficacy: Associations with academic emotions and language proficiency. *System*, 99, 102633. <https://doi.org/10.1016/j.system.2021.102633>



- Jin, K.-Y., Hsu, C.-L., Chiu, M. M., & Chen, P.-H. (2023). *Modeling rapid guessing behaviors in computer-based testlet items. Applied Psychological Measurement*, 47(1), 19–33. <https://doi.org/10.1177/01466216221125177>
- Kane, M. T. (2006). Validation. En Brennan, R. L. (Ed.), *Educational measurement* (4th ed., pp. 17-64). Westport, CT: American Council on Education - Praeger.
- Krippendorff, K. 2018. *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks, CA: Sage Publications.
- León Auris, M. E. (2024). Estrategias Didácticas en el Aprendizaje Significativo en Educación Básica: Didactic Strategies for Meaningful Learning in Basic Education. *Revista Scientific*, 9(33), 212–230. <https://doi.org/10.29394/Scientific.issn.2542-2987.2024.9.33.10.212-230>
- Leontjev, D., Poehner, M. & Huhta, A. (2025). Dynamic and Diagnostic Language Assessment: Learning across Frameworks to Support Second/Foreign Language Education. *De Gruyter Mouton*. <https://doi.org/10.1515/9783111233918>
- Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.
- Messick, S. (1990). Validity of Test Interpretation and Use. *Educational Testing Service*.
- Ministerio de Educación Nacional. (2006). *Estándares básicos de competencias en lenguas extranjeras: Inglés. Formar en lenguas extranjeras: ¡el reto! Lo que necesitamos saber y saber hacer*. https://www.mineducacion.gov.co/1780/articles-115174_archivo_pdf.pdf
- Ministerio de Educación Nacional (2014) Mallas de aprendizaje inglés https://eco.colombiaaprende.edu.co/wp-content/uploads/2021/10/Mallas-de-Aprendizaje_VF-1-1.pdf



Ministerio de Educación Nacional. (2016). *Derechos básicos de aprendizaje de inglés: Transición y primaria*.

https://www.colombiaaprende.edu.co/sites/default/files/files_public/2022-06/DBA-TRANSICI%C3%93N-Y-PRIMARIA_Ingl%C3%A9s-min.pdf

Ministerio de Educación Nacional. (2006). *Programa Nacional de Bilingüismo 2004–2019*.

Ministerio de Educación Nacional. https://www.mineducacion.gov.co/1759/articles-132560_recurso_pdf_programa_nacional_bilinguismo.pdf

Morales Llano, R. (2022). Política lingüística, bilingüismo y rol del maestro de inglés: Una revisión de la literatura científica desde el entorno global al contexto colombiano.

Colombian Applied Linguistics Journal, 24(2), 203–218.
<https://doi.org/10.14483/22487085/18130>

Negrete Raura, E.C., Pilaguano Umaginga, F. X., Gavilanes Lisintuña, G. E., & Escobar

Cueva, M. J. (2025). Análisis del proceso de validación de instrumentos de evaluación en la Educación General Básica. *Revista Ciencias De La Educación Y El Deporte*, 3(2), 190-204. <https://doi.org/10.70262/rced.v3i2.2025.121>

Ordoñez, M. E., & Rodríguez, J. (2017). *Análisis de las estrategias e instrumentos de evaluación en el aprendizaje del inglés como lengua extranjera*. Paideia, 22, 66.

Recuperado de <https://journalusco.edu.co/index.php/paideia/article/view/1336/2999>

Palomares Ruiz, A., Gracia Zomeño, A., Cebrián Martínez, A., & López Parra, E. (2023).

Diseño y validación de un instrumento de evaluación de los procesos de enseñanza-aprendizaje en el desarrollo de la competencia aprender a aprender “Innovaprende”.

Contextos Educativos. Revista De Educación, (31), 201–222.

<https://doi.org/10.18172/con.5534>





- Ríos, J. A., & Deng, J. (2022). Investigating the effect of differential rapid guessing on population invariance in equating. *Applied Psychological Measurement*, 46(7), 589–604. <https://doi.org/10.1177/01466216221108991>
- Robles Garrote, P., & Rojas, M. D. C. (2015). La validación por juicio de expertos: dos investigaciones cualitativas en Lingüística aplicada. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*. <https://www.nebrija.com/revista-linguistica/la-validacion-por-juicio-de-expertos-dos-investigaciones-cualitativas-en-linguistica-aplicada.html>
- Schorr, I., Plecher, DA., Eichhorn, C, and Klinker, G. (2024). Foreign language learning using augmented reality environments: A systematic review. *Frontiers in Virtual Reality*, 5, 1288824. <https://doi.org/10.3389/frvir.2024.1288824>
- Tárraga-Mínguez, R., & Gómez-Marí, I. (2023). La accesibilidad cognitiva en el punto de mira: Análisis de una experiencia de futuros docentes sobre la evaluación del entorno. *Aula de Encuentro*, 25(1), 109-130. <https://doi.org/10.17561/ae.v25n1.7253>
- Tique-Estrada, N. (2025). *Políticas educativas estandarizadas en la enseñanza del inglés en básica primaria: una revisión sistemática*. *Portal De La Ciencia*, 6 (1), 33–46. <https://doi.org/10.51247/pdlc.v4i3.395>
- Torres-Lara, K. L., Montes-Párraga, J. F., González-Barona, V. B., & Peñaherrera-Larenas, M. F. (2021). Técnicas e instrumentos de evaluación como herramienta para el cumplimiento de los resultados de aprendizaje. *Polo del Conocimiento*, 6(12), 776–785. <https://polodelconocimiento.com/ojs/index.php/es/article/view/3404>

Wang, C., & Xu, G. (2020). *A mixture hierarchical model for response times and response accuracy to identify rapid guessing behavior*. *Psychometrika*, 85(2), 377–402.

<https://doi.org/10.1007/s11336-020-09717-7>

Ye, X. (2024). Classroom environment and student engagement in EFL learning: A systematic review. *Frontiers in Education*, 9, 1415829.

<https://doi.org/10.3389/feduc.2024.1415829>