

**Validación por expertos de materiales de comprensión de lectura generados por
ChatGPT para aprendices de nivel B1 de inglés como lengua extranjera en Colombia**

Miguel Angel Castaño Díaz

202127522

Valeria Nogales Camacho

202123979

Orlando Chaves, PhD

Supervisor

UNIVERSIDAD DEL VALLE

ESCUELA DE CIENCIAS DEL LENGUAJE

DEPARTAMENTO DE LENGUAS Y CULTURAS EXTRANJERAS

LICENCIATURA EN LENGUAS EXTRANJERAS

ENERO 2026

Agradecimientos

Agradecemos, en primer lugar, a Dios por brindarnos la fortaleza, la sabiduría y la perseverancia necesarias para culminar satisfactoriamente esta etapa de formación académica. Expresamos también nuestro más sincero agradecimiento a nuestros padres, abuelos y a toda nuestra familia, por su apoyo constante, su confianza en nosotros y su amor incondicional. De igual manera, extendemos nuestro agradecimiento al profesor Orlando Chavez por su orientación, acompañamiento y valiosos aportes durante el desarrollo de este trabajo investigativo. Finalmente, agradecemos a los participantes de esta investigación por confiar en el proceso, por su disposición y por su valiosa colaboración, la cual fue fundamental para la culminación de este estudio.

Miguel Angel Castaño Díaz y Valeria Nogales Camacho

TABLA DE CONTENIDOS

1. INTRODUCCIÓN.....	9
2. REVISIÓN DE LA LITERATURA.....	13
2.1. Investigación Empírica.....	13
2.1.1. Estudios Internacionales.....	13
2.1.2. Estudios Nacionales.....	17
2.1.3. Estudios Locales.....	19
2.2. Marco Conceptual.....	22
2.2.1. Validación y validez.....	22
2.2.2. Diseño de materiales.....	26
2.2.3. Materiales auténticos.....	28
2.2.4. Comprensión lectora.....	29
2.2.5. La lectura en el nivel B1 de inglés.....	35
2.2.6. La cultura y su relevancia en el aprendizaje de lenguas.....	37
2.2.7. La inteligencia artificial en la educación.....	38
3. METODOLOGÍA.....	43
3.1. Enfoque y diseño.....	43
3.2. Participantes.....	45
3.3. Diseño del material.....	48
3.4. Métodos de recolección de datos y procedimientos.....	49
3.4.1. Métodos de recolección de datos.....	49
3.4.2. Procedimientos.....	52
3.5. Análisis de datos.....	53
3.5.1. Análisis de la rúbrica de validación.....	53
3.5.2. Análisis de las entrevistas.....	53
3.6. Validez y confiabilidad.....	57
3.7. Consideraciones éticas.....	58
4. ANÁLISIS DE RESULTADOS.....	60
4.1. Pertinencia pedagógica de materiales de lectura de ChatGPT para nivel B1-grado noveno.....	60
4.1.1. Resultados.....	60
4.1.2. Discusión.....	73
4.2. Alineación con los estándares de competencia lectora en inglés.....	78
4.2.1. Resultados.....	78
4.2.2. Discusión.....	83
4.3. Adecuación cultural al contexto colombiano.....	85
4.3.1. Resultados.....	85
4.3.2. Discusión.....	89
4.4. Fortalezas y limitaciones de ChatGPT en la elaboración de materiales.....	91
4.4.1. Resultados.....	91
4.4.2. Discusión.....	96
4.5. Implementación de los talleres.....	100
5. CONCLUSIONES.....	106
REFERENCIAS.....	110

ANEXOS.....	118
Anexo 1 Materiales de comprensión lectora generados por ChatGPT-4.....	119
Anexo 2 Prompts elaborados.....	140
Anexo 3 Formulario de consentimiento de participación.....	147
Anexo 4 Rúbrica de Validación.....	149
Anexo 5 Preguntas de la Entrevista.....	164

LISTA DE TABLAS

Tabla 3.1 Resumen de los métodos de recolección de datos.....	44
Tabla 3.2 Características de los participantes.....	46
Tabla 3.3 Experiencia y trayectoria de los participantes.....	47
Tabla 3.4 Categorías y subcategorías identificadas en las entrevistas.....	54
Tabla 4.1 Adecuación al nivel B1.....	61
Tabla 4.2 Variedad y calidad de los textos.....	62
Tabla 4.3 Variedad, calidad y precisión de los ejercicios.....	63
Tabla 4.4 Calidad de los niveles de comprensión lectora.....	65
Tabla 4.5 Coherencia con los objetivos pedagógicos de lectura.....	78
Tabla 4.6 Relevancia cultural y contextual.....	86

LISTA DE FIGURAS

Figura 4.1 Pregunta de Matching con ambigüedad.....	103
--	-----

RESUMEN

La aparición de herramientas de inteligencia artificial (IA), como ChatGPT, ha impulsado innovaciones significativas en la enseñanza de lenguas extranjeras, en particular en el área de diseño de materiales. Sin embargo, los materiales generados con estas herramientas requieren procesos de validación por parte de expertos para garantizar su adecuación pedagógica a contextos educativos específicos. De este modo, este estudio, adoptando un enfoque cualitativo, tiene como objetivo validar materiales de comprensión lectora generados con IA para estudiantes colombianos de nivel B1 mediante el juicio de expertos en el campo del inglés como lengua extranjera. Se generaron cuatro materiales de comprensión lectora con el apoyo de ChatGPT, que fueron validados por un panel de seis expertos con experiencia en educación universitaria y secundaria. Los materiales se diseñaron con base en los Derechos Básicos de Aprendizaje de Inglés (MEN, 2016), el Esquema Curricular Sugerido (MEN, 2016) y los lineamientos del MCER (2020) para el nivel B1. Se utilizó una rúbrica de validación, cuyos datos se analizaron descriptivamente y se llevaron a cabo entrevistas semiestructuradas, examinadas mediante análisis temático.

Los resultados mostraron que los expertos consideraron que los materiales estaban alineados con los estándares para el nivel B1 e incluían una variedad de ejercicios y temas relevantes. Igualmente reconocieron el potencial de la IA en los procesos de enseñanza y aprendizaje, especialmente para fomentar la autonomía en el estudiante, optimizar el tiempo del docente y ayudar en la creación de instrumentos de evaluación. No obstante, los expertos indicaron que los materiales tienen limitaciones relacionadas con su correspondencia al contexto colombiano, la disparidad entre el nivel B1 y la realidad escolar, la falta de recursos multimodales y posibles ambigüedades en ciertas preguntas. Asimismo, con respecto al uso de la IA, destacaron los riesgos asociados con la dependencia inmoderada de los estudiantes y las consideraciones éticas, lo que enfatiza la necesidad de una visión crítica entre los docentes en este contexto.

Palabras clave: Materiales de comprensión lectora para el inglés como lengua extranjera, Desarrollo y validación de materiales, Validación por juicio de expertos, ChatGPT, Nivel B1.

ABSTRACT

The emergence of artificial intelligence (AI) tools, such as ChatGPT, has driven significant innovations in foreign language teaching, particularly in the field of material design. However, materials generated with these tools require expert validation processes to ensure their pedagogical adequacy for specific educational contexts. Thus, this study, adopting a qualitative approach, aims to validate AI-generated reading comprehension materials for Colombian B1 level learners based on experts' judgements in the field of English as a Foreign Language (EFL). Four reading comprehension materials were generated with the support of ChatGPT, which were validated by a panel of six experts with experience in university and high school education. The materials were designed based on the *Derechos Básicos de Aprendizaje de Inglés* (MEN, 2016), the *Esquema Curricular Sugerido* (MEN, 2016) and the CEFR (2020) guidelines for level B1. A validation rubric was used and the data was analyzed descriptively, semi-structured interviews were also conducted and were examined using thematic analysis.

The results showed that the experts considered the materials to be aligned with the guidelines for level B1 and to include a variety of relevant exercises and topics. They also recognized the potential of AI as support for teaching and learning processes, especially in fostering learner autonomy, optimizing teacher's time, and assisting in the creation of assessment tools. However, the experts indicated that the materials have limitations related to their contextualization to the Colombian context, the gap between level B1 and the school reality, the lack of multimodal resources, and some possible ambiguities in certain questions. Additionally, regarding the use of AI, they highlighted risks associated with student overreliance and ethical considerations, which emphasizes the need for a critical stance among teachers in this context.

Keywords: EFL reading comprehension materials, Material development and validation, Validation by expert judgement, ChatGPT, B1 level proficiency.

RÉSUMÉ

L'émergence d'outils d'intelligence artificielle (IA), tels que ChatGPT, a suscité des innovations importantes dans l'enseignement des langues étrangères, en particulier dans le domaine de la création de matériel pédagogique. Cependant, les matériels générés à l'aide de ces outils nécessitent des processus de validation par des experts afin de garantir leur adéquation pédagogique à des contextes éducatifs spécifiques. Ainsi, cette étude, qui adopte une approche qualitative, vise à valider des matériels de compréhension de lecture générés par l'IA pour des étudiants colombiens de niveau B1, à partir des avis d'experts dans le domaine de l'anglais comme langue étrangère. Quatre matériels de compréhension de lecture ont été générés avec l'aide de ChatGPT et validés par un panel de six experts dotés d'une expérience dans l'enseignement universitaire et secondaire. Les matériels ont été conçus sur la base des *Derechos Básicos de Aprendizaje de Inglés* (MEN, 2016), de l'*Esquema Curricular Sugerido* (MEN, 2016) et des normes du CECR (2020) pour le niveau B1. Une grille de validation a été utilisée et les données ont été analysées de manière descriptive. Des entretiens semi-structurés ont également été menés et examinés à travers une analyse thématique.

Les résultats ont montré que les experts considéraient que les matériels étaient alignés avec les normes du niveau B1 et qu'ils comprenaient une variété d'exercices et de thèmes pertinents. Ils ont aussi reconnu le potentiel de l'IA comme soutien aux processus d'enseignement et d'apprentissage, notamment pour favoriser l'autonomie des apprenants, optimiser le temps des enseignants et aider à la création d'outils d'évaluation. Toutefois, les experts ont indiqué que les matériels présentent des limites liées à leur contextualisation dans le contexte colombien, à l'écart entre le niveau B1 et la réalité scolaire, au manque de ressources multimodales et à de possibles ambiguïtés dans certaines questions. De plus, en ce qui concerne l'utilisation de l'IA, ils ont souligné les risques liés à la dépendance excessive des élèves et les considérations éthiques, ce qui met en évidence la nécessité d'une attitude critique de la part des enseignants dans ce contexte.

Mots clés: Matériel de compréhension de lecture pour l'anglais comme langue étrangère, Développement et validation de matériel, Validation par avis d'experts, ChatGPT, Niveau B1.

1. INTRODUCCIÓN

La naturaleza social, política y cultural de la educación exige su adaptación continua a nuevas formas de instrucción. Para la enseñanza de lenguas extranjeras, en concreto, es importante adoptar estrategias innovadoras que fortalezcan las habilidades comunicativas que exige un mundo globalizado. Precisamente, el reciente interés por el uso de la Inteligencia Artificial (IA) ha comenzado a marcar una nueva etapa para el aprendizaje y la enseñanza en esta área. La incorporación de la IA en el aula no debe entenderse como una limitación, sino como un recurso pedagógico que optimiza la experiencia formativa (Kildiusheva, 2025).

El *chatbot* ChatGPT (*Chat Generative Pre-Trained Transformer*), en particular, se ha convertido en un recurso que puede potenciar los procesos de enseñanza-aprendizaje del inglés, específicamente, en lo que concierne a la producción de material didáctico. En efecto, diversos estudios han encontrado que los modelos de IA pueden ser útiles en la creación de recursos pedagógicos (Contreras & Pinzón, 2024; Haggag, 2021; Quintero, 2024; Sihite et al., 2023; Xiao et al., 2023), lo que muestra a esta herramienta como una posible aliada para el docente.

De acuerdo con Junguito y Chenoll (2024), la IA permite el diseño de materiales didácticos adaptados a las necesidades del aula, ya sean textos, audios o imágenes orientadas a contenidos específicos que el docente busca reforzar. Agregan, además, que este proceso de adaptación puede extenderse a la creación de actividades de comprensión y análisis crítico, que pueden ajustarse a los textos que los estudiantes estén abordando en un momento determinado. De manera similar, Murphy et al. (2023) expresan que ChatGPT, por ejemplo, tiene la capacidad de producir una gran cantidad de textos en una amplia diversidad de géneros, como la creación de guiones para *roleplays* o la elaboración de ensayos expositivos, quizzes o exámenes con preguntas de opción múltiple. Asimismo, añaden que es posible adaptar estos textos a distintos niveles de competencia lingüística y pueden ser ampliados o sintetizados según la preferencia del usuario. Lo anterior muestra que la IA, particularmente ChatGPT, puede funcionar como un apoyo para la creación de materiales que incluyan textos y ejercicios de comprensión lectora.

La lectura proporciona a los estudiantes de una lengua extranjera un aporte esencial para mejorar su dominio de la lengua meta (Al-Homoud & Schmitt, 2009). La lectura también provee un marco conceptual que permite a los estudiantes explorar la cultura, desarrollar autonomía y fortalecer sus habilidades interpretativas y analíticas (Ningsih, 2018; Zhao, 2024). Por ello, es necesario fortalecer las habilidades de comprensión lectora en la enseñanza de lenguas extranjeras. Esta competencia va más allá de la simple codificación de textos, ya que se trata de un proceso que implica comprender, analizar, criticar, evaluar y sintetizar fragmentos de información (Bharuthram, 2012). Dada la importancia de su desarrollo e instrucción, se requieren, por lo tanto, materiales

didácticos que contribuyan al fortalecimiento de estas habilidades. En ese sentido, la IA les ofrece y les facilita a los educadores posibilidades para producir dichos materiales, teniendo en cuenta también que la IA es capaz de generar actividades acordes a las necesidades particulares de un contexto en específico.

No obstante, el uso de la IA aún presenta desafíos en el proceso de elaboración de material didáctico. Entre las mayores preocupaciones en torno al uso de ChatGPT se encuentra el riesgo de que el texto generado presente sesgos o produzca información inexacta o engañosa, lo que podría derivar en consecuencias significativas dentro del contexto de aprendizaje de lenguas extranjeras (Murphy et al., 2023). De igual manera, la IA no siempre sigue las instrucciones con gran precisión, lo que hace necesario realizar ajustes y supervisar constantemente (Junguito & Chenoll, 2024). Por consiguiente, es fundamental que la elaboración y producción de estos textos y materiales sean sometidos a un proceso de validación, para que se compruebe si el uso de la IA de este modo resulta adecuado en el contexto educativo.

Si bien estudios internacionales han señalado el uso potencial de la IA en la educación y en la creación de materiales, nacionalmente existe una clara necesidad de procesos de validación precisos que permitan determinar la calidad y la pertinencia de los materiales, en especial los de comprensión lectora, generados por IA para distintos niveles lingüísticos y entornos educativos. La ausencia de participación de expertos, principalmente de docentes familiarizados con el contexto educativo colombiano, representa una brecha crítica, puesto que los materiales no validados podrían no responder a las necesidades curriculares y cognitivas de la población estudiantil. Asimismo, sin una evaluación adecuada, se encuentra el riesgo de que estos materiales refuercen imprecisiones o presenten deficiencias pedagógicas, lo que podría interferir en el aprendizaje y desarrollo lingüístico de los estudiantes en lugar de favorecerlo.

Este estudio, por lo tanto, toma en cuenta la participación de seis profesores expertos de educación bachiller y superior, quienes realizan la evaluación y revisión de materiales de comprensión lectora generados por ChatGPT-4. La validación de expertos, tanto escolares como universitarios, es elemental para garantizar que estos recursos cumplan con estándares teóricos y prácticos. En Colombia dichos estándares están retratados en Los Derechos Básicos de Aprendizaje del Inglés (2016) y en el Esquema Curricular Sugerido (2016) propuesto por el gobierno nacional. De igual forma, el Marco Común Europeo de Referencia para las lenguas (2020) constituye una guía central porque establece criterios claros y estandarizados sobre las competencias lingüísticas que debe alcanzar un aprendiz en cada nivel, incluyendo el B1 en inglés. Por ello, al momento de elaborar algún material se deben tener en cuenta estas directrices para asegurar que estos recursos se ajusten a ellas.

Además, considerando que en el Ministerio de Educación Nacional (2006) establece al nivel B1 en inglés como la meta de competencia mínima para los egresados de educación media y que exactamente en grado noveno se inicia el desarrollo de dichas competencias, se hace necesario la creación de materiales didácticos que puedan contribuir al progreso y fortalecimiento de habilidades lectoras en esta etapa, por lo que la IA puede convertirse en una herramienta pedagógica que permita la elaboración de estos recursos que se ajusten a los retos lingüísticos propios del nivel B1.

Con el fin de abordar la ausencia de este tipo de estudios a nivel nacional, la presente investigación busca validar la efectividad y pertinencia de materiales de comprensión lectora generados por IA para el nivel de competencia lingüística B1 en inglés. Concretamente, se evalúan materiales producidos por ChatGPT-4, teniendo en cuenta la perspectiva de los docentes con experiencia en el campo de la enseñanza del inglés como lengua extranjera. Es así que este estudio se propone dar respuesta a las siguientes preguntas de investigación:

1. ¿Qué tan pertinentes son, desde el punto de vista pedagógico, los materiales de lectura generados por ChatGPT para el desarrollo de habilidades de comprensión lectora en el nivel B1, en términos de vocabulario, sintaxis y adecuación del contenido, según los expertos en la enseñanza del inglés como lengua extranjera?
2. ¿Qué tan eficazmente se alinean los materiales generados por ChatGPT con los estándares de competencia lingüística establecidos para el nivel B1 en comprensión lectora?
3. ¿Qué tan culturalmente relevante es el contenido de los materiales generados por ChatGPT para el contexto educativo colombiano?
4. ¿Cuáles son las fortalezas y limitaciones de la IA en la generación de materiales de comprensión lectora para el nivel B1?

Teniendo en cuenta el creciente uso de la IA en el área de aprendizaje de lenguas extranjeras, la relevancia de este estudio es significativa, ya que aporta evidencia sobre la efectividad de esta herramienta para la creación de material didáctico, en particular, el de comprensión lectora. Igualmente, proporciona a los docentes unos parámetros claros para evaluar y seleccionar materiales que sean generados por IA, determinando su adecuación y relevancia pedagógica, lingüística y cultural. Los resultados de este estudio, de igual forma, pueden permitir que los docentes e instituciones educativas se informen y conozcan sobre la contribución que la IA puede ofrecer en el proceso de enseñanza-aprendizaje, considerando que estos materiales pueden personalizarse de acuerdo con los niveles de lengua extranjera y las necesidades de los estudiantes. Además, en contextos donde el acceso a materiales de calidad es limitado, los recursos generados por IA pueden ofrecer un importante apoyo pedagógico. En ese sentido, los hallazgos, obtenidos de

la aplicación de una rúbrica de validación y entrevistas semiestructuradas organizadas en categorías, pueden contribuir a que se comprendan los alcances y límites de la IA en la educación.

En los siguientes apartados se presenta la revisión de literatura, incluyendo los antecedentes, que muestran algunos estudios relevantes relacionados al uso de IA en la educación, y el marco conceptual, que aborda los constructos fundamentales para el análisis del estudio, tales como validación y validez, diseño de material y comprensión lectora. Posteriormente, se presenta la metodología del estudio que describe un enfoque cualitativo y un diseño descriptivo, así como las características de los participantes, los métodos de recolección de datos y procedimientos. En el capítulo siguiente se exhiben los resultados con sus respectivos análisis, dando respuesta a las preguntas de investigación. Finalmente, se exponen las conclusiones, las limitaciones del estudio y sugerencias para futuras investigaciones.

2. REVISIÓN DE LA LITERATURA

Este capítulo explora, en un principio, la investigación empírica que se ha realizado sobre el uso e integración de la IA en el campo de la educación en lenguas extranjeras, centrándose especialmente en la producción de recursos didácticos. Se revisan estudios previos a nivel global, nacional y local, remarcando sus aportaciones y los retos que puede conllevar la implementación de la IA en la elaboración de estos materiales. Posteriormente, se presenta el marco conceptual que sustenta el estudio, donde se incluyen conceptos importantes como validez y validación, diseño de materiales, comprensión lectora, la cultura como elemento relevante en la enseñanza de lenguas y la incorporación de la IA en la educación, e igualmente, se presenta los parámetros de habilidades lectoras para el nivel B1 en inglés de los documentos estandarizados que se tuvieron en cuenta.

2.1. Investigación Empírica

Esta sección ofrece una visión general de investigaciones realizadas en contextos internacionales, nacionales y locales que funcionan como referentes que demuestran la pertinencia de la temática y el enfoque adoptado, y justifica el diseño, metodología y procedimiento del presente estudio. Los estudios revisados emplearon enfoques cualitativos y mixtos, e involucraron a estudiantes y profesores de escuelas y universidades, proporcionando información sobre el nivel de efectividad en el uso de la IA en la enseñanza de lenguas. A nivel internacional, investigadores han explorado la implementación de herramientas tecnológicas, especialmente ChatGPT, en la educación, destacando su potencial para mejorar las habilidades lingüísticas (Andriyani et al., 2018; Haggag, 2021; Nieves et al., 2013; Quintero, 2024; Sihite et al., 2023; Xiao et al., 2023). A nivel nacional y local, estudios previos se han centrado en el diseño de planes y materiales didácticos de inglés como lengua extranjera eficaces, identificando elementos clave para su desarrollo (Aristizábal & Velásquez, 2016; Cifuentes & Niño, 2010; Contreras & Pinzón, 2024; Erazo, 2018; Gutiérrez & Jurado, 2018; Ojeda, 2018; Motato, 2020). Los estudios también resaltaron la validación de materiales y estrategias educativas por parte de expertos para garantizar su pertinencia y eficacia. En general, todos los estudios demuestran un creciente interés en el uso de la IA, especialmente en ChatGPT, para mejorar las prácticas educativas. De esta forma, se abordarán primero los estudios internacionales, seguidos de los estudios nacionales y locales.

2.1.1. Estudios Internacionales

Con respecto a los procesos de validación realizados por expertos, se revisaron las metodologías de estudios orientados a la evaluación y validación de materiales. En un primer caso, Nieves et al. (2013) realizaron un estudio mixto centrado en la validación de un prototipo de software diseñado para implementar el enfoque de Lectura Extensiva en inglés, dirigido a estudiantes de último año de la Educación Media de una Unidad Educativa Nacional en Venezuela.

Este estudio se efectuó con la participación de tres evaluadores externos, considerados expertos en áreas clave como la instrucción de lectura en inglés como lengua extranjera, la adquisición de vocabulario en una lengua extranjera y el diseño de materiales instruccionales. El prototipo se evaluó mediante un instrumento *ad hoc* dividido en tres partes: una escala de estimación de 20 aspectos relacionados con el diseño de la interfaz y la navegación; preguntas abiertas sobre el diseño, la idoneidad, la consistencia y la facilidad de uso del software; finalmente, preguntas que identificaron ventajas, desventajas y sugerencias para mejorar el prototipo. La valoración del prototipo fue positiva, destacando aspectos como la coherencia con los principios de la LE y el diseño de la interfaz. No obstante, se sugirieron mejoras en la redacción de las instrucciones y la navegación de los vínculos dentro de los textos. Los autores recomiendan hacer uso del prototipo en contextos reales para confirmar si cumple con los objetivos.

De igual manera, Andriyani et al. (2018) emplearon un diseño de investigación de método mixto para explorar el desarrollo y la validación de materiales didácticos para un curso de evaluación lingüística basado en un enfoque de aprendizaje basado en problemas (ABP). Participaron estudiantes de este curso y expertos en materiales en inglés y campos de evaluación, quienes revisaron los materiales desarrollados para garantizar su validez y relevancia. El estudio se desarrolló en varias etapas: primero, se recopilaban datos sobre las necesidades de los estudiantes mediante entrevistas y cuestionarios; luego, en un grupo focal, se crearon los materiales didácticos con la colaboración de los socios de investigación. Posteriormente, los materiales fueron validados por los expertos y se realizaron revisiones basadas en sus comentarios. Se evaluaron aspectos como contenido, coherencia con PBL, calidad didáctica, lenguaje, y diseño gráfico. La validez de constructo se realizó mediante análisis factorial y el porcentaje de evaluación de los materiales didácticos, y la validez de contenido se verificó mediante juicio racional y juicio de expertos. Los resultados indican que los materiales didácticos recibieron evaluaciones positivas y los aspectos evaluados cumplieron con los criterios de validez. Además, el análisis estadístico (KMO y Bartlett) confirmó la adecuación del instrumento y del material evaluado. Los autores recomiendan implementar este tipo de materiales en el campo, afirmando que ayudarían a fortalecer el pensamiento crítico de los estudiantes.

Centrándose ahora en la implementación de la IA en la educación, Haggag (2021) evaluó el efecto de un módulo de inglés basado en herramientas impulsadas por IA para el desarrollo de habilidades de lectura y escritura orientadas al TOEFL ITP. El estudio involucró a 21 estudiantes seleccionados aleatoriamente del Centro de Aprendizaje e Investigación de Idiomas de la Facultad de Educación de Hurgada. El módulo se centró en las secciones de lectura y escritura del examen TOEFL ITP y utilizó nueve herramientas basadas en IA, incluyendo Reader Pro, Andy, Elsa Speak, AI Grammar, Google Docs, Chat Bot, Google Lens, G-Translate y Google Assistant. Se empleó un

diseño cuasiexperimental de un grupo a través de un diseño pre-test y post-test. Los instrumentos de recolección de datos incluyeron cuestionarios para habilidades de lectura y escritura, un cuestionario sobre herramientas impulsadas por IA, un examen TOEFL ITP estandarizado y una rúbrica de escritura basada en las directrices de ETS. Los hallazgos indicaron diferencias estadísticamente significativas entre las pruebas previas y posteriores, con una mejora significativa en las puntuaciones medias de los participantes después de la implementación de las herramientas impulsadas por IA. El estudio concluyó que las herramientas de IA son efectivas para la preparación de exámenes en entornos de pruebas estandarizadas y recomendó integrar más tecnologías de IA en la preparación de exámenes como TOEFL-IBT y TOEFL-ITP.

En el campo del desarrollo de material didáctico, investigadores internacionales han explorado el uso de herramientas basadas en ChatGPT para generar materiales educativos que faciliten el aprendizaje del inglés como lengua extranjera. Sihite et al. (2023) examinaron la eficacia de ChatGPT 3.5 en la creación de preguntas de comprensión lectora, alineadas con las habilidades cognitivas de orden superior de la Taxonomía de Bloom, para estudiantes del Programa de Estudios de Educación en Inglés de la Universidad Al Washliyah de Medan. Mediante un enfoque de métodos mixtos, los autores realizaron un análisis de contenido cualitativo para evaluar los niveles cognitivos de las preguntas y emplearon métodos estadísticos para analizar las respuestas de los estudiantes. La recopilación de datos implicó la administración de preguntas generadas por IA y la calificación de las respuestas de los estudiantes. Los hallazgos revelaron que ChatGPT 3.5 generó preguntas que abarcaron los niveles cognitivos desde el análisis, la evaluación y la creación. Sin embargo, solo 10 de las 30 preguntas cumplieron los criterios de validez, lo que evidencia que la generación de preguntas aún requiere mejoras. El análisis de fiabilidad mediante el alfa de Cronbach indicó una consistencia interna moderada. El estudio concluye que, si bien las preguntas generadas por IA tienen potencial para las evaluaciones educativas, se necesitan mejoras en los modelos de IA para aumentar su validez y fiabilidad. Los autores recomiendan explorar el uso de la IA en el desarrollo de herramientas de evaluación, abordando diferentes tipos de preguntas e incorporando la retroalimentación de los estudiantes para optimizar el uso de la IA en la educación.

De un modo similar, Xiao et al. (2023) desarrollaron y evaluaron un sistema basado en Large Language Models (LLM), específicamente ChatGPT, para generar ejercicios personalizados de comprensión lectora para estudiantes de inglés de secundaria en China. Este estudio mixto evaluó la eficacia del sistema mediante un sistema de puntuación multidimensional donde evaluadores calificaron los materiales generados por IA y los escritos por personas. Participaron estudiantes, docentes y hablantes nativos, quienes proporcionaron retroalimentación mediante evaluaciones automatizadas y manuales. Los hallazgos mostraron que los materiales generados por IA a menudo superaban la calidad de los modelos de IA anteriores y, en algunos casos, superaban la

calidad de ejercicios escritos por personas. Los profesores participantes valoraron la capacidad del sistema para diversificar los recursos didácticos, destacando su idoneidad para el uso educativo. Sin embargo, se observaron limitaciones como en la creación de opciones de distracción eficaces para preguntas de opción múltiple, requiriendo mediación docente. El estudio concluyó que el sistema basado en ChatGPT y su integración en la educación pueden mejorar la calidad del material, reducir costos, apoyar el aprendizaje personalizado y ahorrar tiempo a los educadores. El estudio recomienda explorar el uso de la tecnología de inteligencia artificial e integrarla en los contextos educativos con el fin de fortalecer el aprendizaje de los estudiantes.

Siguiendo el enfoque en la creación de materiales generados por IA, Quintero (2024) evaluó la efectividad de ChatGPT para ayudar a los profesores de lenguas extranjeras a crear materiales para el aula de inglés como lengua extranjera. Este estudio cualitativo explora la contribución de ChatGPT en el proceso de creación de textos, su análisis y adaptación, de diseño de tareas o de elaboración de instrumentos de evaluación mediante la ingeniería de *prompts*. El estudio compara los productos y adaptaciones obtenidas por la IA a través del uso de los *prompts persona pattern*, *zero-shot*, *wish list* y *few-shot*. Los resultados evidenciaron que los distintos prompts fueron efectivos, aunque produjeron resultados ligeramente diferentes; el uso de cada uno puede depender de los objetivos y necesidades del docente. Quintero concluye que ChatGPT podía crear materiales adecuados para la enseñanza del inglés como lengua extranjera, con niveles de precisión y relevancia condicionados por la consigna, y que su uso no podría reemplazar el rol del profesor, quien sigue siendo responsable de adaptar las metodologías a las necesidades de los estudiantes. El estudio enfatiza la importancia de políticas educativas que brinden recursos y capacitación a los docentes para integrarlos en la exploración de estas herramientas tecnológicas.

Adicionalmente, un artículo reciente sobre el uso de ChatGPT para el desarrollo y la evaluación de materiales (Koraishi, 2023) ha demostrado que esta herramienta tiene el potencial de revolucionar la enseñanza del inglés como lengua extranjera. El autor, mediante una serie de indicaciones, destaca los beneficios de usar ChatGPT para crear recursos pedagógicos relevantes adaptados a las necesidades del alumnado y proporcionar una retroalimentación instantánea y personalizada para la evaluación de textos. Los docentes pueden solicitar modificaciones a sus planes de clase y la herramienta puede proporcionar versiones mejoradas basadas en las sugerencias. Sin embargo, Koraishi señala que, si bien ChatGPT puede realizar estas modificaciones, a menudo se requieren aclaraciones adicionales sobre ciertos detalles, lo que significa que la intervención del docente sigue siendo esencial para garantizar la eficacia del material. El autor concluye que la eficacia de la herramienta depende en gran medida de la participación activa del profesorado, quien debe guiar y supervisar el uso de la IA en el aula, y que

su integración exitosa en la educación requiere un enfoque colaborativo entre la tecnología y la pedagogía.

En resumen, los estudios internacionales muestran la importancia de integrar eficazmente ChatGPT en la educación, especialmente en el ámbito del inglés como lengua extranjera. Los autores comparten un enfoque en el uso de la inteligencia artificial para mejorar habilidades lingüísticas, como la comprensión lectora y el pensamiento crítico, demostrando cómo la IA puede ser útil para crear materiales para el aprendizaje del inglés como lengua extranjera que beneficien tanto a los docentes como a los estudiantes. Aún así, esto no implica que esta herramienta artificial cree materiales perfectos, dado que, a pesar de sus avances tecnológicos, la inteligencia artificial puede fallar en ciertos aspectos, como señalan los investigadores. Por ello, resulta pertinente validar los materiales generados por ChatGPT. Es evidente resaltar también que, si bien estos estudios se relacionan con la presente investigación, difieren en muchos aspectos como en el desarrollo del prompt para ChatGPT-4, el reconocimiento de los elementos que debe tener un material de comprensión lectora y el proceso de validación de expertos de educación bachiller y superior. De esta forma, el presente estudio busca cerrar la brecha existente sobre si ChatGPT puede llegar a ser una herramienta útil para el maestro e incluso estudiantes en la creación de materiales de comprensión lectora en el nivel de inglés B1.

2.1.2. Estudios Nacionales

A nivel nacional, se observa una falta de estudios actuales que se centren en el proceso de validación de materiales didácticos. Cifuentes y Niño (2010) llevaron a cabo una investigación de carácter cualitativo exploratorio, con el propósito de examinar los posibles criterios pedagógicos para el diseño de materiales didácticos digitales que sirvan de apoyo a la enseñanza de lenguas extranjeras. Los autores revisaron la bibliografía para desarrollar parámetros adecuados, y una vez establecidos, evaluaron dos materiales disponibles. Los resultados mostraron que estos materiales no se ajustaban plenamente a los criterios propuestos por los investigadores. Sin embargo, la diferencia entre los aspectos que se cumplían con los estándares y los que no era significativa. El estudio ofrece la posibilidad de seguir explorando el área de diseño de materiales digitales para la enseñanza de lenguas extranjeras, sugiriendo que se realicen investigaciones adicionales para aplicar los criterios establecidos y evaluar su eficacia en contextos educativos reales.

Por otra parte, abordando la optimización de los materiales de comprensión lectora en la enseñanza de lenguas extranjeras, Ojeda (2018) realizó un estudio cualitativo con un enfoque proyectivo, que consistió en la elaboración de un protocolo para el fortalecimiento de la comprensión lectora en inglés en alumnos de cuarto grado de una escuela bilingüe de Floridablanca. Los participantes fueron catorce estudiantes quienes tomaron un examen internacional

estandarizado, una prueba diagnóstica metacognitiva de lectura y, seguidamente, una encuesta de percepción. El estudio identificó factores que fortalecieron las habilidades de comprensión lectora (motivación por el tipo de texto, trabajo en equipo, organizadores gráficos), así como diversos obstáculos (reconocer la idea principal, organizar ideas y la autoevaluación). Estos hallazgos condujeron al desarrollo de estrategias de enseñanza, como el uso de conocimientos previos, apoyo visual, interacción con el texto, protocolo de preguntas y evaluación del uso y aprendizaje de las estrategias. Sobre esta base se definieron los componentes del protocolo, centrándose en la concentración, la autoevaluación, la motivación y la organización de ideas. El estudio concluyó que un diseño bien estructurado y adaptado al contexto colombiano podría mejorar significativamente la comprensión lectora. Entre las limitaciones se encuentran la dificultad de adaptar los estándares internacionales al contexto local y la falta de estudio previos que sirvan de guía para la creación de protocolos específicos.

En lo que concierne al desarrollo de materiales pedagógicos, Contreras y Pinzón (2024) llevaron a cabo una investigación-acción para implementar una secuencia didáctica sobre la historia del Antiguo Egipto, diseñada con IA y dirigida a estudiantes de sexto grado de una institución educativa en la ciudad de Cartagena de Indias. El estudio incluyó observación participante, un diario de campo, entrevistas semiestructuradas con docentes y alumnos, encuestas y listas de cotejo. Se desarrollaron diversos materiales didácticos basados en recursos generados mediante IA, tales como cuentos ilustrados, juegos de cartas y presentaciones interactivas. Los resultados mostraron que estos recursos tuvieron una recepción positiva por parte de los estudiantes y contribuyeron a mejorar su comprensión y retención de los conceptos históricos enseñados. La integración de estos materiales en el aula permitió una experiencia de aprendizaje más dinámica y participativa, fomentando el interés de los estudiantes por la materia. Aún así, una limitación de este estudio radica en que los materiales no fueron validados por expertos en el campo de creación de material o enseñanza de historia, lo que podría haber garantizado una evaluación más precisa de su eficacia.

En general, aunque existe un interés por innovar en el desarrollo de materiales en distintas áreas, incluida la enseñanza del inglés, es fundamental asegurar que estos materiales se ajusten a los estándares necesarios y estén alineados con las normativas pedagógicas que respaldan su uso. Por un lado, algunas investigaciones exponen que los materiales diseñados con herramientas tecnológicas innovadoras, como la IA, pueden captar el interés de los estudiantes y afianzar su aprendizaje. Por otro lado, otros estudios destacan la importancia de establecer criterios pedagógicos claros y estrategias centradas en las necesidades específicas de los estudiantes. Asimismo, es importante mencionar que la falta de procesos de validación, como se evidencia en algunos casos, limita la capacidad de medir el impacto real de estos materiales en diferentes contextos educativos. Por lo tanto, es esencial que los esfuerzos de innovación no solo se enfoquen

en el desarrollo, sino también en procesos de evaluación que garanticen la calidad, pertinencia y eficacia de los materiales en el aula.

2.1.3. Estudios Locales

A nivel local, no se encontraron muchos estudios relacionados con la validación de material didáctico, lo que evidencia la escasez de investigaciones sobre cómo garantizar la eficacia y calidad de los recursos educativos en ese contexto. Gutiérrez y Jurado (2018) desarrollaron un estudio dirigido a identificar y analizar los posibles errores en la elaboración de pruebas de opción múltiple de respuesta única, enfocándose en cinco tipos de errores que afectan a su validez. Este estudio, igualmente, pretendía contribuir a la cualificación de los futuros creadores de ítems y destacar la importancia de la revisión colaborativa en la creación de pruebas de comprensión lectora. El proceso contó con la participación de un equipo de tres profesionales con formación en literatura, psicología y lingüística, que evaluaron individualmente los ítems de un corpus obtenido del pilotaje de una prueba de comprensión lectora del Sistema de Evaluación de Aprendizajes de la Universidad Autónoma de Occidente. Se diseñó una rejilla de evaluación para el análisis de los ítems de opción múltiple, en la que cada evaluador emitía un juicio sobre cada ítem, que podía ser *“Aceptado”*, *“Aceptado con correcciones”* o *“Rechazado”*. Los resultados se discutían en grupo hasta alcanzar un consenso. Si dos de los tres evaluadores rechazaban un ítem, este se eliminaba. De un total de 270 rejillas evaluadas, se encontró una discrepancia en el 39% de los ítems. Se identificaron cinco tipos de errores: ambigüedad semántica, contenido léxico ambiguo, fiabilidad mínima, inferencias originadas en el saber del constructor y relaciones lógicas evidentes. Estos errores afectan a la calidad de las pruebas y cuestionan su validez.

Por otro lado, existen diversos estudios sobre intervenciones pedagógicas con énfasis en mejorar las habilidades y estrategias de lectura, tanto para estudiantes escolares como universitarios en Cali, realizados por alumnos de la Universidad del Valle (Aristizábal & Velásquez, 2016; Erazo, 2018; Motato, 2020). Aristizábal y Velásquez (2016) realizaron un estudio para fortalecer la comprensión lectora en inglés en los estudiantes de décimo grado del colegio bilingüe Montessori de Cali. La investigación abordó la falta de textos literarios en las clases de inglés y los bajos niveles de comprensión de los estudiantes al trabajar con dichos textos. Para explorar esta cuestión, los investigadores diseñaron, implementaron y evaluaron una secuencia didáctica utilizando cuentos cortos para mejorar la habilidad lectora. Se trató de un estudio de caso, comenzando con un análisis de campo para caracterizar al grupo y hallar la problemática. Seguidamente, los investigadores desarrollaron y aplicaron la secuencia didáctica, asegurándose de que se basara en los conceptos teóricos esenciales. Se utilizaron diversos instrumentos de recolección de datos, entre ellos encuestas, entrevistas a docentes, diarios de campo, listas de control, cuestionarios de autoevaluación y pruebas de diagnóstico previas y posteriores. Los resultados demostraron que esta

secuencia con la implementación de cuentos cortos mejoró significativamente las habilidades de comprensión lectora en inglés del alumnado. El estudio concluyó que la incorporación de textos literarios en el plan de estudios no solo fortalece las habilidades lingüísticas, sino que también fomenta la apreciación cultural y la motivación entre los estudiantes. Se recomienda incluir, por lo tanto, textos literarios en la enseñanza del inglés para crear experiencias de aprendizaje más estimulantes.

De manera similar, Erazo (2018) llevó a cabo un proyecto de investigación orientado a diseñar una secuencia didáctica de comprensión lectora para la enseñanza de inglés con fines específicos a estudiantes de primer año de medicina de la Universidad del Valle. El estudio, realizado entre febrero y octubre del 2017, empleó una metodología de investigación-acción. Los datos se recopilaban mediante observaciones en el aula, encuestas, entrevistas y pruebas de diagnóstico de entrada y de salida. El estudio incluyó la implementación y evaluación de la secuencia didáctica que incorporaba los enfoques de Aprendizaje Integrado de Contenidos y Lenguas Extranjeras (AICLE) y Enseñanza para la Diversidad. Los hallazgos revelaron que conocer las necesidades de los estudiantes, reforzar los conocimientos de gramática inglesa y analizar las funciones retóricas de textos auténticos mejoraba significativamente la comprensión lectora de los estudiantes en una lengua extranjera. La identificación y el uso eficaz de estrategias de lectura demostraron que una secuencia didáctica bien diseñada que aborde las funciones retóricas puede fortalecer la comprensión lectora. Además, cuando se les proporciona a los alumnos las herramientas necesarias para superar los retos lingüísticos y afectivos, su rendimiento académico en lectura muestra mejoras. Este estudio destaca la importancia de alinear el diseño pedagógico con las necesidades académicas de la población estudiantil para fomentar un aprendizaje significativo y recomienda seguir explorando diferentes enfoques para mejorar las habilidades lingüísticas en contextos específicos.

Un estudio más reciente, realizado por Motato (2020), se centró en promover la lectura en una lengua extranjera como base para desarrollar las habilidades comunicativas entre los estudiantes de cuarto y quinto grado de una escuela pública de Cali. La investigación implementó una secuencia didáctica acorde con el enfoque basado en tareas, integrando ciencias sociales, tecnologías de la información y comunicación (TIC), lecturas, libros ilustrados y trabajos en grupo. Esta investigación-acción consistió en diagnosticar las necesidades de los estudiantes, diseñar una unidad didáctica y evaluar continuamente sus efectos. Los instrumentos de recolección de datos incluyeron encuestas durante la fase de diagnóstico, observaciones de las clases, entrevistas y grupos focales. Los resultados mostraron un impacto positivo en el aprendizaje del inglés en los estudiantes y un cambio en su percepción de la lengua. La integración de las TIC mejoró la interacción y participación en el aula. Se concluyó que la combinación del aprendizaje basado en

tareas con recursos interdisciplinarios impulsó un aprendizaje más enriquecedor. Las recomendaciones incluyeron una mayor exploración del uso de las TIC en el apoyo al desarrollo de la lengua y la adaptación de materiales a distintos contextos educativos. Entre las limitaciones están el tamaño reducido de la muestra y los retos que plantea la implementación de la tecnología en el entorno de una escuela pública.

En resumen, se recomienda la integración de la tecnología y estrategias interdisciplinarias para satisfacer las diversas necesidades educativas. Los investigadores se centran en fortalecer la comprensión lectora mediante intervenciones didácticas creadas para responder a las necesidades específicas de los estudiantes, ya sea utilizando textos literarios, materiales auténticos o pruebas estandarizadas. Remarcan la importancia de abordar factores cognitivos y afectivos, validar la calidad de los materiales y emplear enfoques innovadores. La investigación de Gutiérrez y Jurado (2018), especialmente, se destaca por llevar a cabo un proceso de validación de los ítems de una prueba, involucrando a diferentes expertos.

En conjunto, todos los estudios anteriores, tanto internacionales como nacionales y locales, proporcionan información y aportan perspectivas sobre la creación y personalización de *prompts* para ChatGPT en el diseño de materiales de lectura para la enseñanza de lenguas extranjeras. Entre los elementos principales se encuentran la extensión, el uso de historias, nivel de dificultad y los temas, así como el diseño de las actividades, en cuanto al número, tipos y cantidad de opciones por pregunta, esto como indicaciones que pueden servir para el prompt en ChatGPT (Contreras & Pinzón, 2024; Quintero, 2024; Sihite et al., 2023; Xiao et al., 2023). Estos aspectos permiten un enfoque más estructurado para los materiales generados por IA, asegurándose de que sean útiles para los alumnos en caso de implementación. Asimismo, los estudios describen los factores a tener en cuenta al momento de diseñar materiales de comprensión lectora, como la activación de conocimientos previos, el uso de recursos visuales, la identificación de ideas principales, la organización de la información, el uso de estrategias de lectura y la adaptación de los materiales al contexto (Aristizábal & Velásquez, 2016; Contreras & Pinzón, 2024; Erazo, 2018; Motato, 2020; Ojeda, 2018).

Este último punto revela la importancia de considerar los documentos oficiales relativos a la enseñanza del inglés en Colombia, tales como los Derechos Básicos de Aprendizaje de Inglés (2016) y el Esquema Curricular Sugerido (2016), ya que existe una deficiencia en cuanto a su implementación en el diseño de materiales. Al adaptar los materiales al contexto colombiano, los educadores pueden garantizar su pertinencia y eficacia para cumplir con los estándares pedagógicos nacionales. Por último, algunos estudios enfatizan la necesidad de validar los materiales con la

opinión y juicio de expertos. La evaluación de los expertos brinda información valiosa sobre el nivel de eficacia de los materiales, asegurando que estos se ajusten a los parámetros de calidad y alcancen los objetivos de aprendizaje previstos (Andriyani et al., 2018; Gutiérrez & Jurado, 2018; Nieves et al., 2013). Por lo tanto, es indispensable incluir este proceso de validación, reconociendo la experiencia de los expertos tanto en los aspectos teóricos como prácticos del campo.

A continuación, se presentan los constructos y conceptos que ayudaron a sustentar el análisis de los resultados.

2.2. Marco Conceptual

Anteriormente, se revisó la literatura sobre el uso de la IA, especialmente ChatGPT, en contextos educativos, aspectos a tener en cuenta para el diseño de materiales de comprensión lectora y validación de materiales con la opinión y el juicio de expertos. Si bien existen investigaciones sobre temas relacionados al estudio, existe una notable escasez de estudios que busquen validar los materiales de comprensión lectora con inteligencia artificial, específicamente ChatGPT. Por ello, este apartado busca comprender los componentes de un material de comprensión lectora nivel B1 de inglés, con el fin de indicarle a ChatGPT, mediante el prompt, toda la información necesaria. Posteriormente, expertos en la enseñanza del inglés como segunda lengua podrán validar si el material cumple con todos los requisitos para considerarse un buen material aplicable al contexto colombiano. De esta forma, el presente capítulo analiza los conceptos que guían el estudio, incluyendo la validez y validación, el diseño de materiales y la comprensión lectora. Asimismo, se analiza la cultura como elemento relevante en la enseñanza de lenguas, la incorporación de la IA en la educación y los parámetros en la habilidad de lectura para el nivel B1 en inglés según los documentos oficiales (el Marco Común Europeo de Referencia para las lenguas, el Esquema Curricular Sugerido y los Derechos Básicos de Aprendizaje) que se tuvieron en cuenta.

2.2.1. Validación y validez

Aunque esta investigación se enfoca en la validación de materiales didácticos y no en la de pruebas o exámenes, la definición de validación asociada a la evaluación de estos últimos puede ser utilizada y adaptada al diseño de materiales para el aprendizaje de lenguas extranjeras. Tanto la validación de materiales como la validación de pruebas requieren una verificación rigurosa de la fiabilidad, validez y pertinencia del recurso para garantizar que se concreten y efectúen los objetivos de aprendizaje. En ese sentido, los materiales deben diseñarse justamente para permitirle a los estudiantes el desarrollo de habilidades de comprensión lectora, y además, se debe asegurar que dichos materiales sean relevantes y estén alineados con las necesidades educativas.

La validación de pruebas es el proceso de generar evidencia que respalde la solidez de las inferencias que se realizan sobre determinados rasgos a partir de los resultados

obtenidos en un test, es decir, la evaluación debe esencialmente centrarse en una validez basada en la evidencia (Weir, 2005, p.1).

Weir asegura además que el proceso de validación es la relación establecida entre el método (el cómo) y los constructos (el qué), lo que significa que la validación garantiza una unión clara y respaldada por evidencias entre el instrumento de evaluación y las habilidades lingüísticas que se van a medir. De la misma manera, de acuerdo con AERA et al. (2014) el proceso de validación implica la acumulación de evidencia relevante que brinde un apoyo sólido a las interpretaciones de los resultados obtenidos en una prueba. Por su parte, Arribas (2004) define a la validez como “el grado en que un instrumento de medida mide aquello que realmente pretende medir o sirve para el propósito para el que ha sido construido” (p. 28). Adicionalmente, Sireci (2007) argumenta que la validez no es una evaluación única y definitiva, sino un proceso que debe revisarse y actualizarse continuamente. Es así que la validez se entiende como un indicador dinámico que asegura la calidad y fiabilidad de un determinado instrumento de evaluación, lo que contribuye a la optimización de estos recursos en el ámbito educativo.

De acuerdo con lo anterior, en este estudio la validez se entiende, por lo tanto, al grado en que los materiales de comprensión lectora generados por IA se alinean con los propósitos pedagógicos que se proponen cumplir. La revisión realizada por varios expertos constituye la evidencia que respalda la validez de los materiales, por ello es necesario definir el concepto de validez basada en contenido y el juicio de expertos.

2.2.1.1. Validez basada en contenido

Los estándares de AERA et al. (2014) proponen cinco fuentes de evidencia de validez que son *evidencia basada en el contenido*, *evidencia basada en los procesos de respuesta*, *evidencia basada en la estructura interna*, *evidencia basada en las relaciones con otras variables* y *evidencia de la validez y consecuencia de la prueba* (p.14-21). El marco propuesto se considera importante para este estudio, ya que, aunque se centra principalmente en la validación de materiales didácticos, de igual forma podría apoyar en la evaluación del aprendizaje de los estudiantes. Se puede afirmar que los materiales también son capaces de medir las competencias de los estudiantes, tanto de manera formativa como sumativa, lo que certifica su efectividad en el seguimiento del proceso de aprendizaje.

Asimismo, este estudio se enfoca exclusivamente en la evidencia basada en el contenido, debido a que se refiere al grado en que el contenido de una prueba es coherente con su finalidad y representa los conocimientos, competencias o habilidades que se buscan medir (Arias & Sireci, 2021). Según Sireci (1998) la validez basada en contenido se sustenta en tres aspectos:

Definición del dominio: se refiere a la definición operativa del contenido que abarca la prueba y suele representarse mediante una tabla bidimensional de especificaciones del test. Una dimensión describe las áreas de contenido y la otra dimensión cubre las áreas cognitivas.

Evaluar la pertinencia de esta definición implicaría cuestionar si otras áreas de contenido o cognitivas también pueden ser relevantes, si algunas de las áreas enumeradas resultan ser superfluas y si las proporciones relativas de ítems asignadas a cada área de contenido-cognitiva son apropiadas (Sireci, 1998, p. 300).

Representación del dominio: se centra en la correspondencia entre el test y su contenido y sus especificaciones cognitivas. Se evalúa juzgando en qué medida los ítems de la prueba reflejan el contenido y las habilidades descritas en la definición del dominio.

Relevancia del dominio: se analiza si cada ítem del test es relevante para el dominio definido.

Generalmente, la representación y relevancia del dominio se enfocan más en el test que en sus especificaciones y también se evalúan a través del juicio de expertos del contenido, quienes comparan los conocimientos y las habilidades que cada ítem pretende medir con las especificaciones del test.

Akbari (2012), igualmente, argumenta que la validez de contenido “se ocupa de la adecuación de un test en términos de su contenido o de la muestra de comportamiento que pretende medir” (p. 30), en otras palabras, la validez de contenido busca procurar que la prueba aborde lo que realmente necesite evaluarse. Akbari igualmente afirma que cuando se trata de diseñar una prueba de competencia lingüística, utilizar este enfoque puede resultar más complicado, debido a que la competencia es un concepto más abstracto y no tiene un dominio de contenido fijo del que extraer preguntas. Esto presenta el riesgo de “subrepresentación del constructo”, lo que significa que la prueba puede no incluir todos los aspectos importantes que deberían evaluarse.

En el caso de materiales de comprensión de lectura generados por IA, es importante asegurar que el contenido aborde adecuadamente el dominio de conocimientos necesarios para desarrollar habilidades de comprensión lectora. El reto de los materiales elaborados mediante IA consiste en que, al estar diseñados automáticamente, como señala Akbari, es posible que no incluyan todos los aspectos relevantes que deberían evaluarse. Esto puede pasar si la IA no entiende completamente los objetivos pedagógicos o las necesidades específicas del contexto educativo. Por ejemplo, los materiales pueden no representar adecuadamente las áreas de contenido o las áreas de habilidades cognitivas esenciales para la comprensión lectora. Por lo tanto, es fundamental verificar la validez de este tipo de materiales para certificar que se cubran los contenidos y habilidades centrales para los estudiantes y que sean también coherentes con los objetivos del plan de estudios.

2.2.1.2. Juicio de expertos

Arribas (2004) señala que en la validez de contenido el cuestionario es sometido a la valoración de expertos, quienes juzgan si este permite evaluar todas las dimensiones que se pretenden medir y, por lo tanto, se recurre solamente a las valoraciones cualitativas emitidas por dichos expertos. Carley et al. (2018) definen a un experto como un individuo que posee conocimiento extenso y autoritario en un campo específico, que no es comúnmente comprendido por la población general. De igual modo, Escobar-Pérez y Cuervo-Martínez (2008) expresan que “el juicio de expertos se define como una opinión informada de personas con trayectoria en el tema, que son reconocidas por otros como expertos cualificados en éste, y que pueden dar información, evidencia, juicios y valoraciones” (p. 29).

Además, Escobar-Pérez y Cuervo-Martínez (2008) describen los siguientes pasos que pueden facilitar la organización de la información para que el juicio de expertos sea más eficiente:

- El juicio debe tener un propósito, ya que puede haber diferentes objetivos, como establecer la equivalencia semántica de una prueba que está validada en otra lengua, evaluar la adaptación cultural de los ítems para asegurar que midan lo mismo en contextos culturales diferentes o validar el contenido de un test creado por un grupo de investigadores.
- Los jueces deben ser seleccionados teniendo en cuenta su formación académica, su experiencia en el área y su reconocimiento en la comunidad.
- Es esencial especificar las dimensiones e indicadores que cada ítem del test está midiendo, debido a que esto permite a los expertos evaluar su relevancia, suficiencia y pertinencia. No debe darse por sentado que el experto puede identificar el constructo adecuadamente con solo una descripción general, ya que el mismo constructo puede tener diferentes definiciones dependiendo del contexto.
- Es necesario proporcionar a los expertos información detallada sobre el propósito del test y el uso que se les dará a los resultados obtenidos. Esto ayuda a contextualizar la prueba para los expertos.
- Se pueden establecer estimaciones diferenciales para las dimensiones del test con el fin de asignar diferentes niveles de valor a las distintas áreas o aspectos a evaluar según su relevancia para el objetivo específico de la prueba en cuestión. Esta práctica es utilizada cuando algunas dimensiones son más significativas que otras dependiendo de la situación en la que el test se lleva a cabo.

- Las plantillas de evaluación deben diseñarse y estructurarse de acuerdo a los objetivos de la evaluación.
- Finalmente, se elaboran las conclusiones del juicio de expertos para describir las características psicométricas de la prueba.

Estos criterios son fundamentales en el presente estudio al proporcionar un enfoque estructurado y sistemático del proceso de validación, específicamente en el contexto de juicio de expertos. Estos pasos muestran la importancia de elegir expertos cualificados, concretar el propósito del juicio y definir de manera precisa los constructos que se miden, haciendo que el proceso sea más riguroso y confiable.

2.2.2. Diseño de materiales

Este concepto constituye la base teórica para la creación efectiva de recursos pedagógicos. El diseño de material no solo implica la selección del contenido, sino que también la adaptación y estructuración de los materiales para promover habilidades específicas, que en este estudio es la comprensión lectora en el nivel B1 de inglés. Dado que los materiales didácticos están profundamente relacionados al contexto en el que se utilizan, cuando un docente es encargado de diseñarlos debe tomar ciertas decisiones asociadas a la realidad del aula, decisiones que lo llevan a pensar en cómo y qué enseñar (características de los estudiantes, propósitos de aprendizaje, tiempo y recursos disponibles, etc.), es decir, el diseño de materiales debe buscar un equilibrio entre las condiciones del contexto y las necesidades de los estudiantes (Augusto-Navarro, 2015). Por consiguiente, los materiales generados por IA pueden presentar un riesgo al intentar conseguir este equilibrio, ya que pueden no cumplir con criterios pedagógicos, contextuales o curriculares si no son revisados y validados antes por expertos.

Teniendo esto en cuenta, es importante definir y comprender qué características debería poseer un material educativo apropiado y alineado a los objetivos de aprendizaje. Por ello, los conceptos de material didáctico y desarrollo de material se exponen a continuación:

2.2.2.1. Material didáctico

Los materiales pueden ser cualquier recurso que se utilice para ampliar los conocimientos y la experiencia de los estudiantes en la lengua, por lo que pueden ser CDs, videos, libros de texto, fotocopias, entre otros (Tomlinson, 2011). Asimismo, Tambo (2003), citado por Pechele (2016), describe los materiales didácticos como herramientas para la enseñanza y vías de aprendizaje, haciendo énfasis en que estos materiales o recursos pueden enseñar por sí mismos o ayudar a un profesor a enseñar de forma más efectiva. Sin embargo, para que los docentes logren tener un aula exitosa, requieren de materiales y recursos transformativos (Ngum & Tetieuh, 2022). Es así que se vuelve necesario, como se mencionó previamente, explorar cómo el material didáctico creado

mediante la asistencia de IA puede ser válido y fiable para optimizar las habilidades de lectura de los estudiantes de inglés como lengua extranjera.

Morales Muñoz (2012) señala que es fundamental considerar ciertas características de un material didáctico para que este sea efectivo:

- El material debe estar diseñado para alcanzar los objetivos que pretende cumplir.
- Los contenidos deben concordar con los temas de la asignatura.
- Se deben tener en cuenta algunos rasgos del diseñador en cuanto a sus habilidades, estilos cognitivos, conocimientos previos, experiencia e intereses.
- Se debe tener presente el contexto en el que el material será implementado.

De la misma manera, Szököl y Osifčín (2024) expresan que los criterios principales para seleccionar los materiales didácticos son el objetivo que se desea alcanzar, el contenido de la enseñanza, la naturaleza de los fenómenos que se desean demostrar, la edad, el nivel educativo y conocimientos previos de los estudiantes, la finalidad de la enseñanza, así como el nivel del profesor y su habilidad para usar los materiales elegidos.

En general, este concepto destaca la importancia de los materiales didácticos como un elemento clave para el aprendizaje, especialmente cuando se trata de mejorar las habilidades de los estudiantes. Además, la necesidad de materiales transformadores y el cumplimiento de criterios específicos para su selección, tales como el nivel y las características de los estudiantes y el diseñador, revela el valor de revisar los documentos nacionales, Los Derechos Básicos de Aprendizaje de Inglés (2016) y El Esquema Curricular Sugerido (2016), ya que estos proveen las pautas para que los materiales consigan estar alineados con los objetivos y las necesidades del sistema educativo.

2.2.2.2. Desarrollo de materiales

Los materiales didácticos, como herramientas clave para el aprendizaje, deben ser diseñados cuidadosamente para alcanzar los objetivos pedagógicos y responder a las necesidades de los estudiantes. Por lo tanto, es necesario considerar que implica el desarrollo de materiales:

Tomlinson (2019) argumenta que el campo de desarrollo de material involucra investigar los principios y procedimientos relacionados con el diseño, redacción e implementación de materiales con fines educativos, así como analizar los efectos que estos materiales pueden tener sobre los estudiantes. Adicionalmente, Rahmawati (2015) afirma que el desarrollo de materiales para la enseñanza del inglés como lengua extranjera abarca tres procesos: producción, evaluación y adaptación. No obstante, la mayoría de educadores se limitan al proceso de producción y no consideran la evaluación ni la adaptación. Nguyen (2011) expone que “la evaluación de materiales didácticos suele incluir un análisis de las tareas y contenidos, lo que permite encontrar posibles

discrepancias entre los materiales y un grupo de aprendices en particular” (p. 13). Es así que la evaluación de materiales se vuelve un importante factor al momento de saber si satisfacen las necesidades de aprendizaje y sus objetivos lingüísticos. Nguyen añade igualmente que los docentes deben desarrollar habilidades para adaptar actividades que nos son pertinentes y seleccionar las más apropiadas de entre diferentes recursos. Esto les permite cumplir las necesidades específicas de los estudiantes y cerrar brechas entre el contenido estándar de los libros de texto y los intereses, habilidades o niveles de los estudiantes.

Este concepto demuestra que la validación de materiales va más allá de su simple creación, ya que involucra aspectos como la evaluación del contenido y su alineamiento con los objetivos educativos y toma en consideración las características y demandas de los estudiantes. Este enfoque es crucial para asegurar que los materiales no solo sean apropiados, sino que también faciliten el desarrollo de habilidades de comprensión lectora en el contexto del aula de inglés como lengua extranjera.

2.2.3. Materiales auténticos

Los materiales auténticos son recursos utilizados por hablantes nativos en un contexto de comunicación verdadero para interactuar entre sí, en lugar de estar diseñados para propósitos pedagógicos y educativos (Kirana, 2014). Estos materiales muestran un lenguaje “real” y están destinados a un público real, centrándose principalmente en el mensaje y utilizando a menudo elementos como formatos, diseño, estilo y contexto para facilitar la comunicación (Adam et al., 2010). Por el contrario, materiales no auténticos son desarrollados siguiendo el currículo y las políticas educativas de cada país o región y están intencionalmente creados para objetivos de enseñanza y aprendizaje (Febrina, 2017).

De acuerdo con Kusumarasdyati (2019)

[E]n las clases de lectura, los textos auténticos se refieren a las lecturas tomadas de varias fuentes, excepto de los libros escritos para el aprendizaje de lengua, tales como los periódicos o los artículos de revista, obras literarias (novelas, historias cortas o poemas) y textos funcionales breves (anuncios, invitaciones o comunicados) (p. 84).

Los textos auténticos pueden resultar más atractivos considerando que pueden parecer más genuinos y con una dimensión cultural más vívida, lo que hace que la lengua sea más dinámica en lugar de estructurada (Daskalos & Jellum Ling, 2006). Sin embargo, usualmente debido a la complejidad y el reto que el contenido cultural y la lengua meta pueden plantear a los estudiantes, muchos académicos abogan por la adaptación de este tipo de materiales, de modo que los profesores puedan modificar estos recursos para que se adapten mejor al contexto de enseñanza (Febrina, 2017). Es así que la adaptación de materiales auténticos facilita el proceso de aprendizaje

al identificar las dificultades y ayudar a los estudiantes a que interioricen el contenido de manera natural (Ebrahimpourtaher & Hamidi, 2015).

Existen varios métodos involucrados en el proceso de adaptación de un texto, siendo los más comunes la simplificación y la elaboración. Como establece Mousavi (2011) “la simplificación puede definirse como el control del texto dirigido a aprendices de una segunda lengua mediante la eliminación de elementos lingüísticos poco familiares (por ejemplo, construcciones gramaticales y léxico desconocido) con el fin de facilitar la comprensión” (p. 36). No obstante, la simplificación puede resultar contraproducente, debido a que previene a los estudiantes a que se expongan y adquieran la lengua en la que está escrita el texto, lo que a su vez les impide desarrollar habilidades generales de lectura que puedan aplicarse más allá de ese texto en particular (Ross et al., 1991). En cambio, la elaboración, aunque también modifica la lengua para facilitar la comprensión, conserva el vocabulario desconocido sin alterar las estructuras lingüísticas y mantiene el significado original del texto (Luís, 2024). Además, según Long (2014) la elaboración consiste en añadir redundancia a un texto a través de distintas estrategias lingüísticas, como la repetición, la paráfrasis, el uso de sinónimos, la segmentación y la reiteración. Ambos métodos pueden utilizarse para adaptar textos auténticos en función de los objetivos de aprendizaje y las necesidades de los alumnos.

En este estudio se seleccionaron cuatro textos auténticos que fueron sometidos a un proceso de adaptación a nivel B1 mediante el uso de IA. A esta herramienta se le solicitó explícitamente en el *prompt* que adaptara cada texto a través del proceso de elaboración, con el fin de ajustar el contenido a las habilidades del nivel B1 sin perder el sentido ni la intención comunicativa que se quería expresar originalmente.

2.2.4. Comprensión lectora

Considerando que la lectura es una destreza central cuya eficacia se desea evaluar en los materiales creados por ChatGPT, es imperativo abordar una percepción clara de su conceptualización. Por un lado, según Snow (2002), la lectura es el proceso de extracción y construcción de significado mediante la interacción y la participación del lector, el texto y el propósito o la actividad de lectura, todos ellos dentro de un contexto sociocultural. Al comprender un texto, el lector debe reconocer sus capacidades, conocimientos, experiencias y motivaciones para construir significado, establecer relaciones entre ideas y definir propósitos de lectura, mientras su conexión con su entorno social y cultural influye en su comprensión y aprendizaje. Es decir, la comprensión lectora involucra que el lector use lo que sabe, ha vivido y le interesa.

Por otro lado, Cline, Johnstone y King (2006) conciben la lectura como un proceso compuesto de decodificación y de comprensión. De acuerdo con estos autores,

leer es decodificar y comprender textos escritos. Decodificar requiere traducir los símbolos de los sistemas de escritura (incluido el braille) a las palabras habladas que representan. La comprensión está determinada por los propósitos de la lectura, el contexto, la naturaleza del texto y las estrategias y conocimientos del lector. (p. 2).

Es decir, la lectura es un proceso que no solo tiene en cuenta la decodificación del lenguaje escrito. Leer es una secuencia dinámica que implica que el lector se involucre y traiga sus conocimientos previos, demostrando un compromiso profundo y una participación consciente en la construcción del significado del texto.

Para Grabe y Yamashita (2022),

cuando nosotros leemos, coordinamos el reconocimiento rápido y automático de palabras, el análisis sintáctico, la formación de significado, la comprensión textual, la inferencia, la evaluación crítica y la vinculación con conocimientos previos. Lo hacemos aparentemente sin esfuerzo y con todos los procesos sincronizados. (p. 16).

Esta habilidad es un acto cognitivo y coordinado que conlleva a que el lector interactúe con el texto y desarrolle una interpretación. El lector decodifica, interpreta, construye ideas, monitorea su proceso y usa estrategias para lograr su objetivo. De esta forma, durante la lectura se produce una ejecución sincronizada y automática de una serie de habilidades jerárquicas para la construcción de significado.

Por lo tanto, la lectura es un proceso activo y dinámico que implica interpretar el significado, conectar el conocimiento previo con la nueva información y construir una comprensión personal del texto. Dada su relevancia en el proceso de aprendizaje de la lengua y su capacidad de fomentar un conocimiento activo, su definición es útil como orientación sobre las cualidades que debe ofrecer un material de comprensión lectora. Por ello, es necesario evaluar y validar dichos materiales hechos con ChatGPT con el fin de que estos expertos puedan determinar si esta habilidad se puede desarrollar de forma efectiva mediante dichos recursos.

2.2.4.1. La lectura en una segunda lengua (L2)

Dado que se espera que los materiales elaborados con ChatGPT fomenten el aprendizaje de la lectura en una segunda lengua (L2 por su sigla en inglés), es esencial precisar lo que se comprende por esta habilidad en dicho contexto. Brown (2002) afirma que esta habilidad no se desarrolla por medio de una simple exposición a los textos; la lectura requiere un desarrollo consciente y un procesamiento de palabras y frases, junto con los conocimientos previos, contexto y experiencia cultural, para interpretar el texto. El lector en L2 debe anticipar las convenciones textuales y aplicar estrategias que le permitan interactuar con el texto y construir un significado

global a partir de su contenido. Por lo tanto, la lectura en una L2 debe estar guiada para que los estudiantes logren una comprensión eficiente.

Centrándose en la relación entre la lengua materna y la lengua meta, Grabe (2006) explica que la lectura en L2 está influenciada por la lengua materna del lector, tanto facilitando como interfiriendo con la adquisición de esta habilidad. Leer en la lengua materna comparte los mismos procesos cognitivos que en una segunda lengua, pero el desempeño del lector en L2 puede verse condicionado frente al vocabulario, la gramática y el conocimiento previo. Asimismo, Grabe y Yamashita (2022) afirman que, cuanto mayor es la diferencia entre la lengua materna y la segunda lengua, mayor es la probabilidad de interferencia en el aprendizaje del lector. Las personas no aprenden una lengua desde cero, sino que se apoyan en sus habilidades de comprensión y sus conocimientos adquiridos en su lengua materna al leer en la lengua meta.

Por ello, Mlakar (2020) señala que las exigencias en la lectura en L2 son mayores debido a sus diferencias lingüísticas y ortográficas, así como al nivel de dominio del lector en la lengua meta. La autora señala que el proceso de lectura en una segunda lengua está condicionado por el grado de dominio en aspectos como la decodificación, la comprensión del vocabulario, la gramática y la comprensión auditiva, las cuales permiten al lector acceder progresivamente a un mayor nivel de comprensión, como la inferencia, el pensamiento crítico y la integración de ideas. De hecho, según Hamada et al. (2022), la comprensión lingüística, especialmente la comprensión oral, tiene mayor influencia que la decodificación en la comprensión lectora. Asimismo, las habilidades metalingüísticas, como la consciencia fonológica, morfológica y ortográfica, no influyen de manera directa, sino indirecta, al fortalecer la decodificación y la comprensión lectora. En este sentido, la lectura en L2 no se entiende únicamente por habilidades individuales, sino también por las oportunidades de exposición y el uso significativo de la lengua en su contexto.

En definitiva, la lectura en L2 es una habilidad que requiere un proceso guiado que integra decodificación, comprensión lingüística y conocimientos previos. Aunque comparte procesos cognitivos con la lectura en la lengua materna, su desarrollo puede verse condicionado por el nivel de dominio de la lengua meta, las diferencias entre las lenguas y la exposición a la L2, aspectos que son fundamentales a considerar en el diseño de materiales pedagógicos. Por lo tanto, dado que la lectura en una segunda lengua implica la interacción entre la decodificación y la comprensión lingüística, es fundamental tener esto en cuenta a la hora de diseñar los materiales con ChatGPT, con la esperanza de que dichos materiales contribuyan al desarrollo adecuado de esta habilidad en una segunda lengua.

2.2.4.2. Tipos y niveles de comprensión lectora

En el campo de la investigación sobre la lectura, los términos *tipos* y *niveles* de comprensión lectora se utilizan de forma intercambiable. Por lo que en el presente estudio ambos términos forman un marco unificado que, siguiendo lo propuesto por Westwood (2001), incluye la comprensión y niveles literal (superficial), inferencial (interpretativo), crítico (evaluativo) y aplicado (creativo) que se explican a continuación:

2.2.4.2.1. Comprensión literal

La comprensión literal es entendida como el nivel más elemental del proceso de comprensión lectora, puesto que se centra en la identificación directa de la información explícita del texto, como el reconocimiento de detalles y el orden de los acontecimientos (Kagami, 2020). En este nivel los estudiantes deben ser capaces de reconocer la o las ideas principales y secundarias, así como organizar, categorizar y resumir información. (Naniwarsih & Andriani, 2018). En efecto, es común que los docentes evalúen primero este nivel para comprobar si los estudiantes han entendido el significado superficial de una lectura (Day & Park, 2005). Dado que las preguntas de tipo literal exigen a los aprendices simplemente recuperar el contenido del texto tal y como se presenta (Kurniawan, 2015) es usual que se incluyan preguntas básicas como qué, cuándo, dónde y quién, así como el uso de pruebas objetivas, entre ellas ejercicios de verdadero o falso, opción múltiple y completar espacios (Suhadi, 2016). En general, la comprensión superficial o literal abarca habilidades esenciales en el proceso de aprendizaje, especialmente en el de lengua extranjera, que permiten el progreso hacia niveles más complejos de comprensión.

2.2.4.2.2. Comprensión inferencial

En este nivel, el lector comienza a interpretar y predecir información que se encuentra implícitamente dentro del texto. Este tipo de comprensión conlleva la interpretación del lenguaje figurado, la formulación de conclusiones y la anticipación de resultados e ideas del autor (Safitri & Melati, 2023). Asimismo, el proceso inferencial ocurre mediante la interrelación entre distintas partes del texto y la asociación del contenido con los saberes previos (Valdez-Asto, 2022), lo que implica una alta capacidad de abstracción por parte del lector (Flores, 2019). Es por ello que, en un principio, puede que los estudiantes enfrenten dificultades para responder preguntas de tipo inferencial, debido a que el análisis del texto se vuelve un poco más complejo (Day & Park, 2005; Páez-López et al., 2024). En suma, es un nivel que necesita ser fortalecido de manera sólida a través de la práctica constante, guiando al estudiante a encontrar los significados implícitos del texto, puesto que este proceso abre paso a una comprensión más crítica.

2.2.4.2.3. Comprensión crítica

En este nivel se emplea un proceso evaluativo, ya que el lector produce valoraciones del texto, lo acepta o lo rechaza de manera fundamentada dependiendo de su criterio, su formación y sus conocimientos sobre lo leído (Flores, 2019). Algunas formas de desarrollar este nivel incluyen analizar el contenido, la intención y el punto de vista del autor y las ideas transmitidas (Valdez-Asto, 2022).

Las preguntas típicas en este nivel serían: *¿Qué es lo que posiblemente sucederá a continuación en el texto?*; *¿Es lógica la argumentación del escritor?*; *¿Existe algún argumento alternativo?*; *¿Es un hecho o la opinión del escritor?*; *¿Estás de acuerdo con la argumentación del escritor?*; *¿Cuál es la mejor solución para el caso que se plantea en el texto?* (Safitri & Melati, 2023, p. 249).

Debido a factores culturales, es posible que algunos estudiantes se encuentren reacios a cuestionar o discrepar el contenido del texto, por lo que el docente puede guiar el proceso mostrando ejemplos de respuestas a preguntas de carácter evaluativo que incluyan tanto elementos favorables como desfavorables (Day & Park, 2005). En general, este nivel se plantea como aquel que le permite al estudiante desarrollar su pensamiento crítico y a argumentar su opinión con evidencia textual, fundamental para avanzar a una comprensión que vaya más allá del texto, que fomente las habilidades creativas.

2.2.4.2.4. Comprensión aplicada

En el nivel creativo, el lector puede tomar información o ideas de lo leído y transformarlas en pensamientos originales e innovadores (Westwood, 2001). Se trata de trascender el contenido presentado por el autor, lo que requiere que los estudiantes piensen activamente durante la lectura, empleen su imaginación, amplíen el mensaje del texto y hagan uso de sus conocimientos y experiencias previas (Helwa, 2010, citado en Hassan et al., 2018). Se considera el nivel más alto de comprensión lectora dado que es necesario haber desarrollado habilidades en los otros tipos de comprensión, desde la decodificación de palabras hasta la transformación de estas en conceptos internos y la construcción de representaciones mentales complejas que permitan entender y recrear los elementos del texto (Hassan et al., 2018). En definitiva, el desarrollo de la comprensión aplicada es imprescindible porque no limita al estudiante a quedarse solo con el texto, sino que lo invita a reflexionar, generar nuevas nociones, soluciones o perspectivas sobre un tema.

Normalmente, todos los tipos de comprensión lectora se enseñan de manera simultánea y, de igual modo, se considera pertinente que se integren estos niveles en conjunto, debido a que así se puede garantizar un aprendizaje y desarrollo equilibrado y progresivo de habilidades lectoras. Por

consiguiente, para esta investigación se tuvieron en cuenta estos cuatro niveles al momento de crear el prompt que generaría los materiales.

2.2.4.3. Fases de lectura

Los materiales didácticos que refuercen las habilidades lectoras en lenguas extranjeras requieren de una organización que le facilite al estudiante la comprensión del texto. Esto implica que, para las actividades, se lleven a cabo diferentes etapas relacionadas entre sí y que permitan activar los conocimientos previos, guiar la lectura y afianzar o aplicar lo leído posteriormente. Los ejercicios, entonces, pueden dividirse en tres fases: pre-lectura, durante la lectura y post-lectura, que se explican a continuación.

2.2.4.3.1. Pre-lectura

Las actividades de pre-lectura tienen como propósito activar los conocimientos previos de los estudiantes y, al mismo tiempo, ofrecer orientaciones lingüísticas o socioculturales que les permitan enfrentarse de manera más adecuada al texto que van a leer (Napitupulu, 2018). El acercamiento previo al texto pretende captar la atención de los estudiantes y guiarlos hacia una lectura más consciente y con objetivos definidos, ya que el intercambio inicial de ideas los impulsa a analizar los temas y situaciones que se desarrollan en el contenido del texto (Toprak & Almacioğlu, 2009). “Las actividades de pre-lectura incluyen: predicciones basadas en el título, predicciones basadas en el vocabulario, predicciones basadas en preguntas de verdadero o falso, skimming, scanning, juegos de palabras, debates y lluvias de ideas” (Khamraeva, 2016, p.1). Es así que las actividades de pre-lectura les permiten a los estudiantes prepararse lingüísticamente y familiarizarse con el texto, estimulando su interés y saberes previos, de modo que puedan construir un esquema mental del tema o del contenido.

2.2.4.3.2. Durante la lectura

Durante la etapa de lectura, los estudiantes deben tener la oportunidad de comprobar hipótesis iniciales, procesar y estructurar la información, y empezar a desarrollar generalizaciones a partir de los nuevos aprendizajes obtenidos en la interacción con el texto (Tarshaei & Karbalaeei, 2015). Las actividades realizadas durante la lectura contribuyen al desarrollo de estrategias lectoras, profundizan el dominio de la lengua extranjera y facilitan la comprensión de fragmentos del texto que puedan ser difíciles para los estudiantes (Toprak & Almacioğlu, 2009). Esta etapa incluye actividades que fomentan la participación del estudiante mientras lee, permitiéndole alcanzar los objetivos de la clase a través de distintas formas de interacción y trabajo con el texto (Khamraeva, 2016). Las actividades buscan el desarrollo de habilidades de inferencia y el aprendizaje lingüístico y sociolingüístico, recordarles a los estudiantes la importancia del vocabulario como pista contextual para la interpretación del texto, ayudarlos a aprender a generalizar sobre el tema que se

discute, entre otros (Saricoban, 2002). En las actividades durante la lectura, por lo tanto, los estudiantes interactúan con el texto y construyen su significado. Igualmente, monitorean y ajustan, con ayuda del docente, estrategias de lectura que le brinden la posibilidad de una mejor comprensión.

2.2.4.3.3. Post-lectura

Las actividades posteriores a la lectura se enfocan en cimentar y ampliar la comprensión del texto, profundizando en las ideas trabajadas anteriormente (Harida, 2016). Los estudiantes reconstruyen el texto a través de la evaluación, la síntesis, el análisis y responden a preguntas críticas (Güzel, 2022). La post-lectura también puede incluir las percepciones de los estudiantes tanto del texto como de las actividades desarrolladas durante la lectura, por consiguiente, pueden manifestar su valoración personal sobre el interés y la utilidad del material trabajado (Peñañiel Arcos, 2013). En general, es una fase que revisa el entendimiento del estudiante sobre el texto y lo invita a reflexionar. Se trata de un nivel más crítico, donde se le brinda al estudiante la oportunidad de opinar, reinterpretar, establecer conexiones con su realidad o ejercer su creatividad.

Las tres fases de lectura (pre, durante y post) se tuvieron en cuenta al momento de diseñar la instrucción para generar los materiales de comprensión lectora, con el fin de presentar una estructura que promoviera la interacción continua con el texto. Además, esta organización favoreció la formulación de actividades variadas y con secuencia lógica, y al mismo tiempo, facilitó la distribución de ejercicios para cada nivel de comprensión (literal, inferencial, crítico y creativo). De este modo, se realizaron materiales didácticos estructurados y completos.

2.2.5. La lectura en el nivel B1 de inglés

El presente estudio se centra en analizar la eficacia y la pertinencia de los materiales de comprensión lectora generados por ChatGPT, específicamente diseñados para el nivel de inglés B1, según el Marco Común Europeo de Referencia para las Lenguas (2020), el Esquema Curricular Sugerido (2016) y los Derechos Básicos de Aprendizaje de Inglés (2016), ambos documentos propuestos por el Ministerio de Educación de Colombia (MEN). En este sentido, resulta pertinente definir el alcance de estos documentos y su concepción del nivel B1 en la enseñanza del inglés como lengua extranjera.

Por un lado, a nivel global, el Marco Común Europeo de Referencia para las lenguas (MCER) es un estándar internacional que define las competencias de la lengua y establece una base común para el diseño de programas, currículos, exámenes y materiales de enseñanza del inglés como lengua extranjera (Council of Europe, 2001). El MCER es un instrumento político lanzado en 2001 que busca contribuir a la misión del Consejo de Europa de lograr una mayor unidad, promoviendo el derecho a una educación de calidad en la enseñanza de lenguas para todos (Council

of Europe, 2020). Este instrumento se utiliza para definir los alcances y competencias que los estudiantes deben desarrollar, organizadas en una escala de niveles de inglés que va desde A1, nivel básico de inglés, hasta un C2, correspondiente al dominio más avanzado de la lengua. De esta forma, el MCER ha tenido una gran influencia en los contextos donde se enseñe un nivel de inglés, ya que reconoce las competencias lingüísticas a partir de descriptores como la comprensión oral, la comprensión escrita, la producción oral, la producción escrita, la interacción oral, la interacción escrita, la interacción en línea, entre otros.

En los descriptores del MCER, según el Council of Europe (2020), el nivel de inglés B1 se describe como un individuo comprende y se comunica en contextos cotidianos, entiende ideas y detalles básicos, interactúa con fluidez, expresa opiniones, relata experiencias y produce textos sencillos y comprensibles. Además, aunque el individuo puede presentar algunas limitaciones, puede mantener una conversación y transmitir información. En la habilidad de lectura, el aprendiz comprende textos cotidianos, reconoce información relevante, entiende descripciones, narraciones y textos argumentativos simples, distingue conectores, identifica ideas principales, hace inferencias y deduce el significado de palabras que aún no conoce.

Por otro lado, a nivel nacional, el Esquema Curricular Sugerido (ECS) es una propuesta orientadora, diseñada por el MEN (2016), que presenta un marco curricular estructurado destinado a mejorar la enseñanza del inglés como lengua extranjera. Este currículo establece los estándares de alcance y secuencia, los componentes del programa de estudios, los enfoques metodológicos, las estrategias de evaluación y se alinea con los niveles del MCER, desde el nivel A1 hasta B1.3. Asimismo, se concibe como una guía flexible centrada en proporcionar insumos para la planificación y ejecución de una enseñanza eficaz del inglés, en coherencia con el objetivo nacional de promover el bilingüismo en los diversos contextos educativos. En los niveles de secundaria, los grados están ordenados de acuerdo con el nivel de dominio en la lengua meta: el sexto grado corresponde al nivel A1, séptimo a A2.1, octavo a A2.2, noveno a B1.1, décimo a B1.2 y undécimo a B1.3. Cada grado se ordena en cuatro módulos temáticos, los cuales integran un tema transversal, metas, funciones, objetivos, indicadores de desempeño (saber, saber hacer, saber ser) y contenidos sugeridos, que abarcan los componentes léxico, gramatical, fonológico, discursivo y sociolingüístico/intercultural.

Dado que la investigación se centra en el nivel B1 en inglés, se analizaron los contenidos y módulos del grado 9 en la habilidad de lectura. En estos apartados, se consideran temáticas como globalización (fenómenos sociales), sostenibilidad (reducción de inequidades), salud (prevención de enfermedades) y democracia y paz (derechos humanos). De acuerdo con el ECS (2016), el nivel B1 corresponde a un nivel de desempeño en el cual el estudiante es capaz de comprender y analizar textos escritos y orales y sobre temas de mayor complejidad. En el área de la lectura, el individuo es

capaz de identificar y analizar información relevante de textos orales y escritos de extensión corta y media, reconociendo relaciones de causa y efecto, contraste y adición, así como argumentos a favor y en contra en textos académicos y sociales.

En suma, los Derechos Básicos de Aprendizaje de Inglés (DBA) constituyen un documento gubernamental, elaborado desde el proyecto de bilingüismo del MEN (2016), que describe los derechos y expectativas de los estudiantes en el contexto colombiano. Con un enfoque específico en el aprendizaje del inglés como segunda lengua, su objetivo es ser una herramienta de apoyo para los docentes en el desarrollo de propuestas curriculares con el fin de garantizar el aprendizaje integral, evaluar el logro de aprendizajes y definir acciones de mejora continua. En este sentido, este documento recopila los saberes y competencias comunicativas que los estudiantes deben desarrollar en su aprendizaje del inglés durante su escolaridad, organizado por grados desde sexto hasta once de secundaria. Los DBA reconocen las cinco habilidades que los estudiantes deben fortalecer: la escucha, la lectura, la escritura, el monólogo y la conversación.

Dado que tanto los DBA como el ECS se conciben como herramientas articuladas que buscan potenciar las habilidades comunicativas en el inglés como lengua extranjera, se toma como referencia la descripción del grado noveno para la conceptualización del nivel B1. De acuerdo a los DBA, en el contexto del aprendizaje del inglés como lengua extranjera en grado noveno, el estudiante es capaz de procesar información más extensa, saber argumentar, organizar ideas y analizar textos con estructuras más complejas. En el ámbito de la lectura, se espera que el individuo comprenda y analice información académica, reconozca relaciones de causa y efecto, exprese y justifique sus ideas, e identifique la estructura y el propósito de los textos.

En resumen, una persona con un nivel B1 en inglés, es un usuario independiente capaz de comprender y producir el inglés en diversos contextos, manejar diversas temáticas, y expresar opiniones e ideas coherentemente. En la habilidad de lectura, el aprendiz es capaz de analizar textos sobre temas académicos y cotidianos, hacer conexiones lógicas, identificar propósitos, expresar opiniones y comunicar ideas a partir de lo que lee. De esta forma, para la creación del *prompt* que comandó a ChatGPT en la elaboración de materiales de comprensión lectora, se consideraron todos los aspectos que involucran el nivel de B1 en la habilidad de lectura, según estos tres documentos oficiales (MCER, DBA y ECS). Estos formatos permiten garantizar, desde una dimensión pedagógica, materiales alineados con dichos estándares, para su posterior validación por expertos.

2.2.6. La cultura y su relevancia en el aprendizaje de lenguas

Aunque este estudio no aborda la cultura en un sentido muy amplio, es evidente que, al leer, el lector usa sus propios conocimientos culturales y experiencia para construir un significado (Mckay, 2006). Este aspecto se tomó como un aspecto importante dentro de la rúbrica de

validación. Por ello, resulta importante comprender a qué se refieren estos aspectos culturales y su relevancia en el aprendizaje de una segunda lengua. El término de "cultura" se refiere a un grupo de individuos que comparten una serie de constituyentes (visiones del mundo, comportamientos, valores, símbolos, actividades cotidianas, maneras de expresarse y comunicarse, etc.) que, al compartirse y desarrollarse en una sociedad, permiten diferenciarse de otros (Muñoz, 2023). En suma, la cultura constituye un conjunto de normas que forman parte de la vida de quienes comparten pensamientos, historia y un espacio común, y permanece en sus identidades incluso al salir de las comunidades (Kramsch & Zhu, 2016).

En el ámbito de las lenguas, se comprende que la lengua y la cultura están vinculadas (Coracini, 2010). Al comunicarse, se consideran las pertenencias socioculturales tanto del emisor como del receptor, realizando un intercambio o negociación de identidades (Giménez, 2011). La lengua es dinámica y permite a los individuos construir su identidad, interpretar el mundo e intercambiar perspectivas y cultura. Por ello, dentro de la enseñanza de lenguas, es fundamental promover una competencia intercultural. Según Liddicoat y Scarino (2013), esto implica reconocer que no hay una única forma correcta de hacer las cosas, encontrar formas personales de interactuar con otras culturas, usar su conocimiento cultural existente para aprender sobre nuevas culturas y, finalmente, reflexionar sobre su identidad y estilo intercultural en la comunicación. En este contexto, el maestro debe actuar como guía, acompañando a los estudiantes en el proceso de descubrimiento de su identidad y en la construcción de una comprensión compartida, valorando la diversidad como un aspecto enriquecedor (Hossain, 2024).

En definitiva, la cultura y la lengua son aspectos inseparables que forman parte de las identidades que tienen los individuos. Por ello, se consideró el contenido cultural durante la selección de los textos que serían adaptados. Las temáticas elegidas son coherentes con el Esquema Curricular Sugerido (2016), lo que proporciona una orientación clara en contexto, y son relevantes para motivar a los estudiantes a participar, con el fin de que puedan intercambiar sus identidades y compartir sobre sus perspectivas y sus culturas. Aunque el estudio no se centra en el desarrollo de una competencia intercultural, se consideró la cultura y su relevancia en la enseñanza de lenguas para desarrollar un material de lectura que promueva el aprendizaje inclusivo y activo.

2.2.7. La inteligencia artificial en la educación

La inteligencia artificial es un sistema capaz de realizar tareas propias del pensamiento y la inteligencia humana (Díaz-Guio et al., 2023). El término fue acuñado por John McCarthy en 1956 y, con el paso del tiempo, ha experimentado cambios continuos que han demostrado avances significativos (Haenlein & Kaplan, 2019). Por esta razón, su conceptualización siempre ha sido objeto de debate. Sin embargo, desde un enfoque más acertado con este estudio, esta herramienta es

la capacidad de un sistema para comprender datos externos, aprender de ellos y utilizar ese conocimiento para cumplir objetivos y tareas específicas (Kaplan & Haenlein, 2019). La IA es una tecnología avanzada que, mediante el procesamiento de grandes cantidades de datos y el uso de algoritmos, puede mejorar sus sistemas, razonar y analizar para realizar diversas tareas y resolver problemas, simulando la inteligencia humana. De igual forma, esta herramienta es capaz de comprender el lenguaje y crear nuevos contenidos, como imágenes, vídeos, textos, música o códigos. Los sistemas de la IA se basan en modelos de aprendizaje profundo entrenados con datos, lo que les permite no solo interpretar instrucciones en lenguaje natural y crear nuevos datos o contenidos desde ellas (Dand et al., 2022), sino también adaptarse a las preferencias, anticiparse a las necesidades y aprender del usuario (Levy, 2024). Esta tecnología tiene un gran avance en el procesamiento del lenguaje natural y la gramática, capaz de abordar prácticamente cualquier tema. Gracias a su capacidad de brindar información al instante, la IA ofrece positivas aplicaciones no solo en la industria, sino también en el aprendizaje y la educación.

Por ello, es un hecho que a medida que sus múltiples ventajas han aumentado, estas herramientas tecnológicas se han popularizado a nivel global. Su alta eficacia en la resolución de problemas y su relevante desempeño de tareas, se ha extendido a los entornos educativos. La inteligencia artificial permite replantear la enseñanza y apoyar el aprendizaje. Este sistema permite personalizar la enseñanza, favorecer la autonomía del estudiante, dar asistencia o tutoría, proporcionar una retroalimentación y ser un apoyo pedagógico (Holmes et al., 2019). Por lo tanto, esta poderosa herramienta es útil tanto para docentes como para estudiantes, creando oportunidades para mejorar las experiencias de aprendizaje, optimizar el trabajo docente y la enseñanza para personas en todos los niveles educativos. Al proporcionar experiencias de aprendizaje efectivas adaptadas a sus necesidades, capacidades y preferencias, esta tecnología fomenta una relación positiva entre estudiantes y docentes.

Por una lado, para los estudiantes, estos modelos de IA pueden ayudar en el desarrollo de habilidades lingüísticas, proporcionar retroalimentación, tutoría y práctica para el inglés como lengua extranjera, fomentar el pensamiento crítico y la capacidad de resolución de problemas, ofrecer consejos de aprendizaje personalizados, buscar información y recursos, generar actividades o cuestionarios para practicar y proporcionar explicaciones sobre temas o lecturas específicas (Haggag, 2021; Quintero, 2024; Xiao et al., 2023). Así, la IA demuestra un gran potencial con diversos beneficios y oportunidades que contribuyen a un proceso de aprendizaje eficaz.

Por otro lado, para el profesorado, estas tecnologías avanzadas ofrecen posibilidades para revolucionar el aprendizaje y facilitar sus procesos. Estos beneficios incluyen la asistencia en el desarrollo o la creación de materiales didácticos, el apoyo en la planificación de clases, la facilitación de la evaluación de las actividades del alumnado, la aportación de sugerencias y la

presentación de ejemplos de actividades didácticas para el aula (Contreras & Pinzón, 2024; Quintero, 2024; Sihite et al., 2023; Xiao et al., 2023). Además, estas herramientas pueden ser útiles para su implementación en el aula, ya que esta innovación puede involucrar al alumnado, mejorando tanto el dominio de las herramientas innovadoras por parte del profesorado como el desarrollo de sus habilidades de aprendizaje y tecnológicas.

En definitiva, la implementación de la inteligencia artificial puede producir aportes significativos al aprendizaje y presentarse como un apoyo para los docentes. Por ello, el presente estudio busca evaluar el nivel de efectividad con el que, dadas a las múltiples funciones de la IA, esta herramienta puede contribuir a la creación de materiales de comprensión lectora en una segunda lengua, a partir de la percepción de docentes expertos, quienes pueden determinar el nivel de pertinencia, su adecuación al contexto y su coherencia con los criterios que debe cumplir un material de este tipo.

ChatGPT

ChatGPT es una de las IA con mayor impacto global, incluyendo la educación. Este chatbot, lanzado por OpenAI en 2022, es capaz de interactuar con personas mediante diálogos, comprendiendo y generando textos como lo haría un ser humano. Según OpenAI (2022), ChatGPT está diseñado para interactuar conversacionalmente, lo que le permite responder preguntas de seguimiento, admitir errores, cuestionar premisas incorrectas y rechazar solicitudes inapropiadas. Esta IA tiene una capacidad avanzada para comprender el lenguaje humano (Lim et al., 2023), implementar técnicas de aprendizaje automático como las personas (Segarra et al., 2024) y examinar la relación entre los datos (Haleem et al., 2022). Además, ChatGPT interactúa con el usuario a través de prompts (indicaciones) un prompt claro y específico puede hacer que la comunicación sea más efectiva y su propósito sea cumplido (Morales-Chan, 2023). Es desde esta indicación que la herramienta, con una capacidad de búsqueda rápida y potentes habilidades gramaticales y de escritura, puede ayudar con su amplia gama de tareas, como ofrecer consejos, crear textos, resolver problemas eficientemente y generar imágenes. Su efectividad ha llegado incluso a entornos académicos y, aunque su creciente uso ha generado múltiples retos y cuestionamientos (Ray, 2023; Molina-Montalvo, 2025), esta herramienta ha demostrado que, si se usa adecuadamente, puede permitir un ambiente de aprendizaje activo y autónomo para los estudiantes (Haggag, 2021; Holmes et al., 2019), así como ser una herramienta de apoyo para la optimización de tareas de maestros (Quintero, 2024; Sihite et al. 2023; Xiao et al., 2023).

Es fundamental explicar que existen diferencias entre las arquitecturas de dos recientes modelos de ChatGPT, GPT-3.5 y GPT-4, dado que el presente estudio se centrará en el modelo GPT-4. Estas diferencias no solo se limitan a que GPT-3.5 es de acceso gratuito mientras que GPT-4

es de pago, sino que también involucran aspectos relacionados con sus diseños y capacidades. Como explica OpenAI (2023), GPT-4, la versión que se utilizará en esta investigación, ofrece un sistema más versátil, detallado y avanzado. Este modelo de lenguaje multimodal entiende y genera un lenguaje más natural, mejora la precisión de sus respuestas, proporciona un contexto más amplio y es más creativo que su predecesor, demostrando un rendimiento relevante en entornos profesionales y académicos. También es capaz de aceptar entradas de imágenes y texto y generar una respuesta con base en eso. De hecho, este modelo ya ha presentado resultados positivos en el desarrollo de preguntas para actividades en entornos educativos (Doughty et al., 2024).

En definitiva, este chatbot ha ofrecido diversas utilidades y ha demostrado una efectividad positiva en la creación de textos y el cumplimiento de tareas. Sus efectos favorables se han evidenciado incluso en contextos académicos, al apoyar tanto al maestro como al estudiante en un ambiente que promueva el aprendizaje. De esta forma, resulta pertinente evaluar su nivel de efectividad en este ámbito y su capacidad de adaptación de textos y creación de materiales de comprensión lectora, haciendo uso de un modelo reciente, la versión GPT-4.

Este capítulo analizó la noción de validación mediante el juicio de expertos como un proceso que puede proporcionar una evaluación clara y precisa de los materiales, garantizando que cumplan con los requisitos pedagógicos necesarios para su eficacia en el aula. Se consideró la conceptualización de material didáctico, el cual enfatiza la importancia de comprender qué tipo de material se puede utilizar y cómo debe estructurarse. Se identificaron aspectos como la comprensión lectora, la lectura en una segunda lengua y las fases de lectura, las cuales proporcionaron un marco teórico y práctico para el diseño de materiales que favorezcan el desarrollo de las habilidades lectoras en un contexto educativo. Asimismo, se detalló la habilidad de lectura en el nivel B1 en inglés, teniendo en cuenta las conceptualizaciones de los documentos oficiales (ECS, DBA y MCER) y la relevancia de la cultura como un contenido fundamental en la enseñanza de lenguas. En la revisión también se reconoció la eficacia y las oportunidades que ofrece la IA en la educación, como ChatGPT, como una herramienta de apoyo tanto para docentes como para estudiantes. En conjunto, estos constructos proporcionan las herramientas necesarias para el diseño, la adaptación y la validación del material a crear con IA para el contexto educativo nacional, garantizando que sea relevante, efectivo y culturalmente apropiado.

El siguiente capítulo describe la metodología empleada para el desarrollo de este estudio. Se presenta el enfoque y el diseño, así como los métodos de recolección de datos que se utilizaron para analizar la percepción de los expertos sobre los materiales de comprensión lectora generados por

ChatGPT. Los procedimientos, el perfil de los participantes, la validez de los resultados y las consideraciones éticas de la investigación también se explican en detalle.

3. METODOLOGÍA

En este estudio se propone analizar la efectividad y la pertinencia de materiales de comprensión lectora generados por IA para aprendices colombianos de inglés nivel B1 de acuerdo con Los Derechos Básicos de Aprendizaje de Inglés (2016), al Esquema Curricular Sugerido (2016) por el Ministerio de Educación colombiano y al Marco Común Europeo de Referencia para las Lenguas (2020). Mediante la validación de expertos, esta investigación intenta asegurar la calidad y coherencia de los materiales con los estándares establecidos, evaluando su relevancia en términos de vocabulario, sintaxis y contenido, su concordancia con las competencias en inglés del nivel B1, su inclusividad cultural, identificando, al mismo tiempo, sus puntos fuertes y limitaciones. Es así como al ofrecer estas perspectivas basadas en la experiencia, los resultados podrían aportar una potencial guía para los educadores sobre cómo incorporar eficazmente materiales generados por IA en su enseñanza. Para lograr estos objetivos, en esta sección se expone a detalle la metodología de la investigación, explicando los enfoques, procedimientos e instrumentos que se utilizaron para la recolección y el análisis de datos.

3.1. Enfoque y diseño

El estudio adopta un enfoque cualitativo, ya que, tal como es descrito por Creswell y Cresswell (2018) la investigación cualitativa se basa en la recolección de distintos tipos de datos, como entrevistas, observaciones, documentos o materiales audiovisuales, en lugar de apoyarse en una sola fuente. Añaden, además, que todos los datos recopilados se analizan, se interpretan y se organizan en diferentes códigos y temas que atraviesan todas las fuentes de información, buscando construir una comprensión profunda y detallada del fenómeno estudiado según las diferentes perspectivas y los múltiples factores que intervienen en una situación. En ese sentido, este estudio busca explorar y comprender las percepciones y opiniones de expertos en la enseñanza de lenguas extranjeras para evaluar la validez del contenido de los materiales de comprensión lectora elaborados por IA, mediante la interpretación y el análisis descriptivo de los datos extraídos de diferentes fuentes, tales como los documentos gubernamentales, una rúbrica de evaluación y entrevistas.

De la misma manera, dado el propósito de la investigación, se opta por un diseño descriptivo, ya que se enfoca en examinar las características y particularidades del objeto de estudio a partir de la información recopilada (Siedlecki, 2020). Para ello, se emplean instrumentos de carácter cualitativo que facilitan la descripción del aspecto a analizar, que en este caso las percepciones de los expertos.

La Tabla 3.1 muestra los métodos de recolección de datos que se utilizaron para obtener las percepciones de los expertos sobre los materiales de comprensión lectora para nivel B1 elaborados

por ChatGPT. Asimismo, en la tabla se muestran los objetivos que sustentan la elección de los instrumentos en relación a las preguntas de investigación y qué tipo de análisis se llevó a cabo.

Tabla 3.1

Resumen de los métodos de recolección de datos

Preguntas de Investigación	Método de recolección de datos	Objetivos	Análisis de datos
P1: ¿Qué tan pertinentes son, desde el punto de vista pedagógico, los materiales de lectura generados por ChatGPT para el desarrollo de habilidades de comprensión lectora en el nivel B1, en términos de vocabulario, sintaxis y adecuación del contenido, según los expertos en la enseñanza del inglés como lengua extranjera?	Rúbrica de validación Entrevista	Explorar la percepción de los expertos acerca de la adecuación pedagógica de los materiales generados por ChatGPT, teniendo en cuenta diversos aspectos lingüísticos	Análisis descriptivo para explicar en detalle las percepciones de los expertos acerca de los materiales Análisis temático para identificar y discutir patrones encontrados en las respuestas de los expertos
P2: ¿Qué tan eficazmente se alinean los materiales generados por ChatGPT con los estándares de competencia lingüística establecidos para el nivel B1 en comprensión lectora?	Rúbrica de validación Entrevista	Explorar la percepción de los expertos acerca del grado de alineación de los materiales generados por ChatGPT con los estándares de competencia lingüística del nivel B1 para comprensión lectora	Análisis descriptivo para explicar en detalle las percepciones de los expertos acerca de los materiales Análisis temático para identificar y discutir patrones encontrados en las respuestas de los expertos
P3: ¿Qué tan culturalmente relevante es el contenido de los materiales generados por ChatGPT para el contexto educativo colombiano?	Rúbrica de validación Entrevista	Explorar la percepción de los expertos sobre la pertinencia cultural de los materiales generados con ChatGPT	Análisis descriptivo para explicar en detalle las percepciones de los expertos acerca de los materiales Análisis temático para identificar y discutir patrones encontrados en las respuestas de los expertos
P4: ¿Cuáles son las fortalezas y limitaciones de la IA en la generación de materiales de	Entrevista	Explorar la percepción de los expertos acerca de las ventajas y desventajas del uso de	Análisis temático para identificar y discutir patrones encontrados en las respuestas de los

3.2. Participantes

En este estudio participaron dos profesores de universidad, un profesor de secundaria y tres profesores tanto de universidad como de secundaria, todos con experiencia en la enseñanza del inglés como lengua extranjera o en la elaboración y desarrollo de material didáctico. La selección de estos tipos de participantes fue necesaria debido a los diferentes enfoques y percepciones que pueden aportar.

Por un lado, los profesores de la universidad que igualmente son formadores de docentes de educación secundaria ofrecieron una perspectiva académica y teórica centrada en el contexto educativo en el que ejercen, donde la enseñanza del inglés suele orientarse a niveles intermedios o avanzados y se fundamenta en teorías y prácticas del aprendizaje de lenguas. Su formación y prácticas de enseñanza pudieron haberles permitido evaluar la calidad pedagógica de los materiales desde un enfoque teórico y en términos de alineación con los estándares internacionales. Además, dado que normalmente trabajan con estudiantes de niveles más altos, existía la posibilidad de que examinaran los materiales con una mirada centrada en el nivel de proficiencia y dificultad, evaluando si los contenidos eran apropiados para un nivel de competencia intermedio-alto y considerando aspectos relacionados con la relevancia y la complejidad de los temas tratados en los textos.

Por otro lado, los profesores con experiencia en educación secundaria provienen de un contexto de educación media, donde la enseñanza del inglés se desarrolla en niveles más básicos a intermedios. Estos docentes pudieron aportar una perspectiva más práctica, lo que les pudo haber permitido evaluar cómo podrían implementarse los materiales en el trabajo pedagógico habitual, teniendo en cuenta las necesidades específicas de estudiantes en edad escolar y su nivel de dominio del inglés. Los profesores de secundaria igualmente podrían identificar qué materiales se ajustaban mejor a las realidades culturales y sociales de los estudiantes colombianos, garantizando que sean culturalmente pertinentes, comprensibles y funcionales dentro del aula. Además, sus percepciones resultaban relevantes considerando que los documentos institucionales utilizados en este estudio para la generación de los materiales están dirigidos precisamente a orientar a la educación media.

La Tabla 3.2 presenta las características demográficas de los participantes. En general, estos cuentan con titulaciones en educación superior, cuatro de ellos con título de doctorado y dos con título de pregrado. En lo que concierne a la experiencia docente, sus trayectorias oscilan entre 7 a 25 años de ejercicio profesional. Respecto al nivel educativo en el que ejercen actualmente, la

mayoría se dedica a la enseñanza universitaria y solo unos cuantos a la secundaria. La intensidad horaria semanal en la que se dedican a su profesión es diversa, distribuyéndose entre menos de 10 horas, 10-20 horas, 21-30 horas y 31-40 horas de enseñanza. Por último, los participantes reportan la frecuencia en la que deben adaptar sus materiales didácticos para sus clases, indicando que lo realizan frecuentemente, a veces o siempre.

Tabla 3.2

Características de los participantes

Participantes	Nivel de formación académica	Años de experiencia docente	Nivel educativo de enseñanza actual	Horas dedicadas a la enseñanza semanalmente	Frecuencia en la que adapta sus materiales
A	Pregrado (Licenciatura)	7	Universitario	10-20 horas	Frecuentemente
B	Doctorado	36	Universitario	21-30 horas	Frecuentemente
C	Doctorado	51	Secundaria	21-30 horas	A veces
D	Doctorado	42	Universitario	Menos de 10 horas	A veces
E	Doctorado	45	Secundaria y universitario	31-40 horas	Siempre
F	Pregrado (Licenciatura)	47	Universitario y escolar	10-20 horas	Siempre

En relación con la entrevista y con el propósito de profundizar aspectos del formulario, también se les solicitó a los participantes que se presentaran y describieran su formación académica, su trayectoria profesional y su experiencia docente. De hecho, la tabla 3.3 resume la información aportada por cada uno en lo referente a ello. En general, son docentes que cuentan con una formación extensa en el área de la enseñanza del inglés y han participado en diversos proyectos educativos.

Tabla 3.3*Experiencia y trayectoria de los participantes*

Participantes	Resumen de lo dicho en la entrevista
A	Docente de inglés y portugués en la universidad. Lleva dedicándose al diseño de materiales educativos desde 2018, cuando comenzó a trabajar en la Universidad del Valle como diseñador de módulos en inglés, francés y portugués. Además, trabaja en el diseño curricular y pedagógico en la Universidad ICESI.
B	Su formación y experiencia profesional se centran principalmente en literatura en lengua inglesa, aunque ha tenido otros cursos de habilidades en inglés en la Universidad del Valle. Cuenta con un título de maestría en el área de Estudios Literarios y un doctorado en Lenguas y Literaturas Modernas.
C	Cuenta con un título de licenciatura en Lenguas Modernas, maestría en Estudios Interlingüísticos e Interculturales y un doctorado en educación. Tiene 25 años de experiencia docente en inglés y español, de los cuales 20 se han desarrollado en el sector oficial, en una institución en el municipio de Yumbo. Durante la última década se ha centrado en la enseñanza media técnica (grados décimo y undécimo) y actualmente ejerce en noveno grado.
D	Tiene 25 años de experiencia en la enseñanza del inglés, habiendo trabajado en todos los niveles educativos. Actualmente, es docente de inglés, de investigación y de pedagogía didáctica a nivel de licenciatura, aunque también ejerce en programas de posgrado y doctorado. Sus áreas de investigación incluyen la evaluación, enseñanza de lenguas, desarrollo profesional docente y estadística en lingüística aplicada.
E	Es egresado de la Licenciatura en Lenguas Extranjeras Inglés-Francés y del programa profesional de Estudios Políticos y Resolución de Conflictos de la Universidad del Valle. Realizó una maestría en Gestión de la Tecnología Educativa, desarrollando un proyecto de integración de las TIC en la enseñanza del inglés en un colegio público. Recientemente obtuvo un doctorado en Ciencias de la Educación, con un trabajo centrado en la descolonización de la enseñanza del inglés en instituciones etnoeducativas de Colombia. Tiene casi 16 años de experiencia en el magisterio, actualmente es docente en un colegio oficial etnoeducativo de Cali, y lleva 8 años trabajando como profesor contratado en la Universidad del Valle.
F	Cuenta con un título de licenciatura en Lenguas Extranjeras de la Universidad del Valle. Comenzó su experiencia docente en 2006, trabajando en colegios de bachillerato y participando en programas bilingües como <i>Valle bilingüe</i> , gestionado por la gobernación del Valle. Posteriormente, enseñó inglés en diversos colegios bilingües privados y a partir del 2015, se vinculó a la Universidad del Valle. Es fundador de un colegio en Siloé, Cali, orientado a promover el bilingüismo en comunidades de estratos bajos.

La selección de los participantes responde a uno de los criterios propios del juicio de expertos, puesto que son individuos que tienen un conocimiento extenso y una experiencia amplia en el área en el que se desempeñan, siendo en esta ocasión la enseñanza de lenguas extranjeras. Es así que, considerando su trayectoria y los niveles en los que ejercen, los participantes de este

estudio pueden considerarse expertos, lo que respalda la importancia de sus aportes acerca del proceso de validación de material generado por IA.

3.3. Diseño del material

Para el diseño de los materiales ([Anexo 1](#)), fue necesario desarrollar una serie de etapas para llegar a un producto adecuado que cumpliera con el objetivo de la investigación. Estas fases fueron las siguientes:

- **Selección de materiales auténticos:** se realizó una búsqueda y selección de textos auténticos a partir del concepto de materiales auténticos propuestos por Kusumarasdyati (2019). La selección se fundamentó en los contenidos propuestos para el nivel B1 según el Marco Común Europeo de Referencia para las lenguas (2020), el Esquema Curricular Sugerido (2016) y los Derechos Básicos del Aprendizaje (2016) para el grado noveno, establecido como nivel B1.1 en los documentos nacionales. Inicialmente, se seleccionó un conjunto amplio de textos auténticos, pero se optó por elegir cuatro textos que más se ajustaban a los objetivos formativos. Las temáticas de los textos se alinearon con el enfoque del Esquema Curricular de Referencia con la globalización (fenómenos sociales), sostenibilidad (reducción de inequidades), salud (prevención de enfermedades) y democracia y paz (derechos humanos). Asimismo, los textos permitieron dar cumplimiento a los objetivos planteados en los estándares del MCER, el ESC y los DBA. Los materiales seleccionados desarrollaron temáticas sobre la equidad de género, la ansiedad, la inteligencia artificial y la autoestima, abordadas respectivamente en *"Emma Watson: Gender Equality Is Your Issue Too"*, *"The Imagined Suffering"*, una reseña sobre la película *Her* y *"Lucy's Body Image"*.
- **Adaptación de los textos auténticos:** con ayuda del supervisor, se elaboró un primer prompt ([Anexo 2](#)) con el fin de adaptar los textos a un nivel de inglés B1. En conversaciones separadas por cada material, a ChatGPT se le proporcionó un archivo que contenía la lectura en formato PDF y, mediante el prompt, se le solicitó adaptar para facilitar su comprensión y simplificarlo, proceso conocido como *elaboración* (Long, 2024), orientándolo a aprendices de inglés B1, específicamente estudiantes de bachillerato hispanohablante en Colombia. Con el fin de ofrecer una contextualización más alineada con los objetivos de la investigación, se le pidió que siguiera los criterios del MCER, el ESC y los DBA. Para ello, se le suministraron algunos descriptores u objetivos pertinentes que los aprendices de nivel de inglés B1 deben alcanzar en el proceso de comprensión lectora. Tras la entrega del texto adaptado, se le solicitaron ajustes con respecto a la extensión de los párrafos y con

cambios que alteraban en exceso el contenido original. Finalmente, cuando el producto resultó adecuado, se le pidió a la IA organizar el texto en un documento en formato PDF y entregarlo como resultado final.

- **Creación de los materiales de comprensión lectora:** con base en el texto adaptado, se redactó el segundo prompt ([Anexo 2](#)) que incorporaba también los descriptores y objetivos establecidos en los documentos oficiales. Asimismo, se solicitó la inclusión de preguntas a nivel superficial, interpretativo, evaluativo y creativo, por lo que se proporcionaron diversos ejemplos de tipos de actividades. De igual manera, se indicó que los materiales debían dividirse en las etapas de *pre-reading*, *while-reading* y *post-reading* y, con el fin de contextualizar la instrucción, se incluyó una breve descripción sobre el concepto de comprensión lectora. No obstante, los resultados iniciales no cumplieron con las expectativas, por lo que se le pidió a ChatGPT realizar modificaciones en las actividades, en la cantidad de preguntas y en el orden de presentación del material. Finalmente, una vez el producto resultó adecuado, se le pidió a la IA la elaboración de un documento con las actividades. Aunque en algunos casos no se incluyeron todas las actividades, al final se consolidaron tanto el texto adaptado como las actividades en un único documento en formato PDF.

En resumen, esta etapa inició con la búsqueda y la selección de los textos auténticos que serían adaptados por ChatGPT-4. Posteriormente, se diseñaron los *prompts* que orientaron a la IA en el proceso de creación de los materiales. Se elaboraron dos instrucciones, una para la adaptación del texto auténtico a nivel B1 y otra para la producción de las preguntas y actividades. Finalmente, se realizaron algunos ajustes para asegurar la calidad de los materiales y se entregó el producto final. Es importante señalar que los materiales fueron generados por ChatGPT-4, únicamente se le proporcionó cierta orientación para garantizar el cumplimiento de los criterios del estudio y su adecuada organización para la respectiva revisión.

3.4. Métodos de recolección de datos y procedimientos

3.4.1. Métodos de recolección de datos

Se emplearon diversos instrumentos con el fin de recolectar los aportes de los expertos en cuanto a su percepción de los materiales de comprensión lectora generados por IA. Esta investigación utilizó, en primer lugar, un formulario de participación, no solo con el fin de solicitar el apoyo voluntario de los expertos, sino también para establecer el perfil de los participantes. Posteriormente, se presentó una rúbrica de validación en escala de estimación para que los participantes evaluaran los materiales de acuerdo a ciertos criterios. Por último, se llevaron a cabo entrevistas semiestructuradas para ampliar algunas respuestas de la rúbrica y obtener más

información del juicio de los expertos sobre los materiales. Estos instrumentos permitieron construir una visión más completa del uso de la IA para la creación de materiales didácticos.

3.4.1.1. Formulario de consentimiento de participación

Como primer paso para la recolección de los datos, se realizó una encuesta corta en *Google Forms* ([Anexo 3](#)), donde en un principio se solicitaba explícitamente a los profesores su participación y seguidamente, se indagaba sobre su experiencia y trayectoria profesional, aspectos que se consideraron importantes para establecer su perfil como evaluadores expertos. Por ello, se incluyeron preguntas sobre su experiencia, su nivel de formación académica alcanzado, los niveles educativos en los que enseñó y enseña actualmente, la intensidad horaria en la que se dedica a ejercer, la frecuencia en la que adapta o crea material didáctico y las principales dificultades que enfrenta al trabajar con materiales para la enseñanza del inglés.

3.4.1.2. Rúbrica de validación

Según Rochford y Borchert (2011, citados en Ito, 2015), una rúbrica es un instrumento de evaluación que define los criterios para evaluar un determinado trabajo. Es un tipo de matriz que proporciona niveles en escala de un logro para un conjunto de criterios o dimensiones de calidad del desempeño o rendimiento de algo en particular (Allen & Tanner, 2006). En este estudio se aplicó una rúbrica de validación ([Anexo 4](#)), por medio de *Google Forms*, con el propósito de evaluar los materiales generados por ChatGPT, ya que este tipo de instrumento ofrece indicadores claros que permiten una valoración consistente y organizada de los materiales.

La rúbrica estuvo compuesta por treinta preguntas, más una sección de comentarios, organizadas en seis criterios específicos que permitieron evaluar los aspectos importantes de los materiales, que son: 1) adecuación al nivel B1 (*Adequacy to level B1*), 2) variedad y calidad de los textos (*variety and quality of the texts*), 3) variedad, calidad y precisión de los ejercicios (*variety, quality, and accuracy of the exercises*), 4) relevancia cultural y contextual (*cultural and contextual relevance*), 5) coherencia con los objetivos pedagógicos de lectura (*coherence with the pedagogical reading objectives* y 6) calidad de los niveles de comprensión lectora (*quality of the levels of reading comprehension*). Cada criterio se propuso para medir las dimensiones fundamentales del material de acuerdo a los objetivos del estudio.

Cada pregunta se valoró mediante una escala numérica. Este tipo de escala consiste en otorgar un valor numérico dentro de un rango determinado, en el que cada valor corresponde a una categoría ordenada que representa un aspecto específico del objeto de evaluación (Dörnyei y Taguchi, 2010). Se designó una escala breve de 1 (mínimo) a 5 (máximo), lo cual resulta práctico tanto para los participantes que evalúan como para el análisis de datos, sin hacer el instrumento demasiado extenso o complejo.

Finalmente, para asegurar su fiabilidad, la rúbrica fue sometida a pilotaje. Se contó con la colaboración de dos profesores universitarios y uno de secundaria. Esto permitió comprobar su funcionalidad, identificar posibles errores o ambigüedades en los criterios y en las preguntas y realizar los ajustes necesarios para mejorar su aplicación.

3.4.1.3. Entrevistas

De acuerdo con los objetivos planteados, se llevaron a cabo entrevistas ([Anexo 5](#)) con todos los participantes, puesto que esta constituye un método de recolección de datos cualitativo en el que se busca acceder a la perspectiva personal del entrevistado sobre un fenómeno (Acquah, 2023). En específico, se trataron de entrevistas semiestructuradas, lo que significa que se tenía un conjunto de preguntas establecidas, pero se contó con una cierta flexibilidad para ampliar o profundizar en las respuestas de los participantes (Ellis, 2016). Las entrevistas, además, fueron grabadas con el permiso de cada uno de los expertos entrevistados. Se crearon doce preguntas de base, organizadas en tres secciones que buscaban indagar sobre diferentes aspectos:

- **Sección 1, sobre la experiencia:** En esta sección se tenía el propósito de caracterizar a los participantes. Para ello, se recopilaron datos relacionados a su formación académica y su experiencia y contexto profesional. Si bien fueron aspectos abordados en el formulario de consentimiento de participación, la entrevista permitió profundizar en detalles más específicos. Asimismo, se incluyeron preguntas acerca de experiencias previas con el uso de la IA en su área y su entendimiento sobre las características de un material didáctico adecuado dirigido a nivel B1 en inglés. Estos datos permitieron establecer el perfil de los participantes para catalogarlos como expertos.
- **Sección 2, sobre la rúbrica y los materiales:** En esta sección se buscó profundizar en las valoraciones dadas en la rúbrica. Se trataron cuestiones de aplicabilidad o modificaciones de los materiales. Se exploraron los criterios que los participantes consideraron al momento de evaluar los materiales, teniendo en cuenta si estos permitían desarrollar las estrategias de lectura propias de B1, sus tipos de actividades, si presentaban desfases con relación a los documentos oficiales y si eran culturalmente pertinentes. Estos datos permitieron obtener la percepción de los expertos con respecto a la efectividad de los materiales.
- **Sección 3, sobre la IA:** Esta última sección trató de recoger la percepción de los participantes sobre el uso de la IA en el contexto educativo. Se consultó la utilidad de la IA como herramienta de apoyo para el docente, sus ventajas y desventajas (o riesgos) en cuanto a su uso general y su uso para la creación de materiales. De igual

manera, se solicitaron recomendaciones para la optimización de los materiales. Estos datos permitieron encontrar las posibles fortalezas o limitaciones de la IA en la educación según los expertos.

La entrevista fue diseñada con acompañamiento del supervisor considerando, igualmente, su alineación con los objetivos del estudio. Su estructura permitió organizar el desarrollo de las preguntas con una secuencia lógica. Además, una vez grabadas, se realizó la transcripción por medio de diversas herramientas de IA, como *Otter* y *Fireflies*, con el fin de facilitar la recolección y análisis de la información.

3.4.2. Procedimientos

Primera etapa: Diseño de los *prompts* y de los materiales

Esta etapa inició con la búsqueda y la selección de los textos auténticos que serían adaptados por ChatGPT-4. Posteriormente, se diseñaron los *prompts* que harían parte del proceso de la creación de los materiales. Se elaboraron dos de ellos mediante la guía del supervisor, uno para la adaptación del texto auténtico a nivel B1 y otro para la producción de las preguntas y actividades. Se realizaron varios intentos y ajustes para asegurar la calidad de los materiales, hasta seleccionar aquellos que cumplieran con los criterios del estudio y que serían presentados a los participantes.

Segunda etapa: Diseño de los instrumentos

Una vez listos los materiales, se comenzó a considerar a los posibles participantes del estudio. Se contactaron distintos profesores, tanto de universidad como de colegio, por medio de correo electrónico. Al mismo tiempo, se diseñaron los instrumentos de recolección, iniciando con el formulario de consentimiento, que se procedía a enviar al momento de recibir la confirmación de los participantes. Se continuó con la elaboración de la rúbrica de validación. Se discutieron los posibles criterios que se evaluarían y el formato en el que sería presentada a los expertos. Después del pilotaje, se les remitió a los participantes para que comenzaran con el proceso. Por último, se prepararon las preguntas de la entrevista con ayuda del supervisor.

Tercera etapa: Recolección de los datos de la entrevista

Tan pronto como estuvo terminada la entrevista, se convocó a los profesores que ya habían realizado la evaluación de los materiales para su realización. Nuevamente, al correo electrónico, se les envió un link de *Google Meet* y se programaron las reuniones de acuerdo a sus horarios disponibles. En esta etapa se buscó recopilar y obtener la percepción de los participantes sobre los materiales, profundizando en las respuestas de la rúbrica.

3.5. Análisis de datos

3.5.1. Análisis de la rúbrica de validación

La rúbrica de validación se diseñó utilizando la plataforma *Google Forms*, la cual permite la recopilación de respuestas de manera organizada y proporciona gráficos y una tabla en Excel, lo que facilita la visualización de los datos. La información obtenida se organizó en tablas, elaborando una por cada criterio evaluado. Cada tabla contiene las calificaciones (en escala de 1 a 5) asignadas por los participantes, el promedio por sección y el promedio por experto. Es importante mencionar que las valoraciones obtenidas de la rúbrica y los promedios resultantes fueron interpretados con base en una escala establecida en el estudio de Nieves et al. (2013). En dicha escala, el rango comprendido entre 3.9 y 5 se considera “Muy bueno a excelente”, aquel ubicado entre 2.5 y 3.8 se considera como “Aceptable” y las valoraciones entre 1 y 2.4 se clasificaron como “Pobre a deficiente”. Estas medidas permitieron darle un sentido cualitativo a la evaluación de los expertos, lo que facilitó el análisis.

Los datos de estas tablas fueron examinados a través del análisis descriptivo. Este tipo de análisis consiste en explorar y detallar un fenómeno tal como se presenta, con el fin de comprender sus características principales. Su propósito es organizar y presentar la información recopilada sobre el objeto de estudio, identificando patrones o regularidades que no han sido previamente documentadas (Loeb et al., 2017). Es aquí que, mediante el promedio calculado por sección y por participantes, fue posible identificar las percepciones de los expertos sobre cada criterio evaluado y alcanzar un consenso sobre la validación general de los materiales.

3.5.2. Análisis de las entrevistas

Como se mencionó previamente, las entrevistas fueron registradas y transcritas por medio de diferentes programas de IA (otter.ai y fireflies.ai). Sin embargo, cada grabación fue nuevamente reproducida, con el fin de revisar y editar la transcripción en caso de errores de redacción. Las transcripciones finales se organizaron en documentos de *Google Docs* y se almacenaron en una carpeta compartida en *Google Drive* para facilitar el procesamiento de la información.

Los datos obtenidos de la entrevista se examinaron a través del análisis temático (AT), que según Braun y Clarke (2012) consiste en un procedimiento sistemático para detectar, ordenar e interpretar los patrones de significado que se repiten dentro de un conjunto de datos, no enfocándose en aportes aislados, sino en encontrar sentido a las experiencias y significados colectivos. En el AT, como su nombre lo indica, se identifican temas en común; “un tema recoge algo importante de los datos en relación con la pregunta de investigación y representa un patrón de respuestas o significados que se repite en un grupo de datos” (Braun & Clarke, 2006, p. 10).

Por consiguiente, la identificación de las categorías y subcategorías se llevó a cabo mediante el análisis de las transcripciones. Durante la revisión de los planteamientos de los expertos, se destacaron y subrayaron aquellos fragmentos que resultaban relevantes por su relación precisa con las preguntas de investigación. Las ideas importantes, expresadas de manera recurrente o con especial énfasis por los participantes, se agruparon en temáticas comunes. A partir de todo este proceso, se establecieron las categorías principales y, posteriormente, se delimitaron subcategorías que permitieron profundizar en los distintos aportes de los expertos. La tabla 3.4 muestra, por cada criterio de validación, las categorías de análisis con sus respectivas definiciones, así como las subcategorías identificadas. Estas últimas se ilustran, de igual forma, mediante ejemplos textuales extraídos de las entrevistas.

Tabla 3.4

Categorías y subcategorías identificadas en las entrevistas

Criterio de validación	Categorías	Definición	Subcategorías	Ejemplo
<i>Pertinencia pedagógica de los materiales de lectura de ChatGPT para nivel B1-grado noveno</i>	Concepciones docentes sobre un material didáctico de comprensión lectora adecuado	Percepciones de los expertos acerca de los aspectos que consideran esenciales al diseñar o elegir materiales de comprensión lectora	1. Estructuración del material de comprensión lectora. 2. El desarrollo de habilidades de comprensión lectora. 3. Las características de los textos en términos de interés y complejidad.	<i>“Es prudente empezar por acercamiento a palabras claves [...] y después ir construyendo algo mucho más avanzado”</i> (Entrevista del experto A)
	Calidad del diseño de los materiales generados para el estudio	Valoraciones de los expertos sobre la estructura y características de los materiales elaborados por ChatGPT-4 para esta investigación.	1. Valoración de los ejercicios propuestos. 2. Estructuración del material a través de las fases de lectura. 3. Manejabilidad de la complejidad temática. 4. Relevancia de los contenidos para el contexto estudiantil objetivo.	<i>“Las [preguntas que veo son] pertinentes, está true or false, multiple choice, de todo un poquito”</i> (Entrevista del experto F) <i>“Los temas fueron muy interesantes, muy contemporáneos y apropiados para el nivel y el interés de estudiantes de noveno grado y para el nivel B1 estuvieron apropiados”</i> (Entrevista del

			experto D)
	Limitaciones identificadas en los materiales	Valoración de los expertos con respecto a los elementos de los materiales que presentan oportunidades de mejora.	1. Desajuste entre el nivel propuesto y el contexto educativo. 2. Planteamiento del ejercicio de causa-efecto. “se supone que este material estaría enfocado en noveno, que según el currículo nacional dice que tiene que tener un B1, pero pues sabemos que en realidad no es así” (Entrevista del experto E) “no me parece que a los estudiantes les suene tan divertido [el ejercicio de causa-efecto]” (Entrevista del experto C)
	Propuestas para el fortalecimiento de los materiales	Recomendaciones de los expertos para optimizar los materiales elaborados por ChatGPT-4.	1. Incorporación de recursos multimodales para dinamizar el material. 2. Clarificación de orientaciones e instrucciones de trabajo. 3. Inclusión de preguntas orientadas a la opinión personal y al contexto cercano. “¿Sabes que les pondría yo? De pronto el audio, como en la pronunciación, que ellos lo puedan escuchar también” (Entrevista del experto F) “Yo creo que ahí hace falta como un tipo de pregunta de comprensión de lectura [...] que tienen que ver sobre tú qué piensas sobre lo que el texto está diciendo” (Entrevista del experto B)
Alineación con los estándares de competencia lectora en Inglés	Coherencia normativa de los materiales.	Aspectos relacionados con la coherencia tanto normativa como pedagógica de los materiales diseñados en este estudio.	1. Alineación con documentos oficiales. 2. Pertinencia de las temáticas alineadas al nivel B1. 3. Resignificación de los Derechos Básicos del Aprendizaje “Me pareció que todo estaba como muy apropiado y de acuerdo a sus objetivos” (Entrevista del experto B).

			(DBA).	
	Tensiones entre normativa y contexto colombiano	Correspondencia entre los lineamientos establecidos en los documentos oficiales y su adecuación al contexto educativo colombiano.	1. Influencia de referentes externos. 2. Limitaciones de los estándares frente a la realidad educativa nacional.	“En mi realidad oficial no podemos decir que en grado noveno yo tenga B1” (Entrevista del experto C)
	Limitaciones pedagógicas del material en la práctica	Aspectos asociados a las limitaciones pedagógicas que presentaron los materiales en la implementación. Esto último tratándose de la iniciativa de algunos expertos.	1. Diversidad de niveles de competencias en los contextos educativos. 2. Necesidad de articulación en los materiales con los DBA.	“No veo cómo el material tenga en cuenta o incluya a los estudiantes que no han llegado a ese nivel, porque los hay” (Entrevista del experto E)
Adecuación cultural al contexto colombiano	Requerimiento de mayor contextualización a la realidad local	Valoraciones de los expertos que indican que los materiales evidencian limitaciones en cuanto a la relación con el contexto cultural propiamente colombiano.	1. Referenciar el contexto local en los ejercicios propuestos. 2. Enfoque globalizado de los contenidos.	“[los materiales] no hablan pues como un contexto cultural tan específico, pero yo creo que eso se podría lograr [...] a través de las preguntas y de las actividades que se propongan” (Entrevista del experto B)
	Pertinencia de los temas abordados para estudiantes de noveno grado	Valoraciones de los expertos sobre la relevancia de las temáticas de los materiales para estudiantes de noveno grado.	1. Temas relacionados con la experiencia personal. 2. Temas sociales contemporáneos.	“Los contenidos creo que son acordes para los estudiantes” (Entrevista del experto F)
Fortalezas y limitaciones	Contribuciones percibidas del uso	Aspectos que evidencian las	1. Diseño de actividades de aula	“Para los profes, [la IA] les facilita

de ChatGPT en la elaboración de materiales	de la IA en el contexto educativo	ventajas del uso de la IA en contextos educativos, según los expertos.	mediadas por inteligencia artificial. 2. Adaptación de textos al nivel y al contexto educativo. 3. Generación de preguntas a partir de textos. 4. Apoyo a la elaboración de instrumentos de evaluación. 5. Optimización del tiempo.	<i>mucho la revisión de materiales o, por ejemplo, la creación, pero partiendo como de materiales auténticos”</i> (Entrevista del experto D) <i>“Sí, yo creo que es útil en muchos aspectos. Primero porque ayuda a agilizar el trabajo”</i> (Entrevista del experto F)
	Limitaciones percibidas del uso de la IA en el contexto educativo	Aspectos limitantes identificados por los expertos con relación al uso de la IA en contextos educativos.	1. Falta de conocimiento contextual por parte de la IA. 2. Necesidad de pensamiento crítico por parte del docente. 3. Posible despersonalización de los textos adaptados y materiales generados por la IA. 4. Riesgo de facilismo para los estudiantes. 5. Consideraciones éticas en el uso de la IA.	<i>“Riesgo en cuanto a evaluación también, el estudiante también puede llegar esto, subir el taller y que la IA lo responda”</i> (Entrevista del experto A) <i>“[La IA no tiene] ese conocimiento más concreto del contexto y de las condiciones materiales pues que sí tiene el profesor.”</i> (Entrevista del experto B)

3.6. Validez y confiabilidad

Golafshani (2003) afirma que "la calidad de una investigación está relacionada con la generalización de los resultados y, por lo tanto, con la comprobación y el aumento de la validez o fiabilidad de la investigación" (p. 603, nuestra traducción). En constancia con esta perspectiva, y con el fin de garantizar la solidez de este estudio cualitativo, se implementaron procedimientos específicos para demostrar tanto su validez como su fiabilidad. La validez se entiende como el conjunto de estrategias utilizadas para comprobar la precisión de los hallazgos y credibilidad de los

datos obtenidos, mientras que la confiabilidad se refiere a la consistencia de los datos recopilados y de los instrumentos utilizados en el proceso (Creswell & Creswell, 2018). En este sentido, el estudio presenta una descripción exhaustiva de la metodología y de los procedimientos usados. Asimismo, se empleó la rúbrica y la entrevista como instrumentos para obtener una evaluación detallada de los resultados obtenidos y verificar la información reportada. Además, se recurrió a la triangulación como técnica para demostrar la credibilidad de los hallazgos obtenidos (Bans-Akutey & Tiimub, 2021). Al comparar los documentos oficiales (MCER, ESC y DBA), la entrevista y las valoraciones de la rúbrica, se logró una comprensión más amplia y profunda de la evaluación realizada por los expertos sobre los materiales elaborados con ChatGPT-4, enriqueciendo así el análisis y garantizando una exploración sólida. Finalmente, el estudio incluye una documentación detallada del perfil de los expertos, así como de sus respuestas en la rúbrica y transcripciones de sus entrevistas, permitiendo un análisis riguroso de los datos y contribuyendo al fortalecimiento de la validez y confiabilidad del estudio.

3.7. Consideraciones éticas

En cuanto a las consideraciones éticas, es esencial crear un entorno seguro y respetuoso para los participantes y proporcionarles información clara sobre el estudio para asegurarse de una participación informada y voluntaria (Mirza et al., 2023). En ese sentido, se mantuvo una relación amigable y respetuosa con todos los participantes, quienes fueron informados detalladamente sobre el objetivo del estudio antes de proceder. En la primera encuesta, se les solicitó que dieran su consentimiento para ser parte de la investigación, por lo que sus aportes fueron totalmente voluntarios. Sus identidades fueron protegidas mediante el anonimato y los datos obtenidos no serán utilizados para fines distintos a los establecidos en el estudio. Del mismo modo, las grabaciones de las entrevistas se mantendrán confidenciales. Cabe mencionar que el uso de herramientas de IA en este estudio también supuso la consideración de las implicaciones éticas relacionadas a la adaptación y generación de materiales didácticos. Se comunicó explícitamente que los textos utilizados fueron adaptados y se incluyó el enlace a la fuente original. Asimismo, a lo largo del estudio, se mantuvo la transparencia sobre el uso de la IA en cuanto a la creación de los materiales y se garantizó que el único propósito de los materiales fuera académico e investigativo, sin fines de comercialización.

Este capítulo presentó la metodología desarrollada a lo largo de este estudio de carácter cualitativo y descriptivo, cuyo objetivo se trata de la validación de materiales de comprensión lectora para nivel B1 generados por ChatGPT mediante el juicio expertos en el área del inglés como lengua extranjera. Los métodos de recolección de datos y análisis incluyeron el uso de una rúbrica de evaluación y entrevistas semiestructuradas con cada participante con el fin de obtener las

percepciones de los expertos sobre los materiales. Los procedimientos se llevaron a cabo en tres etapas, primero con el diseño de los prompts y de los materiales, seguido del diseño de los instrumentos y, por último, con las entrevistas. Las consideraciones éticas incluyeron el uso responsable de la IA y el consentimiento informado de los participantes. El método de triangulación para comparar los datos de todos los instrumentos, aseguran que los resultados de esta investigación sean válidos y fiables.

En el siguiente capítulo se muestra el análisis de los resultados, examinando los datos de la rúbrica y de la entrevista para dar respuesta a cada una de las preguntas de investigación.

4. ANÁLISIS DE RESULTADOS

Este capítulo presenta el análisis de los resultados del estudio, integrando la discusión de los hallazgos obtenidos. La investigación lleva a cabo un proceso de validación de materiales de comprensión lectora para nivel B1 de inglés generados por ChatGPT-4, mediante la aplicación de una rúbrica de evaluación y la realización de entrevistas semiestructuradas a expertos en la enseñanza del inglés como lengua extranjera. En este sentido, los resultados permiten explorar el potencial de la IA como herramienta para la creación de materiales didácticos. Los resultados se organizan en cinco secciones, las primeras cuatro respondiendo a las preguntas de investigación planteadas. La primera se centra en la relevancia pedagógica de los materiales para el nivel B1, considerando aspectos como el vocabulario, sintaxis y la adecuación del contenido. La segunda parte aborda la alineación de los materiales con los estándares pedagógicos establecidos para ese nivel. La tercera sección analiza la pertinencia cultural de dichos materiales en relación con el contexto colombiano. La cuarta sección examina las fortalezas y limitaciones de la IA en la producción de materiales didácticos, según las percepciones de los expertos. Por último, la quinta sección corresponde a los resultados de la implementación de los materiales en un contexto educativo real, factor que en un principio no fue tenido en cuenta para este estudio, pero fue una iniciativa de algunos participantes que proporcionaron retroalimentación sobre su aplicación.

4.1. Pertinencia pedagógica de materiales de lectura de ChatGPT para nivel B1-grado noveno

4.1.1. Resultados

Teniendo en cuenta que los materiales generados para esta investigación fueron diseñados con base en parámetros específicos del nivel B1 de inglés, es crucial que estos se ajusten pedagógicamente a las competencias lingüísticas y lectoras que pretenden desarrollar. Es así que, considerando aspectos como el vocabulario, la sintaxis y la pertinencia del contenido, se propuso una rúbrica de validación que incluía diversas preguntas enfocadas en explorar qué tan adecuadamente se reflejan estos elementos en los materiales. Dichas preguntas se distribuyeron en cuatro criterios a evaluar: *adecuación al nivel B1*, *variedad y calidad de los textos*, *variedad, calidad y precisión de los ejercicios* y *calidad de los niveles de comprensión lectora*. Cada uno de ellos cuestionó diferentes dimensiones, condensándose en componentes que fueron valorados por los expertos.

La tabla 4.1 muestra el criterio de adecuación al nivel B1, un aspecto importante a considerar en el proceso de validación, ya que así se asegura que los materiales sean apropiados para el desarrollo de habilidades lectoras de estudiantes de este nivel. En ese sentido, los componentes evaluados se relacionan con *estrategias de lectura B1*, *inclusión de vocabulario para alumnos de 9° grado (nivel B1)* y *dificultad manejable para los estudiantes*.

Tabla 4.1*Adecuación al nivel B1*

Componentes	Experto A	Experto B	Experto C	Experto D	Experto E	Experto F	<i>Promedio por componente</i>
Estrategias de lectura B1	5	4	5	5	4	5	4,6
Inclusión de vocabulario para alumnos de 9º grado (nivel B1)	4	5	3	5	4	5	4,3
Dificultad manejable para los estudiantes	5	5	3	5	5	5	4,6
<i>Promedio por Experto</i>	4,6	4,6	3,6	5	4,3	5	4,5

A grandes rasgos, los resultados de la rúbrica en lo referente al criterio de adecuación al nivel B1 presentan una recepción favorable, pues el promedio global (4,5) refleja una puntuación cercana a 5. Por componentes, se evidencia que en cuanto a *estrategias de lectura B1* se obtuvo una cifra excelente de 4,6, considerando que solo los expertos B y E asignaron una valoración de 4, mientras que los demás otorgaron la puntuación máxima de 5. De manera similar, el componente de *dificultad manejable para los estudiantes* presenta un promedio excelente de 4,6, siendo el experto C el único que le asignó una calificación aceptable de 3 sobre 5. El promedio más bajo registrado por componente, correspondiente a la *inclusión de vocabulario para alumnos de 9º grado (nivel B1)*, fue de 4,3, aún así manteniéndose en un rango de valoración excelente. En este componente se muestran valoraciones ligeramente más variadas, tres de los expertos (B, D y F) asignaron la calificación máxima en la escala de 5, dos (A y E) coincidieron en una puntuación de 4 y solo el experto C otorgó una estimación aceptable de 3. En lo que concierne al promedio por experto, solo C evidencia un resultado aceptable (3,6), a diferencia del resto, que muestra un promedio de valoración excelente.

Se continúa con la tabla 4.2, que presenta la validación de la variedad y calidad de los textos adaptados por ChatGPT. En esta sección se evaluó si las lecturas son apropiadas en términos de contenido, estructura, coherencia, fluidez y adecuación al nivel B1, de modo que favorezcan una comprensión adecuada y el desarrollo de la comprensión lectora pertinente. De esta forma, los componentes validados consideran los aspectos de *coherencia del texto y fluidez natural, adecuación de la longitud del texto (nivel B1), convenciones de género y tipo de texto, estructura lógica del párrafo (textos argumentativos) y organización y equilibrio de párrafos*.

Tabla 4.2*Variedad y calidad de los textos*

Componentes	Experto A	Experto B	Experto C	Experto D	Experto E	Experto F	Promedio por componente
Coherencia del texto y fluidez natural	4	5	5	5	5	5	4,8
Adecuación de la longitud del texto (nivel B1)	5	5	3	5	5	5	4,6
Convenciones de género y tipo de texto	4	5	5	5	4	5	4,6
Estructura lógica del párrafo (textos argumentativos)	4	5	5	5	3	5	4,5
Organización y equilibrio de párrafos	4	5	5	5	4	5	4,6
Promedio por Experto	4,2	5	4,6	5	4,2	5	4,6

En esta sección, la validación realizada por los expertos demuestra una percepción general positiva, reflejada en un promedio global excelente de 4,6. En cuanto a los componentes evaluados, la *coherencia del texto y fluidez natural* obtuvo un promedio muy alto de 4,8, con valoraciones de 5 por la mayoría de los expertos; únicamente el experto A otorgó una calificación de 4. El componente de *adecuación de la longitud del texto (nivel B1)* presenta un promedio excelente de 4,6, donde la mayoría de los expertos evaluó con 5, a excepción del experto C, quien evaluó con 3. Por su parte, *las convenciones de género y tipo de texto* alcanzaron un promedio de 4,6, con valoraciones excelentes de 5, excepto por los expertos A y E, quienes calificaron con 4. Respecto a la *estructura lógica del párrafo (textos argumentativos)*, se obtuvo un promedio positivo de 4,5, donde cuatro expertos evaluaron con 5, el experto A con 4 y el experto E con 3. Finalmente, el componente de *organización y equilibrio de párrafos* registró un promedio excelente de 4,6, con valoraciones de 5 por la mayoría de expertos, mientras que los expertos A y E otorgaron 4. Con respecto a la valoración por expertos, los expertos B, D y F obtuvieron un promedio de 5, seguidos del experto C con 4,6, y los expertos A y E con 4,2.

A continuación, considerando la importancia de establecer diversas actividades que sean coherentes con lo que se intenta desarrollar en los materiales, la tabla 4.3 presenta el criterio de *variedad, calidad y precisión de los ejercicios*, organizado en componentes que incluyen *claridad*

de las instrucciones, variedad de tipos de tareas, activación de la estrategia de prelectura, desarrollo de estrategias durante la lectura, desarrollo de la estrategia posterior a la lectura, consistencia entre el texto y la respuesta, alineación entre las tareas y las habilidades evaluadas y calidad y plausibilidad de las opciones de opción múltiple.

Tabla 4.3

Variedad, calidad y precisión de los ejercicios

Componentes	Experto A	Experto B	Experto C	Experto D	Experto E	Experto F	Promedio por componente
Claridad de las instrucciones	5	5	5	5	4	5	4,8
Variedad de tipos de tareas	5	5	5	5	4	5	4,8
Activación de la estrategia de prelectura	5	4	4	5	4	5	4,5
Desarrollo de estrategias durante la lectura	5	4	5	5	5	5	4,8
Desarrollo de la estrategia posterior a la lectura	5	5	4	5	5	5	4,8
Consistencia entre el texto y la respuesta	3	5	5	5	5	5	4,6
Alineación entre las tareas y las habilidades evaluadas	4	5	5	5	5	5	4,8
Calidad y plausibilidad de las opciones de opción múltiple	5	5	3	5	5	5	4,6
Promedio por Experto	4,6	4,7	4,5	5	4,6	5	4,7

En términos generales, el promedio global excelente de 4,7 muestra que las percepciones del criterio sobre la *variedad, calidad y precisión de los ejercicios* son positivas. Con relación al promedio por componente, se evidencia en *claridad de las instrucciones* y *variedad de tipos de tareas* un patrón de evaluación similar, ya que en ambos casos se obtuvo un promedio excelente de 4,8 y el mismo experto (E) asignó una calificación de 4, mientras que los demás atribuyeron a estos aspectos una valoración máxima de 5. En lo referente a las fases de lectura, se observa que la

activación de la estrategia de prelectura consigue un promedio de 4,5, en el que tres de los expertos (B, C y E) otorgan una puntuación de 4 sobre 5. En el *desarrollo de estrategias durante la lectura* se destaca una valoración de 4,8, correspondiendo al experto B la única calificación de 4. Del mismo modo, el componente de *desarrollo de la estrategia posterior a la lectura* obtuvo un promedio de 4,8 y, a diferencia del anterior, fue el experto C quien asignó un 4 como calificación. En cuanto a la *consistencia entre el texto y la respuesta*, dado que el experto A concedió a este componente un 3 aceptable, en contraste con los otros expertos, quienes lo valoraron con 5, el promedio resultó en 4,6. En el componente de *alineación entre las tareas y las habilidades evaluadas*, se presenta un resultado de 4,8, siendo el experto A quien le asignó una puntuación de 4. Finalmente, en *plausibilidad de las opciones de opción múltiple* se destaca, nuevamente, un promedio de 4,6 y sólo el experto C concedió una valoración aceptable de 3 sobre 5.

Cabe destacar que, en la sección de comentarios de la rúbrica, el experto A manifiesta que las actividades propuestas ordenadas a lo largo de las fases de pre-lectura, durante la lectura y la post-lectura son variadas, versátiles y de notable calidad. Sin embargo, el experto B agrega que se pueden realizar algunos ajustes en los materiales, particularmente en la fase de prelectura y en el nivel creativo, ya que considera que se pueden formular preguntas más específicas:

Me parece que las preguntas más abiertas (como las de creatividad al final) pueden ser más concretas. Lo mismo con las preguntas de pre-reading. Creo que a veces es insuficiente preguntar sobre qué crees que va a hablar el texto. Es necesario hacer preguntas más específicas y quizás trabajar con algo de vocabulario antes de leer.
(Comentario escrito por el experto B en la rúbrica de validación)

La tabla 4.4 aborda la calidad de los niveles de comprensión lectora. Esta sección tiene como objetivo validar si los materiales cumplen adecuadamente con los distintos niveles de comprensión, que incluyen el nivel literal, inferencial, el crítico y el de comprensión creativa, así como una progresión de los ejercicios entre estos niveles. De esta forma, los componentes evaluados comprenden el *rango de niveles de comprensión lectora*, la *comprensión literal progresiva*, las *habilidades de lectura inferencial*, las *habilidades de lectura crítica*, la *comprensión creativa y aplicada*, y la *progresión de dificultad a través de los niveles*.

Tabla 4.4*Calidad de los niveles de comprensión lectora*

Componentes	Experto A	Experto B	Experto C	Experto D	Experto E	Experto F	Promedio por componente
Rango de niveles de comprensión lectora	5	5	5	5	5	5	5
Comprensión literal progresiva	5	4	4	5	4	5	4,5
Habilidades de lectura inferencial	5	5	5	5	4	5	4,8
Habilidades de lectura crítica	5	5	5	5	5	5	5
Comprensión creativa y aplicada	5	4	5	5	5	5	4,8
Progresión de dificultad a través de los niveles	5	5	5	5	5	5	5
Promedio por Experto	5	4,6	4,8	5	4,6	5	4,8

Como se visualiza en la tabla 4.4, la valoración de los expertos presenta un promedio global excelente de 4,8 con respecto a esta sección. En relación a los componentes evaluados, el *rango de niveles de comprensión lectora* alcanzó un promedio perfecto de 5, ya que todos los expertos otorgaron dicha valoración. El componente de *comprensión literal progresiva* registra un resultado positivo de 4,5, con valoraciones de 5 por parte de la mitad de los expertos (A, D y F), mientras que los expertos B, C y E asignaron una calificación de 4. Respecto a las *habilidades de lectura inferencial*, se obtuvo un promedio alto de 4,8, donde la mayoría de los expertos que calificaron con 5, a excepción del experto E, quien otorgó una valoración de 4. Por su parte, las *habilidades de lectura crítica* reflejaron un promedio perfecto de 5, con una valoración unánime de todos los expertos. El componente de *comprensión creativa y aplicada* muestra un promedio de 4,8, donde la mayoría de los expertos evaluó con 5, a excepción del experto B, quien evaluó con 4. Finalmente, la *progresión de dificultad a través de los niveles* alcanzó un promedio de 5, con una valoración excelente de 5 por parte de todos los expertos. Con respecto a la valoración por expertos, los expertos A, D y F obtuvieron un promedio de 5, seguidos del experto C con 4,8, y de los expertos B y E con 4,6.

En esencia, las valoraciones aportadas por los expertos en la rúbrica de validación de los materiales señalan una recepción positiva en los criterios relacionados con la pertinencia pedagógica para nivel B1 (equivalente a noveno grado). Es así que se procede a destacar los

resultados obtenidos en las entrevistas, las cuales favorecieron la profundización de las calificaciones proporcionadas en la rúbrica. Siguiendo el enfoque de análisis temático presentado por Braun y Clarke (2012), los datos se examinaron de manera detallada para identificar patrones recurrentes en las respuestas de los expertos, prestando especial atención a las percepciones en común. De esta forma, se lograron reconocer temas frecuentes que reflejan aspectos relevantes relacionados a las preguntas de investigación. A partir de este análisis se establecieron las siguientes categorías: *concepciones docentes sobre un material didáctico de comprensión lectora adecuado, calidad del diseño de los materiales generados para el estudio, limitaciones identificadas en los materiales y propuestas para el fortalecimiento de los materiales*, las cuales se describen a continuación.

A. Concepciones docentes sobre un material didáctico de comprensión lectora adecuado

Esta categoría corresponde a las percepciones y criterios de los docentes expertos sobre lo que consideran fundamental al elaborar o seleccionar materiales para el desarrollo de habilidades de comprensión lectora en inglés. De modo que lo expresado por los expertos se distribuye en tres subcategorías: la *estructuración del material de comprensión lectora*, el *desarrollo de habilidades de comprensión lectora* y las *características de los textos en términos de interés y complejidad*. Cada una de ellas evidencia las concepciones y experiencias de los expertos sobre la manera y el tipo de recursos que han empleado en el aula.

Iniciando con la estructuración del material de comprensión lectora, los expertos A y F señalaron que un material adecuado que busque fortalecer las habilidades de lectura debe organizarse de manera progresiva, comenzando con elementos más sencillos y cercanos a los estudiantes para después avanzar hacia actividades más complejas que fomenten el pensamiento crítico. El experto F, en particular, destacó la importancia de un vocabulario comprensible, pero también desafiante, de manera que no cause frustración en el estudiante:

Yo pensaría que primero, [el material] debe explorar los conocimientos del estudiante a partir de un contexto que sea familiar o alguna manera que lo lleve a ese contexto de una manera natural, un vocabulario que sea alcance del estudiante o que por lo menos lo rete, pero que no lo desanime, porque de eso se trata, de que incentive el conocimiento pero que no lo frustre tanto como para que lo deje sin ganas de seguir aprendiendo. (Entrevista del participante F, 2025).

En otras palabras, los expertos manifestaron que para lograr una secuencia apropiada, se debe partir del conocimiento previo del estudiante y con contenidos familiares para ellos y luego proseguir con ejercicios que gradualmente lo lleven a un proceso de reflexión.

En lo referente al desarrollo de habilidades de comprensión lectora, se encontró que los expertos C y D consideran que un material adecuado debe fortalecer las habilidades literales e inferenciales, entendiéndose como la comprensión de ideas globales y específicas, la intención del autor y el propósito del texto: “Yo pienso que todo material de comprensión lectora tiene que tener un tipo de lectura literal, un tipo de lectura inferencial, algo que haga que los estudiantes propongan por sí mismos para crear un nuevo texto” (Entrevista del participante C, 2025).

Es así que, para dichos expertos, es importante que los materiales didácticos de lectura propongan preguntas de nivel inferencial y literal, así como aquellas que promuevan el pensamiento crítico y que, de esta forma, los estudiantes adquieran la capacidad de analizar, proponer y producir ideas de manera autónoma.

Haciendo referencia a las características de los textos en términos de interés y complejidad, los expertos D y F coincidieron en que los textos que se aborden para este tipo de materiales se deben seleccionar considerando dos aspectos esenciales. Por un lado, los contenidos y temáticas deben ser acordes a los intereses de los estudiantes y, por otro lado, la complejidad de los textos debe ajustarse al nivel requerido, evitando vocabulario muy especializado: “Ahí lo que define la complejidad es el tema, que no sea, pues, como muy específico de una disciplina, y también, pues, el vocabulario, la complejidad del vocabulario” (Entrevista del participante D, 2025).

Los expertos señalan, entonces, un equilibrio entre interés y complejidad. Los contenidos de los textos usados en los materiales necesitan captar la atención de los estudiantes, y a su vez, deben ser apropiados para su competencia lingüística en cuanto a nivel y temática.

B. Calidad del diseño de los materiales generados para el estudio

Esta categoría se enfoca en las estimaciones realizadas por los expertos respecto a la organización y características de los materiales didácticos de nivel B1 generados por ChatGPT-4 en el marco de esta investigación. Con base en lo comentado durante la entrevista, se consiguió identificar diversas subcategorías: *valoración de los ejercicios propuestos, estructuración del material a través de las fases de lectura, manejabilidad de la complejidad temática y relevancia de los contenidos para el contexto estudiantil objetivo*. Se consideró, igualmente, incluir una subcategoría llamada *estrategias docentes en la implementación del material* destinada a describir la aplicación de uno de los materiales generados, factor que fue iniciativa de algunos participantes. Es así que las percepciones de los expertos posibilitan evaluar la adecuación pedagógica de los materiales de nivel B1 generados por IA.

En primera instancia, en la valoración de los ejercicios propuestos, todos los expertos coincidieron en que las actividades planteadas en los materiales eran pertinentes para el nivel B1. Destacaron que las preguntas de opción múltiple, falso/verdadero y las relacionadas al reconocimiento de vocabulario clave, especialmente, estaban bien elaboradas:

Todas, todas [las actividades] son relevantes. Una comprensión de lectura se debe evaluar con diferentes tipos de preguntas: abiertas, cerradas, selección múltiple, falso, verdadero. Entonces, una variedad de estrategias para identificar la comprensión de los estudiantes. O sea, todas son relevantes, todas. No hubo ninguna que fuera irrelevante o que no fuera relacionada con la comprensión de lectura, todas eran necesarias. (Entrevista del participante D, 2025).

Los expertos, por lo tanto, consideraron que las actividades están adecuadamente formuladas y que permiten desarrollar de manera efectiva la comprensión lectora en inglés de los estudiantes de nivel B1.

Prosiguiendo con la subcategoría de estructuración del material a través de las fases de lectura, algunos expertos (D, E y F) encontraron pertinente que se incluyeran las etapas de pre-lectura, durante la lectura y post-lectura dentro de los materiales y, según sus aportaciones, la organización de las preguntas y ejercicios a partir de las fases mencionadas están adecuadamente desarrolladas: "Todos esos materiales tenían bien las fases de la lectura, lo que era el pre, while y post-task." (Entrevista del participante D, 2025).

Así pues, los expertos señalaron que la estructuración del material en fases de lectura está bien lograda y resulta adecuada dentro de los materiales generados.

Con respecto a la manejabilidad de la complejidad temática, los expertos B y D coincidieron en que la dificultad de los temas abordados, la sintaxis empleada y el vocabulario utilizado en los materiales eran comprensibles y se ajustaban a las habilidades lingüísticas de los estudiantes de nivel B1:

Yo ahí pues en los textos vi que había como un vocabulario suficientemente amplio pero no tan complicado y no había estructuras sintácticas complejas, ni digamos como descripciones así como con demasiadas imágenes o figuras literarias sino que era bastante directo el estilo y eso me parece que pues que es como adecuado para ese nivel. (Entrevista del participante B, 2025).

De este modo, algunos expertos señalaron que los temas, la sintaxis y el vocabulario empleado en el contenido de los materiales son adecuados para el nivel B1 de inglés y que pueden ser abordados por estudiantes de este nivel sin mucha dificultad.

En lo que concierne a la relevancia de los contenidos para el contexto estudiantil objetivo, las percepciones de los expertos C y D resultan positivas. Ambos destacaron que los contenidos propuestos en los materiales son relevantes e interesantes para los estudiantes de noveno grado, correspondiente al nivel B1:

Entonces, el hecho de que ustedes eligieran temas muy de moda, muy de moda, muy que les interesa a los muchachos, hace que la parte de la producción de pronto no les dé, o sea, les parezca interesante producir, crear su propio texto. Entonces, he ahí la importancia en la selección de las temáticas que me parece que estuvieron excelentes. (Entrevista del participante C, 2025).

La valoración de estos expertos evidencia que los contenidos y temáticas de los textos seleccionados responden al contexto e interés de los estudiantes, lo que hace que se involucren con más disposición en las actividades.

Ahora bien, es importante resaltar que algunos expertos (A y C) pusieron en práctica los materiales generados por ChatGPT-4 en el aula de inglés. Por lo que la subcategoría de estrategias docentes en la implementación del material corresponde a la manera en la que llevaron a cabo su aplicación. Ambos expertos trabajaron en el taller sobre la reseña de la película *Her*. Por una parte, el experto A inició con la contextualización del tema a través del tráiler de la película y posteriormente adaptó el formato del taller, optando a su vez por el trabajo en grupo. Al final incluyó una sección de comentarios para que los estudiantes dieran retroalimentación:

[...]en este taller yo por ejemplo, hice modificaciones, entonces ahí descargué el taller y a la hora de ver el taller, entonces pues primero puse la fecha, el nombre del estudiante, los códigos y segundo, yo entiendo que tal vez ustedes estaban proyectando para un solo estudiante con un nivel B1 en un colegio público con muchos estudiantes, pero yo dije, no, yo no puedo hacer esto de forma individual, tiene que ser de forma colectiva.

Y a lo último les agregué una pregunta también me parece como buena que sea como de realimentación, yo hice todo el taller, yo lo completé todo esto, pero yo quisiera saber algunos comentarios, por ejemplo, de la actividad, '¿Qué te pareció?' '¿Fue difícil?' '¿Fue fácil?' y que eso también sirva como una realimentación sobre el material. (Entrevista del participante A, 2025).

Cabe destacar que los estudiantes que desarrollaron el taller con el experto A manifestaron que una de las preguntas tipo *matching* (emparejar) resultó ser ambigua, puesto que consideraron que se podía seleccionar más de una opción como respuesta correcta.

Por otra parte, el experto C implementó el material de manera individual, enfocándose en la traducción del texto y del vocabulario desconocido para el estudiante, sin embargo, no se proporcionó las percepciones del alumnado:

Después, en mi metodología, es que cada uno tiene que confrontarse con su vocabulario desconocido. Entonces, de manera individual, ya ellos pasan al cuaderno y yo los hago traducir el texto de oración por oración. ¿Por qué? Porque es un recurso que uso evitando que Google Traductor sea quien haga la tarea. Entonces yo les digo: el texto lo van a trabajar en el cuaderno, de punto a punto, oración a oración. Lo copias en tu cuaderno, cada oración; te enfrentas a la oración despacio. Después mira dentro de esa oración cuáles fueron las palabras desconocidas para ti. Ahí mismo escribes tu vocabulario, yo le llamo glosario, ¿cierto? (Entrevista del participante C, 2025).

Según lo aportado por los expertos A y C, se evidencia que los materiales pueden ser utilizados en contextos educativos reales, ya sea a través del trabajo colectivo o individual. No obstante, de acuerdo con los estudiantes, aún hay algunas áreas en las que se pueden realizar ajustes.

C. Limitaciones identificadas en los materiales

En esta categoría se recogen las percepciones de los expertos sobre los elementos que podrían mejorarse en los materiales diseñados. En este sentido, se identificaron dos subcategorías enfocadas en el *desajuste entre el nivel propuesto y el contexto educativo* y el *planteamiento del ejercicio de causa-efecto* que se muestra en los materiales. Es posible que las aportaciones de los expertos permitan reflexionar sobre los ajustes que propicien la optimización del material.

Respecto al desajuste entre el nivel propuesto y el contexto educativo, algunos expertos (A, C, E y F) señalaron que, aunque los materiales están planteados para noveno grado, quienes según los documentos nacionales deberían encontrarse o buscar desarrollar el nivel B1 de inglés, esta equivalencia no siempre se refleja en la realidad educativa. Desde su experiencia, el experto E incluso indicó que para estudiantes universitarios, los materiales podrían también presentar cierta dificultad. No obstante, el experto F, manifestó que el nivel de los estudiantes puede variar según la institución educativa, ya que en algunos colegios si es posible encontrar estudiantes de noveno con nivel B1. Aún así, en general coincidieron en que no existe una alineación directa entre noveno grado y lo estipulado en los documentos sobre dicho nivel:

El 87% , te puedo decir el 90 % de los estudiantes de grado 11 en el sector oficial salen con A2. Entonces, cuando ustedes me dicen esa pregunta, '¿estos materiales pertenecen a B1?' y entre paréntesis me escriben grado noveno, yo digo: yo no les puedo decir que en grado noveno hay un B1 en mi colegio, con mi realidad. Y yo llevo

en ese colegio 20 años y mis respuestas están basadas en mi realidad de 20 años del mismo colegio. (Entrevista del participante C, 2025).

Es así que, según algunos expertos, se evidencia una brecha entre la realidad educativa de noveno grado y el nivel de competencia lingüística asignado a estos estudiantes en los documentos nacionales. Aunque en algunos casos particulares puede encontrarse un nivel B1, la experiencia de los expertos demuestra que esto no puede asumirse en todo momento.

Continuando con otras de las limitaciones señaladas por los expertos, algunos de ellos, en concreto los expertos B y C, señalaron un inconveniente con el planteamiento del ejercicio de causa-efecto. Ambos expertos reconocieron que este tipo de ejercicios puede no resultar muy pertinente. El experto C, por ejemplo, mencionó que la realización de ejercicios con relaciones causales no es motivadora para los estudiantes. Asimismo, el experto B manifestó que este tipo de ejercicios tiende a simplificar los argumentos del texto y volverlos lineales. Este experto propuso como alternativa mostrar argumentos contruïdos incorrectamente para que los estudiantes lo corrijan:

Y si, no sé, yo ahorita como pensando en esas de las causas pensaba que quizás... como pensar en cosas como en lugar de ponerlas como completar, también se podrían presentar como, como argumentos erróneos, como para que ellos corrijan los argumentos, no sé, como algo como de... Como que a través de algo que está equivocado, que está mal contruïdo argumentativamente, que ellos hagan el ejercicio de corregirlo, pero la verdad no sé por qué me parece muy difícil hacer ese tipo de preguntas, pero esas son las que me parecen menos efectivas. (Entrevista del participante B, 2025).

De acuerdo con las respuestas de los expertos, este tipo de preguntas no captan la atención de los estudiantes, son difíciles de plantear y puede que no logren fortalecer efectivamente la comprensión lectora.

D. Propuestas para el fortalecimiento de los materiales

Esta categoría reúne las sugerencias de los expertos para la optimización de los materiales generados por ChatGPT-4 que se les presentaron. Así pues, se establecieron tres subcategorías enfocadas en las recomendaciones ofrecidas: *incorporación de recursos multimodales para dinamizar el material*, *clarificación de orientaciones e instrucciones de trabajo* e *inclusión de preguntas orientadas a la opinión personal y al contexto cercano*.

En lo que respecta a la incorporación de recursos multimodales para dinamizar el material, la mitad de los expertos (A, E y F) manifestaron la falta de inclusión de otros insumos, como

audios, imágenes, íconos o vídeos, que orienten el aprendizaje de la comprensión lectora dentro de los materiales, como se evidencia en el siguiente extracto:

En ese, un ejercicio de apoyo podría ser anexarle un enlace que remita a los estudiantes a un video, por ejemplo, cuyo contenido sea muy similar al de la lectura, no el mismo, sería una maravilla, cierto, que uno pudiera encontrar el mismo, pues, pero por lo menos similar al contenido de la lectura. O un enlace que lo lleve a uno a que los estudiantes vean el video, o si no que de pronto trabajen un audio, que ellos escuchen (Entrevista del participante E, 2025).

De esta forma, los expertos reconocen la importancia de incluir materiales complementarios que orienten la comprensión lectora, ya sea mediante enlaces o códigos QR, con el fin de ofrecer apoyos visuales y auditivos que faciliten la comprensión y hagan el material más inclusivo. Asimismo, los expertos A y F mencionaron el uso de otras herramientas de IA, como *Deepseek* y *Eleven Labs*, para la creación de insumos adicionales, tales como audios o textos, que apoyen el proceso de comprensión de los estudiantes.

En cuanto a la clarificación de orientaciones e instrucciones de trabajo, los expertos A y E, señalaron que es importante que los materiales incluyan indicaciones más precisas sobre el desarrollo de las actividades. El experto E mencionó la necesidad de especificar claramente si los ejercicios están pensados para realizarse de manera individual o grupal. El experto A, por su parte, manifestó que se podía añadir expresiones que el docente por lo general usa habitualmente en las clases:

Entonces eso también lo recomiendo como modificar de alguna manera, formar el prompt y darle como algunos insumos mucho más amplios, expresiones que el profesor podría también utilizar en su día a día como 'vamos', 'hagamos', 'hagámoslo' o 'ya llegaste hasta aquí, entonces continúa'. Esos elementos son importantes también. (Entrevista del participante A, 2025).

Los expertos, entonces, resaltan la importancia de incorporar instrucciones más claras referentes a la manera en la que se trabajan las actividades y que el material utilice expresiones que simulen el lenguaje cotidiano utilizado por el docente en el aula.

Por último, en la subcategoría relacionada a la inclusión de preguntas orientadas a la opinión personal y al contexto cercano, los expertos A y B destacaron la necesidad de incorporar preguntas que inviten a expresar opiniones, acuerdos o desacuerdos frente a lo leído y actividades que busquen relacionar el contenido con la realidad del estudiante. El experto B expresó que hace falta añadir preguntas que establezcan una conexión entre el texto y la reflexión del estudiante:

Entonces yo creo que ahí hace falta como un tipo de pregunta de comprensión de lectura que sea como un puente entre ese como entre la lectura y ya el salirse un poco del texto y son preguntas que tienen que ver sobre tú qué piensas sobre lo que el texto está diciendo, cuál es tu opinión, si estás de acuerdo, si no estás de acuerdo (Entrevista del participante B, 2025).

De esta forma, los expertos consideran que deben existir preguntas que trasciendan la comprensión literal del texto y fomenten la expresión de opiniones personales, para que así los estudiantes relacionen el contenido leído con sus propias experiencias y contextos.

A modo de síntesis, los resultados de la entrevista evidencian que los materiales de comprensión lectora generados por ChatGPT-4 presentan aspectos positivos, entre los cuales se destacan la adecuación de los ejercicios propuestos para evaluar la comprensión lectora, la estructuración clara del material a través de las fases de lectura, la complejidad temática manejable para estudiantes de nivel B1 y la relevancia de los contenidos para el contexto de noveno grado. Sin embargo, aún existen áreas de mejora. Se observó un desajuste entre el nivel B1 y la realidad educativa. Asimismo, se señalaron dificultades con ciertos tipos de ejercicios como los de causa-efecto, que pueden resultar poco motivadores o simplificar demasiado los argumentos. En lo referente a la optimización del material, los expertos sugirieron complementar con recursos multimodales, clarificar las instrucciones e incluir expresiones que dinamicen el material, así como incorporar preguntas que fomenten la opinión y el acercamiento de los estudiantes al texto.

4.1.2. Discusión

En el área del aprendizaje del inglés como lengua extranjera es importante determinar si los materiales didácticos son pedagógicamente pertinentes, puesto que su efectividad depende de la coherencia de estos con los objetivos de aprendizaje, el nivel lingüístico que pretenden medir y las características del contexto educativo en el que se implementan. Por ello, de forma similar a los estudios de los antecedentes (Andriyani et al., 2018; Cifuentes & Niño, 2010; Gutiérrez & Jurado, 2018; Nieves et al., 2013), se tuvieron en cuenta ciertos criterios pedagógicos claros y específicos para la evaluación de los materiales, lo que permitió determinar con mayor precisión, gracias a la valoración de los expertos, el cumplimiento de los objetivos propuestos dentro del dominio de la comprensión lectora en una lengua extranjera. En ese sentido, el proceso de validación de materiales adquiere una función esencial, permitiendo sustentar dicha pertinencia a partir de evidencias concretas (Weir, 2005). En el presente estudio, este proceso de validación se apoya en la evidencia basada en contenidos, entendida como el grado en que estos materiales de comprensión

lectora mantienen una alineación con sus propósitos pedagógicos y representan las habilidades lingüísticas que se busca desarrollar en el nivel B1 (Arias & Sireci, 2021). Es así que esa evidencia se obtiene en el análisis de componentes lingüísticos como el vocabulario empleado, la complejidad sintáctica y la adecuación del contenido temático. Asimismo, teniendo en cuenta que el procedimiento de validación de contenido se fortalece mediante el juicio de expertos (Escobar-Pérez & Cuervo-Martínez, 2008), para la presente investigación se considera fundamental que la valoración del contenido sea realizada por personas con conocimiento especializado en la enseñanza del inglés como lengua extranjera, contribuyendo, con base en criterios pedagógicos y lingüísticos informados, a que los materiales puedan ser considerados válidos.

De acuerdo con los resultados obtenidos, los materiales de comprensión lectora generados por ChatGPT-4 muestran una pertinencia pedagógica acorde al nivel B1. En conjunto, las percepciones de los expertos proporcionadas en la rúbrica y en la entrevista muestran una valoración mayoritariamente positiva, especialmente en lo relacionado con la adecuación al nivel, la variedad y la calidad de los ejercicios y textos, la organización de los niveles de comprensión y la complejidad de las temáticas. Sin embargo, los hallazgos también revelan ciertas limitaciones ligadas a la realidad educativa.

Referente a la rúbrica de validación, se encuentra que el promedio global de 4,5 representa una recepción positiva en cuanto al criterio de adecuación al nivel B1. Los expertos, en su mayor parte, consideran que los materiales generados por ChatGPT-4 permiten desarrollar estrategias de lectura, incluyen vocabulario y proponen una dificultad manejable para el nivel B1, el cual, según el Esquema Curricular Sugerido (2016), corresponde al noveno grado. No obstante, en el promedio por experto se muestra una valoración que, aunque se mantiene en un rango favorable, es relativamente más baja en comparación con otros componentes evaluados. En concreto, el experto C, quien figura un promedio de 3,6, parece estimar que el vocabulario y la complejidad empleada en los materiales son solamente aceptables, lo que puede deberse a la experiencia exclusiva del experto en el nivel de bachillerato. A partir de su perspectiva como docente de secundaria, el experto ha evaluado estos materiales de acuerdo con las condiciones reales de sus estudiantes, quienes en muchos casos no han desarrollado completamente las habilidades del nivel B1 esperadas para noveno grado (factor que fue mencionado igualmente en la entrevista por parte del experto), por lo que el vocabulario y la dificultad percibida en los materiales puede resultar compleja para este contexto educativo en específico.

Por su parte, la variedad y calidad de los textos adaptados fueron valoradas de forma consistentemente alta por parte de los expertos, especialmente en el componente relacionado con la coherencia del texto y fluidez natural. Estos resultados reflejan que ChatGPT-4 presenta un potencial significativo para la adaptación textual, manteniendo la coherencia y corrección

gramatical sin afectar negativamente los componentes lingüísticos del texto. No obstante, se evidenció una discrepancia procedente de los expertos A, C y E, quienes también manifestaron en las entrevistas un desajuste entre el nivel propuesto y la realidad del contexto educativo. Estas percepciones sugieren que el contexto educativo desempeña un papel relevante en la valoración de los textos adaptados y demuestra que las capacidades de ChatGPT-4 reflejan ciertas limitaciones en la organización de los párrafos y en el manejo de las convenciones del tipo textual, aspectos fundamentales para favorecer la comprensión lectora. En este sentido, si bien la calidad general en las adaptaciones es adecuada, se identifican diferencias puntuales que podrían afectar la uniformidad del material, especialmente en textos de tipo argumentativo.

En lo que concierne a la variedad, calidad y precisión de los ejercicios, todos los expertos proporcionaron valoraciones de rango excelente. De acuerdo con la evaluación, los materiales son claros en sus instrucciones, disponen de una diversidad de tipos de tareas, activan y desarrollan estrategias de lectura en las fases de pre, durante y post-lectura, presentan una consistencia entre el texto y las actividades, demuestran alineación entre las tareas y las habilidades evaluadas, y muestran que los ejercicios de opción múltiple, en particular, son adecuados. De hecho, estos componentes también fueron valorados de forma positiva en la entrevista por parte de todos los expertos. Esto indica que ChatGPT-4 responde apropiadamente cuando, desde el prompt, se especifica claramente el tipo de ejercicios que se buscan generar. Sin embargo, es necesario tener presente que los expertos parecen identificar aspectos que requieren optimización, especialmente en lo que respecta a la activación de estrategias de pre-lectura, tratándose del único componente que fue consistentemente calificado con 4 por tres expertos (B, C y E). Por lo tanto, se puede entrever una falencia en el desarrollo de esta etapa, la cual consiste en la activación de conocimientos previos del estudiante (Napitupulu, 2018), sugiriendo que las especificaciones del prompt no fueron lo suficientemente detalladas en cuanto a la generación de actividades de pre-lectura más variadas, o bien que ChatGPT-4 tiende a realizar esta fase más superficial en comparación con las otras etapas. Aún así, el promedio global de 4.7 revela que esta herramienta, en general, es capaz de producir actividades y tareas variadas que fortalezcan la comprensión lectora.

En cuanto a la valoración de los distintos niveles de comprensión lectora, se muestra que el material favorece una progresión adecuada desde la comprensión literal hasta niveles más complejos, como la comprensión crítica y creativa. Este aspecto es relevante, ya que evidencia que las actividades propuestas en los materiales no se limitan a la recuperación de información explícita, sino que promueven procesos de análisis, inferencia y reflexión. No obstante, los expertos B, C y E, quienes previamente identificaron un pequeño desbalance en los ejercicios de pre-lectura, otorgaron valoraciones ligeramente inferiores de 4, especialmente en lo referente a la comprensión literal progresiva. Esto sugiere que ChatGPT-4 presenta ciertas limitaciones en la creación de actividades

iniciales en la comprensión lectora, como en el nivel literal, el cual implica no solo el reconocimiento de los detalles, sino también el orden de los acontecimientos (Kagami, 2020). Dado que en este nivel se requiere que los aprendices recuperen el contenido del texto (Kurniawan, 2015), se evidencia la necesidad de promover una comprensión inicial y general del mismo en los materiales. Es fundamental que todos los niveles de comprensión lectora sean abordados de forma adecuada en los materiales, de modo que orienten al estudiante en su aprendizaje de una lengua extranjera desde un enfoque superficial hacia niveles de comprensión más profundos.

En relación con los resultados cualitativos obtenidos en las entrevistas, se discutieron las concepciones de los expertos sobre lo que caracteriza a un material didáctico adecuado. En primer lugar, algunos señalaron la relevancia de una estructura continua y progresiva, siguiendo una secuencia que inicie con cuestiones más simples y que, gradualmente, se acerque al desarrollo de habilidades más complejas. En ese sentido, esta progresión se interpreta como un aspecto que necesariamente debe apoyarse en los distintos niveles de comprensión, ya que estos permiten una organización que paulatinamente guía el proceso de aprendizaje y el desarrollo de habilidades lectoras. Siguiendo esta línea de pensamiento, algunos expertos también mencionaron que, para el desarrollo de la comprensión lectora, se debe trabajar ante todo en el reconocimiento de la información explícita del texto, haciendo referencia al nivel literal de comprensión (Kagami, 2020), y la identificación del propósito del autor, aludiendo al nivel inferencial de comprensión (Safitri & Melati, 2023). Además, añadieron que se deben fortalecer las habilidades que les permitan a los estudiantes analizar y producir ideas de manera autónoma, lo que se podría entender como la integración del nivel crítico y aplicado. Por último, otros expertos coincidieron en que las temáticas abordadas en los materiales deben captar la atención de los estudiantes y ajustarse a su nivel, lo que destaca la importancia de seleccionar los recursos que sean más adecuados o la adaptación del contenido de estos, puesto que así, los docentes pueden realizar modificaciones que respondan de manera pertinente al contexto en el que se enseña (Nguyen, 2011; Febrina, 2017). Todas estas concepciones parecen ser referentes que los expertos consideraron al momento de analizar los materiales generados por ChatGPT-4.

Asimismo, los expertos retomaron sus valoraciones sobre el diseño de los materiales elaborados. Los participantes afirmaron que las actividades son adecuadas y permiten fortalecer la comprensión lectora, demostrando que ChatGPT-4 tiene un potencial en la creación de actividades de comprensión lectora. En este sentido, las actividades favorecen a los estudiantes, orientándolos en la construcción del significado del texto a partir de su conocimiento previo, experiencias y contexto, lo cual coincide con las concepciones de la comprensión lectora propuestas por Brown (2002) y Snow (2002). De esta forma, en términos generales, los expertos manifestaron percepciones positivas del material en relación con la estructuración a través de las fases de lectura,

la manejabilidad efectiva de la complejidad temática, sintáctica y léxica, y la relevancia de los contenidos para los estudiantes. En conjunto, señalaron que los materiales resultaron suficientemente pertinentes como para implementarlos en sus propios contextos educativos, tal como lo desarrollaron los expertos A y C. No obstante, también se identificaron algunos desajustes en el material, tales como ambigüedad en una pregunta, dificultad en el planteamiento de actividades de causa-efecto, los cuales constituyen un requerimiento relevante según el MEN (2016), y una limitada contextualización a la realidad educativa nacional. Estos resultados revelan que, aunque el material presenta fortalezas, ChatGPT-4 no ofrece resultados completamente precisos y debe entenderse como una herramienta de apoyo al docente y no como un reemplazo de su rol pedagógico. Tal como ilustra el estudio mixto de Xiao et al. (2023), la IA puede contribuir al diseño de materiales educativos, pero aún presenta limitaciones pedagógicas que requieren mediación y ajustes por parte del docente.

De manera que, considerando las propuestas pedagógicas de los expertos para el fortalecimiento de los materiales, se sugirieron la inclusión de enlaces a videos, audios, imágenes, entre otros recursos multimodales, lo que podría entenderse como apoyos didácticos que contribuirían a facilitar la comprensión e incrementar la motivación de los estudiantes, teniendo en cuenta que estos recursos pueden ofrecerles una aproximación a los temas que se abordan, ayudarles a fortalecer otras habilidades, como la comprensión auditiva, y hacer el material más dinámico e interesante para ellos. Asimismo, algunos expertos destacaron la necesidad de proponer orientaciones e instrucciones claras sobre la modalidad del trabajo. Por tanto, se trata de un aspecto que debe contemplarse en las indicaciones que se le brinden a la IA en la elaboración de materiales, debido a que su inclusión podría orientar las actividades de manera más precisa y apoyar una interacción más fluida con el material. Finalmente, los expertos manifestaron que se requieren preguntas o ejercicios que les permitan a los estudiantes expresar sus opiniones y aludir a su contexto cercano o a sus realidades. Esto podría implicar que los estudiantes se involucren de manera significativa con los textos, estableciendo relaciones entre los contenidos y sus experiencias.

En síntesis, los resultados, recopilados tanto de la rúbrica como de la entrevista, indican que los materiales elaborados cumplen de manera satisfactoria con los criterios de calidad para la comprensión lectora en nivel B1. La IA demuestra un gran potencial para la adaptación de textos y el diseño de actividades de comprensión lectora, al considerar las fases de lectura, las estrategias, los niveles de comprensión y la relevancia del contenido para los estudiantes. No obstante, también se pone de manifiesto que los productos generados no son plenamente efectivos por sí solos, ya que esta herramienta presenta ciertas limitaciones en cuanto a la articulación con el contexto educativo

real, las estrategias para una implementación efectiva y la necesidad de ajustes en algunas actividades. Estos resultados refuerzan la idea de que los materiales generados por ChatGPT-4 deben entenderse como recursos susceptibles de mejora y no como productos pedagógicos definitivos. En este sentido, la IA es una herramienta de apoyo para el maestro en el diseño de las actividades, mientras que la evaluación pedagógica y adaptación contextual continúan siendo la responsabilidad del maestro.

4.2. Alineación con los estándares de competencia lectora en inglés

4.2.1. Resultados

En una primera instancia, dado que el presente estudio analiza si los materiales elaborados con ChatGPT-4 cumplen con los criterios de los estándares de comprensión lectora en el nivel B1 en inglés, fue necesario establecer una sección en la rúbrica orientada a verificar la alineación de los ejercicios y de la adaptación del texto con el Marco Común Europeo de Referencia para las Lenguas (2020), el Esquema Curricular Sugerido (2016) y los Derechos Básicos de Aprendizaje de Inglés (2016). Para ello, en la misma rúbrica, se solicitó a los expertos evaluar el nivel de conexión del material con estos documentos a partir de cuatro preguntas, las cuales se resumen en los siguientes componentes: *alineación con los niveles de lectura del MCER*, *alineación con los niveles de lectura del ECS*, *alineación con los niveles de lectura de los DBA*, y *alineaciones con el texto adaptado*. La siguiente tabla (Tabla 4.5) presenta la evaluación de expertos sobre la adecuación de un material de comprensión lectora para 9º grado (nivel B1).

Tabla 4.5

Coherencia con los objetivos pedagógicos de lectura

Componentes	Experto A	Experto B	Experto C	Experto D	Experto E	Experto F	Promedio por componente
Alineación con los niveles de lectura del MCER	3	5	4	5	5	5	4,5
Alineación con los niveles de lectura del ECS	5	5	5	5	5	5	5
Alineación con los niveles de lectura de los DBA	4	5	4	5	5	5	4,6
Alineaciones con el texto adaptado	5	5	5	5	5	5	5
<i>Promedio por Experto</i>	4,2	5	4,5	5	5	5	4,7

En general, los resultados de la rúbrica evidencian percepciones positivas frente a la alineación del material con los documentos oficiales. Por componentes, la *alineación con los niveles de lectura del MCER* obtuvo un promedio excelente de 4,5; cuatro de los expertos consideraron que el material cumplía completamente con este estándar, uno calificó con 4 (Experto C) y otro con 3 (Experto A). Esto contrasta con la *alineación con los niveles de lectura del ECS* que alcanzó una valoración excelente, con un promedio de 5. La *alineación con los niveles de los DBA* registró un promedio de 4,6, ya que cuatro participantes los evaluaron con 5 y dos con 4 (Experto A y Experto C). Finalmente, las *alineaciones con texto adaptado*, orientadas a determinar si el texto facilitaba la identificación de estructuras discursivas de acuerdo con los estándares nacionales, obtuvieron un promedio excelente de 5. En general, la mayoría de los expertos participantes valoraron con una calificación de 5 el cumplimiento de los materiales con el nivel de los estándares de todos los documentos oficiales (MCER, ECS y DBA) tanto en las actividades como en el texto adaptado. Sin embargo, el experto A y el experto C consideraron que hay algunos aspectos de detalle a mejorar en cuanto a la alineación de los materiales con el MCER y los DBA.

Cabe destacar que, en la sección de comentarios de la rúbrica, el experto E señala la falta de integración de otras habilidades lingüísticas en el desarrollo de los materiales, los cuales considera necesarios para la comunicación moderna. En este sentido, el experto resalta que las actividades se centran principalmente en la lectura y la escritura, sin una articulación con la comprensión auditiva ni la expresión oral, tal como se evidencia en el siguiente extracto:

Las hojas de trabajo se adhieren a un modelo tradicional de habilidades aisladas que ignora las exigencias de la comunicación moderna y la alfabetización digital. Se centran principalmente en la lectura y la escritura, sin indicaciones ni estructuras específicas para integrar la comprensión auditiva (p. ej., ver un videoclip relacionado, escuchar el discurso original) o la expresión oral (p. ej., indicaciones para debates estructurados). (Comentario escrito por el experto E en la rúbrica de validación)

Referente a las entrevistas, en la sección denominada *sobre la rúbrica y los materiales*, se preguntó a los expertos si el material mostraba coherencia con los documentos oficiales de referencia y si reconocían algún desfase al respecto. A partir de sus respuestas, se identificaron ciertos patrones que dieron lugar a tres temas relevantes: *coherencia normativa de los materiales*, *tensiones entre normativa y contexto colombiano*, y *limitaciones pedagógicas del material en la práctica*.

A. Coherencia normativa de los materiales

En esta categoría, se agruparon los aspectos relacionados con la coherencia normativa y pedagógica de los materiales. En este sentido, se identifican tres subcategorías dentro de la temática

que involucra *la alineación con documentos oficiales, la pertinencia de las temáticas alineadas al nivel B1 y la resignificación de los Derechos Básicos del Aprendizaje (DBA)*. Estas subcategorías revelan una valoración general positiva del diseño del material y de su correspondencia con los objetivos oficiales.

En relación con la alineación con los documentos oficiales, todos los expertos afirmaron que los materiales cumplían con los estándares establecidos por el MCER, el SCE y los DBA. Esto se evidencia en la siguiente cita: “Me pareció que todo estaba como muy apropiado y de acuerdo a sus objetivos” (Entrevista del participante B, 2025).

Los expertos destacaron la correspondencia entre los materiales, los objetivos planteados y los niveles propuestos por dichos documentos. Se evidencia que se cumple un diseño pedagógico consciente frente a normativa vigente, cumpliendo con las exigencias curriculares desde los niveles estándares.

Respecto a la pertinencia de las temáticas para el nivel B1, la mitad de los expertos (B, C y D) indicaron que los materiales abordaban temas adecuados para estudiantes de este nivel de lengua. Uno de los participantes enunció:

Todo también me parece que es como muy, o sea, de lo que más me gustó de los textos es que son como muy cercanos a los estudiantes, como de temas muy, sí, como muy cercanos a sus experiencias, a sus problemas, a sus contextos, entonces pues sí me parece que son muy apropiados (Entrevista del participante B, 2025).

Las respuestas de los entrevistados señalaron la cercanía de los contenidos a situaciones cercanas, experiencias y contextos acordes con los estudiantes, lo cual concuerda con el nivel de comprensión lectora y facilita la integración estudiantil.

Finalmente, algunos expertos (A, C y F) hicieron referencia al uso de los DBA en el diseño de materiales. Uno de los participantes afirmó:

Es que esos DBA también son como muy... permiten que nosotros seamos creativos los profes, o sea, también nos da cierta libertad. Tampoco es que sea tan de camisa de fuerza, entonces creo que fueron también muy creativos de cumplir con los DBA... (Entrevista del participante C, 2025).

Los participantes aclararon que los DBA funcionan como marcos de referencia que guían el diseño curricular. Asimismo, afirmaron que estos son herramientas de apoyo para los profesores, sin que limite su autonomía del profesor ni el desarrollo de propuestas pedagógicas contextualizadas.

B. Tensiones entre normativa y contexto colombiano

Para esta categoría se tuvieron en cuenta las respuestas que hicieron referencia a la relación entre los documentos oficiales y a su relación con el contexto educativo colombiano. Esta sección reúne las críticas sobre *la influencia de referentes externos y las limitaciones de los estándares frente a la realidad educativa nacional*, demostrando incongruencias entre la normativa educativa nacional y la realidad dentro de los entornos educativos. Los participantes mencionaron aspectos relacionados con la influencia de referentes externos y con la adecuación de los estándares a la realidad nacional.

De entrada, un experto (E) señaló que los documentos del MEN están condicionados por referentes externos, lo cual limita la pertinencia contextual y su adecuación al contexto colombiano:

Todos estos documentos son contruidos por el Ministerio de Educación Nacional, pero el Ministerio de Educación Nacional es influenciado por el British Council. Entonces, es decir, el consejo [Consejo Británico] le dice al Ministerio de Educación Nacional cuál es la sociedad colombiana y cómo hay que enseñarle y qué tipo de inglés hay que enseñarle a esa sociedad colombiana (Entrevista del participante E, 2025).

De esta forma, el participante cuestiona la legitimidad contextual de los documentos nacionales, argumentando que responden más a modelos externos que a las realidades educativas del contexto colombiano. En consecuencia, sostiene que la normativa se construye a partir de reglas formales más que de las necesidades reales del sistema educativo.

Asimismo, algunos expertos (A, C, E y F) afirman que los niveles de lengua establecidos en los documentos oficiales no coinciden con las condiciones reales de los estudiantes a nivel nacional. Esto se evidencia en la siguiente cita: "Pues digamos que eso está bien según los DBA, ¿cierto? Sobre todo porque esto se crea sobre esa base. El contexto me dice otra cosa, yo no podría aplicar esto en 9°" (Entrevista del participante A, 2025).

Las respuestas de los participantes señalan que los niveles de lengua exigidos por los documentos oficiales no corresponden con las condiciones educativas reales. De esta forma, se evidencia una brecha entre los lineamientos establecidos en los documentos nacionales y la realidad educativa donde está dirigida.

C. Limitaciones pedagógicas del material en la práctica

Finalmente, con relación a esta categoría, se identificaron aspectos vinculados a las limitaciones pedagógicas de los materiales en la práctica educativa nacional. Este apartado incluye aspectos asociados a *la diversidad de niveles de competencias en los contextos educativos* y a la

necesidad de articulación en los materiales con los DBA, particularmente en lo referente a la integración de la producción oral como evidencia de la comprensión lectora.

En primer lugar, algunos expertos (A y E) manifestaron que, aunque los materiales correspondían a un nivel B1 en inglés, no contemplaban estrategias para estudiantes con niveles de competencia lingüística inferiores en la lengua meta. Esto se evidencia en la siguiente cita representativa: "No veo cómo el material tenga en cuenta o incluya a los estudiantes que no han llegado a ese nivel, porque los hay" (Entrevista del participante E, 2025).

Los participantes, al conocer el contexto y la realidad del nivel educativo de los estudiantes en el aprendizaje de inglés como lengua extranjera, reconocen que no todos presentan el mismo nivel. Concluyen que un bajo nivel de competencia lingüística no mejora con materiales que no se ajustan a su nivel.

Adicionalmente, el experto A, similar a lo que anteriormente mencionó el experto E en la sección de comentarios, afirmó que permanecía una falla en los materiales elaborados, la falta de integración de la producción oral en los materiales. El participante explica que incluirlo permite garantizar el pleno cumplimiento de los DBA, como se evidencia en la siguiente cita:

Realiza exposiciones cortas, bueno, no tanto, porque ellos no tienen que presentar, al menos de que ustedes hicieran una pregunta al final y que les dijeran OK, entonces presenta a tu grupo tus hallazgos, entonces esa es otra forma de saber que el estudiante comprendió la lectura, llevarlo a la producción oral (Entrevista del participante A, 2025).

Durante la entrevista, el experto recurrió a presentar el ESC y los DBA para leerlos y evidenciar si los materiales realmente cumplían con lo requerido o si presentaban algún desfase. En general, expresó una percepción positiva del material, pero destacó que la comprensión lectora también debería evidenciarse en el material, lo cual tiene coherencia con lo establecido dentro de los documentos nacionales.

En resumen, durante las entrevistas, los hallazgos más recurrentes indican que los materiales demostraban concordancia con los requerimientos de los documentos oficiales. Asimismo, los expertos manifestaron que las temáticas que se desarrollaban eran acordes con el nivel y eran apropiadas para los estudiantes. También, se destacó la comprensión de que los documentos nacionales deben asumirse como guías de apoyo para los docentes y no como reglas que rigen el sistema educativo. No obstante, algunos expertos también señalaron diversas dificultades relacionadas con los documentos oficiales. En primer lugar, se manifestaron críticas a dichos documentos, al no considerar la realidad del sistema educativo nacional, lo cual puede deberse a la

influencia de referentes externos alejados del contexto colombiano. Finalmente, se identificaron aspectos que difieren del diseño de los materiales presentados, como la falta de consideración de los distintos niveles de competencia lingüística de los estudiantes y la limitada integración de expresión oral que permita cumplir plenamente con los estándares nacionales.

4.2.2. Discusión

El nivel B1, según los estándares, es un nivel donde el usuario es capaz de comprender el inglés en contextos académicos y cotidianos, manejar diversas temáticas, identificar propósitos e interpretar la información a partir de la lectura (Council of Europe, 2020; MEN, 2016). Estas normativas son una guía fundamental para orientar la adaptabilidad de los materiales las cuales son importantes herramientas para el diseño de materiales en el aprendizaje de una lengua. Debido a ello, y tomando como referencia la relevancia del juicio de expertos (Escobar-Pérez & Cuervo-Martínez, 2008) y los antecedentes de estudios que incorporan este tipo de validaciones (Andriyani et al. 2018; Gutiérrez & Jurado, 2018; Nieves et al., 2013), se consultó a estos expertos con experiencia en los entornos educativos nacionales para determinar si los materiales elaborados cumplían con los estándares de la normativa educativa. Los hallazgos obtenidos demuestran que los materiales, elaborados con ChatGPT-4, son acordes con los estándares internacionales y nacionales para el nivel B1 en inglés. En general, tanto las valoraciones de los expertos como sus respuestas en la entrevista muestran una evaluación favorable referente a la alineación con los documentos oficiales (el Marco Común Europeo de Referencia para las lenguas, el Esquema Curricular Sugerido y los Derechos Básicos del Aprendizaje), los cuales fueron incorporados apropiadamente por la inteligencia artificial. Sin embargo, también se exponen ciertas limitaciones que invitan a un análisis sobre el empleo de los documentos oficiales y las falencias de los materiales elaborados.

De entrada, tanto las valoraciones alcanzadas en la rúbrica como en la entrevista, indican que los expertos consideran que los materiales demuestran una acertada concordancia con los estándares (MCER, ECS y DBA). La presencia de promedios superiores a 4,5 en todos los componentes y un promedio global de 4,7, señalan que, desde un punto de vista normativo, el material cumple con los lineamientos esperados para el nivel según los estándares. Esto sugiere que ChatGPT-4 podría ser una herramienta efectiva para ayudar a los docentes en la creación de materiales acordes con los documentos oficiales, disminuyendo la brecha entre las exigencias curriculares y el diseño de materiales. Sin embargo, se evidenciaron resultados un poco más bajos en cuanto a las alineaciones con el MCER y los DBA (expertos A y C). De esta forma, de acuerdo con los expertos, a pesar de que el material cumple con los criterios generales, hay elementos específicos que necesitan más precisión. Lo anterior puede deberse a dos posibles razones, fallas en el diseño del prompt o limitaciones de la IA en la comprensión de la instrucción.

Por un lado, se evidenció que, para el diseño del prompt se seleccionaron únicamente aspectos que se consideraron más relevantes, debido a que el MCER incluye muchos descriptores relacionados con la comprensión lectora de este nivel de lengua. En cuanto a los documentos nacionales, se consideraron aspectos exclusivamente de la lectura y no de otras habilidades. Esto pudo haber influido en las percepciones menos favorables, como expresó el experto A, quien señaló la falta de inclusión de la habilidad de habla para evaluar la comprensión de lectura. Por otro lado, la inteligencia artificial no es una herramienta perfecta y puede cometer errores en el diseño de materiales, como se ha evidenciado en los antecedentes (Sihite et al., 2023; Xiao et al., 2023). Esto sugiere que, a pesar de la utilidad de la IA para crear propuestas de materiales coherentes y organizados, el rol del docente aún es esencial para ajustar ciertos detalles concretos y asegurar una adaptación más exacta a los descriptores de desempeño.

En cuanto a los resultados cualitativos obtenidos de las entrevistas, las percepciones positivas permanecen, resaltando tanto la coherencia a los estándares como la pertinencia de los temas tratados. De esta forma, los docentes afirmaron que los temas en los materiales mantienen el nivel y están relacionados con las vivencias y experiencias de los alumnos, lo que contribuye a su motivación y participación con la lectura. Por lo tanto, los resultados evidencian que los textos adaptados y las actividades creadas por ChatGPT-4 cumplen con los estándares lingüísticos, favoreciendo la vinculación con el entorno del alumno.

No obstante, permanece un debate entre la normativa y el contexto real de la educación nacional. La relevancia contextual de los documentos del MEN (ESC y DBA) fue cuestionado, indicando que estos tienen una gran influencia de entidades extranjeras, como el British Council, lo cual se relaciona con discusiones sobre la adopción de modelos que no toman en cuenta las circunstancias educativas de Colombia. En este sentido, aunque dicha crítica apunta a una problemática asociadas con las políticas lingüísticas colombianas, no puede atribuirse directamente como una limitación del material diseñado en este estudio, ya que este se enmarca en las normativas vigentes. No obstante, como señala Augusto-Navarro (2015), los materiales educativos deben buscar un equilibrio entre las condiciones del contexto y las necesidades de los estudiantes, ya que su aplicación puede verse limitada ante los distintos niveles de lengua de los alumnos y las condiciones institucionales. De esta forma, los docentes indicaron que la alineación no asegura que el material sea pedagógicamente viable, debido a su falta de contextualización a la realidad educativa. En constancia con algunos estudios internacionales (Quintero, 2024; Koraishi, 2023), ChatGPT se concibe como una herramienta que no puede reemplazar la labor del maestro, debido a que la ausencia que tiene de conocimiento pedagógicos y de contextual limita la creación de materiales ajustados a las necesidades del aula.

Desde esta misma perspectiva, se evidenció que los materiales no incluyen estrategias de inclusión que atiendan la diversidad de niveles de competencia de los estudiantes, revelando una limitación en el diseño pedagógico. Además, los hallazgos revelaron una falta de integración de la producción oral y la comprensión oral como otra de las limitaciones del material. En la entrevista, se reconoció que los documentos nacionales plantean indirectamente que la interacción oral es otro factor que evidencia la comprensión lectora. En el diseño, a pesar de que el material cumple con los estándares de lectura, es necesario integrar actividades orales complementarias para fortalecer una visión más integral del aprendizaje del inglés. Lo anterior es evidente dentro del marco de comprensión lectora en una segunda lengua. Tal como afirman Hamada et al. (2022), las habilidades metalingüísticas y la comprensión oral influyen en la comprensión lectora. Por lo tanto, es fundamental integrar todas las habilidades en el diseño de materiales que permitan orientar a los estudiantes hacia el dominio de la segunda lengua.

En resumen, de acuerdo con la evaluación de los expertos consultados, los materiales de comprensión lectora generados con ChatGPT-4 tienen coherencia con los estándares oficiales y son apropiados en términos temáticos para el nivel B1. Sin embargo, para que una implementación efectiva en el entorno colombiano, es necesario el papel del docente crítico que tenga en cuenta las circunstancias del aula, la variedad de niveles lingüísticos y el desarrollo de habilidades orales en su diseño. De este modo, la inteligencia artificial, específicamente ChatGPT, es un instrumento de soporte que necesita ser contextualizado y adaptado para entender las verdaderas necesidades del sistema educativo, más no un reemplazo del maestro.

4.3. Adecuación cultural al contexto colombiano

4.3.1. Resultados

Considerando que uno de los objetivos de la presente investigación se centra en la adecuación cultural y contextual de los materiales de comprensión lectora generados por ChatGPT-4, se diseñó una sección en la rúbrica de validación que incluyó un criterio denominado *relevancia cultural y contextual*, por medio del cual se les solicitaba a los expertos evaluar en qué medida los materiales se ajustaban a las realidades o intereses de estudiantes colombianos de nivel B1 inglés. Es así que, se abordaron cuatro preguntas que abarcan los siguientes componentes: *Relevancia del tema para estudiantes de nivel B1, fomento de la reflexión personal, contenido contemporáneo y relevante y referencias culturalmente relevantes*. La necesidad de analizar estos componentes radica en conocer si estos materiales no solo permiten desarrollar habilidades lingüísticas, sino que, de igual forma, les posibilita a los estudiantes relacionar el contenido a sus realidades, experiencias o entornos. La tabla 4.6 muestra los resultados obtenidos de la evaluación de expertos referente a los componentes mencionados.

Tabla 4.6*Relevancia cultural y contextual*

Componentes	Experto A	Experto B	Experto C	Experto D	Experto E	Experto F	Promedio por componente
Relevancia del tema para estudiantes de nivel B1	5	5	5	5	4	5	4,8
Fomento de la reflexión personal	5	5	4	5	5	5	4,8
Contenido contemporáneo y relevante	5	5	5	5	5	5	5
Referencias culturalmente relevantes	5	5	5	5	3	5	4,6
<i>Promedio por Experto</i>	5	5	4,7	5	4,2	5	4,8

En términos generales, los resultados reflejan percepciones positivas sobre la adecuación cultural y contextual de los materiales. A partir del promedio obtenido por componente, se evidencia que sólo el relacionado al *contenido contemporáneo y relevante* consiguió un resultado excelente de 5. Le sigue en puntuación más alta el componente de *relevancia del tema para estudiantes de nivel B1* con un promedio de 4,8, ya que el experto E registró una calificación de 4; y el componente de *fomento de la reflexión personal*, igualmente con 4,8, siendo el Experto C el único que registró una valoración de 4. Al final, se encuentra el componente de *referencias culturalmente relevantes* con un promedio de 4,6, considerando que el experto E asignó una valoración de 3 (aceptable). Asimismo, en cuanto al promedio por experto, se muestra que solo los expertos C y E mostraron en su valoración un promedio inferior a 5, con resultados de 4,7 y 4,2 respectivamente. Este último promedio se refleja y se justifica igualmente en la sección de comentarios u observaciones de la rúbrica, donde el experto E señala que:

El material se basa en referencias culturales globales (*HeForShe*, películas de Hollywood sobre inteligencia artificial, cultura universal de las redes sociales). Esto es demasiado amplio para el contexto local: Los textos **no** integran aspectos culturales específicos relacionados con Cali, Colombia (por ejemplo, figuras locales, cultura de la salsa, cuestiones nacionales/regionales específicas). Esto limita la capacidad del material para conectar verdaderamente con la vida cotidiana de los estudiantes y aprovechar sus conocimientos culturales previos para una comprensión y un debate más profundos. (Comentario escrito por el experto E en la rúbrica de validación)

En la entrevista llevada a cabo con cada uno de los expertos se incluyó una pregunta en la sección de *sobre la rúbrica y los materiales* relacionada a la relevancia cultural y contextual, la cual estaba enfocada en identificar los elementos de los materiales que a los expertos les resultaba más pertinentes para estudiantes colombianos y cuáles no tanto. De acuerdo con las respuestas proporcionadas, se pueden identificar dos temáticas: *requerimiento de mayor contextualización a la realidad local y pertinencia de los temas abordados para estudiantes de noveno grado*.

A. Requerimiento de mayor contextualización a la realidad local

Esta categoría reúne las percepciones de los expertos que indican que, aunque los materiales son temáticamente relevantes, presentan limitaciones en cuanto a la relación con el contexto cultural específicamente colombiano. Como resultado, se logró identificar dos subcategorías dentro de este tema que son: *referenciar el contexto local en los ejercicios propuestos y enfoque globalizado de los contenidos*.

Con respecto a la primera subcategoría, referenciar el contexto local en los ejercicios propuestos, se encontró que dos expertos (A y B) señalaron la necesidad de incluir preguntas que les permitan a los estudiantes aludir a su realidad social: "Creo que sí se deben ajustar los elementos culturales a ser mucho más profundos, explicar por medio de esas preguntas o hacer esas preguntas un poco más cercanas a las realidades de los estudiantes." (Entrevista del participante A, 2025).

Los expertos destacaron la necesidad de acercar los materiales al contexto de los estudiantes, de modo que estos se sientan representados en los contenidos propuestos. Particularmente, el experto A, en otro momento de la entrevista, manifiesta que la inclusión de preguntas que inviten a los estudiantes a contar o a relacionar su realidad a lo leído, podría incrementar su motivación hacia las actividades.

En el caso de la segunda subcategoría, enfoque globalizado de los contenidos, el experto E expresó que los materiales promueven una noción de ciudadanía global y que, si bien esto no es desfavorable, mencionó que este enfoque puede reproducir formas de pensamiento colonizado. El experto, por lo tanto, declara que en el aprendizaje de una lengua extranjera se debería fortalecer la construcción de la cultura e identidad propia.

Lo que yo me propongo, una de las máximas de mi tesis doctoral, es un concepto que denomino como la etnomotivación, es decir, que usted no se aproxima al aprendizaje de una lengua extranjera porque usted quiere parecerse a un ciudadano de esa cultura.[...] La etnomotivación se trata más bien de que, desde mis bases de identidad como ciudadano indígena, afrocolombiano o mestizo, unas bases sólidas, fuertes, me acerco al aprendizaje de una lengua extranjera con el objetivo de seguirme

fortaleciendo como ciudadano colombiano y no buscando esa idealización de ciudadano global. (Entrevista del participante E, 2025).

Las respuestas de este experto señalan que los materiales parecen tener un enfoque globalizado, por lo que es necesario incluir contenido que permita visibilizar las culturas y contextos locales de los estudiantes

B. Pertinencia de los temas abordados para estudiantes de noveno grado

Esta categoría hace referencia a la relevancia del contenido abordado en los materiales para los estudiantes de noveno grado, que según el Esquema Curricular Sugerido (2016) corresponde al inicio y desarrollo del nivel B1 de inglés en Colombia. Por lo general se trata de estudiantes de edades que oscilan entre 13 a 15 años, de modo que, esta temática recoge las percepciones de los expertos sobre el interés que puedan generar los materiales a este colectivo. Es así que se distinguieron dos subcategorías que se dividen en *temas relacionados con la experiencia personal* y *temas sociales contemporáneos*.

En cuanto a los temas relacionados con la experiencia personal, algunos expertos (B, C, F) señalan que los materiales incluyen tópicos relacionados a la autopercepción y a las emociones, por lo que pueden ser recursos con los que los estudiantes puedan conectarse.

En nuestro colegio hay servicio de restaurante escolar pero no van a almorzar que porque se engordan; entonces aguantan hambre, entonces su autopercepción es muy baja. Y entonces ese texto dice que, a ver, precisamente habla de eso, de que en las redes sociales todo el mundo está pendiente de eso y uno no debe prestar tanta atención. (Entrevista del participante C, 2025).

Según los expertos, los tópicos que incluyen aspectos sobre el cuerpo, las emociones y las relaciones interpersonales pueden ser muy reconocibles y comprensibles para los estudiantes, lo que hace que se pueda captar su atención.

Asimismo, en los materiales se tratan temas sociales contemporáneos que abordan aspectos como roles de género o el uso de la inteligencia artificial, de manera que algunos expertos (B, C) consideran que estas cuestiones pueden ser relevantes para estudiantes de noveno grado.

[...]lo que más me gustó de los textos es que son muy cercanos a los estudiantes, como de temas muy, sí, como muy cercanos a sus experiencias, a sus problemas, a sus contextos, entonces pues sí me parece que son muy apropiados. (Entrevista del participante B, 2025).

Los expertos expresan que en los materiales se proponen contenidos que forman parte del entorno social de los estudiantes, lo que favorece su disposición. En general, la mayor parte de los expertos manifiestan que los materiales son pertinentes para grado noveno, debido a que tratan temas que pueden resultar interesantes para estudiantes de esa edad.

En síntesis, los resultados obtenidos de la entrevista en cuanto a la adecuación cultural al contexto colombiano muestran que los expertos consideran que los materiales presentan, en conjunto, una valoración positiva en términos de pertinencia de los contenidos para grado noveno, puesto que estos desarrollan temas de carácter personal y social que pueden ser de interés para los estudiantes. No obstante, algunos expertos señalan que los materiales necesitan integrar en las actividades preguntas que permitan reflexionar sobre la realidad propia y aludir a un contexto mucho más cercano. Además, un experto menciona que los materiales tienden a mostrar una idea de ciudadanía globalizada, y aunque no rechaza esta noción, es importante que estos recursos también ayuden a construir y a fortalecer la identidad cultural propia.

4.3.2. Discusión

En el aprendizaje de lenguas extranjeras no basta con que los estudiantes desarrollen la competencia lingüística, sino que también resulta necesario integrar elementos culturales que permitan la comprensión de realidades propias y ajenas para un aprendizaje más transformador. La lengua y la cultura mantienen una relación estrecha e inseparable (Coracini, 2010), por lo que es esencial fortalecer ese vínculo en la enseñanza. En el desarrollo de la comprensión lectora, en particular, ese vínculo adquiere aún más relevancia, puesto que la lectura implica activar conocimientos culturales propios para construir un significado (Mckay, 2006). Por consiguiente, en el marco de esta investigación, se consideró la adecuación cultural de los materiales generados por ChatGPT-4 como un aspecto importante a validar. Es así que, siguiendo los módulos propuestos por el Esquema Curricular Sugerido (2016), se seleccionaron los textos que aportarían el contenido temático y cultural de los materiales. Se consultó, entonces, a los expertos en el área de enseñanza del inglés como lengua extranjera, qué tan pertinentes son estos materiales en términos de integración de la cultura. Los resultados muestran que, en general, los materiales sí logran tratar temáticas culturales relevantes para estudiantes de nivel B1-noveno grado, aspecto que se refleja en las valoraciones de la rúbrica y la entrevista. Sin embargo, los expertos también identificaron algunas limitaciones en la falta de inclusión de una cultura aún más cercana al contexto de los estudiantes, lo que hace necesario realizar ajustes enfocados en profundizar estos elementos en los materiales generados por ChatGPT.

En lo que respecta la rúbrica de validación, las valoraciones proporcionadas por los expertos señalan que los materiales efectivamente proponen temas, referencias y contenidos culturalmente relevantes. Se presenta un promedio global excelente de 4,8, lo que puede suponer que ChatGPT-4 podría convertirse en un instrumento útil al momento de integrar aspectos culturales en los materiales didácticos de comprensión lectora, en la medida en que se le proporcione orientaciones claras y específicas. No obstante, algunas observaciones de los expertos aseguran que se podrían considerar ciertos factores culturales al momento de crear estos materiales. Por ejemplo, el experto E, con un promedio de 4,2, manifiesta en la sección de comentarios de la rúbrica, que los materiales limitan la capacidad de los estudiantes para conectarse con aspectos de su vida cotidiana, porque los temas propuestos no abordan el contexto local, sino que se basan en referencias culturales globales (muy generales). Esto puede atribuirse, por un lado, a la falta de especificaciones culturales explícitas en el prompt proporcionado a ChatGPT y, por otro, a que los módulos del Esquema Curricular Sugerido (2016) son de carácter general, sin un énfasis en contenidos culturales situados.

En primer lugar, al momento de diseñar el prompt no se tuvo en cuenta información sobre aspectos culturales específicos del contexto en el que se iban a aplicar los materiales, es decir, no se incluyeron indicaciones sobre la integración de cultura local (caleña). Al no mencionarlas, la IA tendió a producir materiales más generales que se limitaban a un contenido globalizado. Esto evidencia la importancia de formular prompts más precisos cuando se busca la integración intencional de contenidos culturales específicos en materiales didácticos. En segundo lugar, dado que el presente estudio consideró el Esquema Curricular Sugerido (2016) por el MEN como uno de los documentos que se abordarían en la elaboración de los materiales, es importante resaltar que los módulos propuestos son de carácter global, por lo que las temáticas culturales responderían a esta índole. Si bien dichas temáticas igualmente resultan pertinentes, al ser tan generales, reducen oportunidades para que los estudiantes puedan expresar su cultura local. Además, es necesario aclarar que estos documentos oficiales fueron producidos con la colaboración de *The British Council*, lo que implica un enfoque de alcance más global en la selección de los contenidos y una menor consideración de referentes culturales colombianos, factor que de hecho fue señalado por el experto E en la entrevista. En ese sentido, es clave que el docente utilice la IA mediante el diseño de instrucciones claras no solo sobre aspectos culturales generales, sino también insistiendo en que los enriquezca con elementos propios del contexto cercano y local.

En lo que concierne a la entrevista, las percepciones de los expertos sobre la adecuación cultural de los materiales permanecen relativamente positivas. Casi todos los expertos manifiestan que las temáticas culturales son relevantes para estudiantes de noveno grado, que equivale al nivel B1 en inglés según los estándares nacionales (aunque como se explicó en las secciones 4.1 y 4.2, la realidad sugiere un escenario diferente) y a una población estudiantil en etapa de adolescencia.

Puesto que en los materiales generados por ChatGPT-4 se tratan contenidos que podían relacionarse con la experiencia de los estudiantes y abordan también temas sociales de actualidad, estos favorecen a la motivación y a la reflexión. Sin embargo, otros expertos expresan que los materiales requieren algunos ajustes en relación a la integración cultural local.

Los expertos A y B coinciden en que los materiales podrían fortalecerse a través de la inclusión de ejercicios o preguntas que inviten a los estudiantes a reflexionar sobre su realidad y contexto propio. Cuando el contenido se relaciona con experiencias cercanas al entorno, los estudiantes tienden a identificarse con las temáticas, lo que propicia un mayor interés y participación en la construcción de sus respuestas. En el estudio de Ojeda (2018) se destaca que un diseño adaptado al contexto inmediato, en ese caso el colombiano, efectivamente, puede mejorar las habilidades de comprensión lectora. Asimismo, el experto E expresa que, aunque no sea negativo el hecho de mostrar una idea de cultura globalizada en los materiales, se debe incluir contenidos culturales que le brinden al estudiante la capacidad de expresar su propia cultura. El experto, siguiendo la línea de pensamiento de Liddicoat y Scarino (2013), asegura que el aprendizaje de una lengua extranjera implica partir de bases identitarias sólidas con el propósito de fortalecer y reafirmar la cultura propia. Por consiguiente, el uso de la IA para la producción de material didáctico exige, en la formulación de sus indicaciones o prompts, una concepción pedagógica que incorpore contenido del contexto cultural que se requiera.

En general, las percepciones de los expertos indican que los materiales didácticos de comprensión lectora generados por ChatGPT-4 son en cierta medida culturalmente relevantes para los estudiantes de noveno grado - nivel B1, sin embargo, aún se necesitan ajustes para que logren ser completamente pertinentes al contexto colombiano. Aunque los materiales presentan contenidos generales que pueden captar la atención de los estudiantes, porque abordan temas relacionados a experiencias comunes en la etapa adolescente y problemáticas sociales del momento, se hace necesario incorporar actividades que les permitan expresar y conectar con su cultura local y cercana. De esta manera, ChatGPT resulta ser una herramienta de apoyo que necesita indicaciones específicas para la integración de aspectos culturales que se aproximen a las realidades de la población estudiantil.

4.4. Fortalezas y limitaciones de ChatGPT en la elaboración de materiales

4.4.1 Resultados

Debido a que uno de los objetivos de este estudio se centra en identificar las fortalezas y limitaciones de la inteligencia artificial en la generación de materiales de comprensión lectora, en las entrevistas realizadas a los expertos, se incluyó dos secciones en la entrevista llamada: *sobre su experiencia* y *sobre la IA*. Estas secciones estuvieron orientadas a reconocer las ventajas y

desventajas del uso de esta herramienta en los contextos educativos, tanto para docentes como para estudiantes. En este sentido, se preguntó a los expertos cómo han usado la IA y cuáles son sus percepciones frente a su uso. A partir de las respuestas proporcionadas, se identifican dos categorías relevantes: *contribuciones percibidas del uso de la IA en el contexto educativo* y *limitaciones percibidas del uso de la IA en el contexto educativo*.

A. Contribuciones percibidas del uso de la IA en el contexto educativo

A partir de las entrevistas, esta categoría recopila los aspectos que evidencian las ventajas del uso de la inteligencia artificial, especialmente ChatGPT, en entornos educativos como herramienta de apoyo para los docentes, tanto en la adaptación de textos como en la creación de actividades. Esta sección comprende las subcategorías de *diseño de actividades de aula mediadas por inteligencia artificial*, *adaptación de textos al nivel y al contexto educativo*, *generación de preguntas a partir de textos*, *apoyo a la elaboración de instrumentos de evaluación* y *optimización del tiempo*.

En primer lugar, se presenta la subcategoría diseño de actividades de aula mediadas por inteligencia artificial. Los expertos B y C manifestaron la importancia de integrar estas nuevas tecnologías en el desarrollo de ejercicios dentro de las clases, como se evidencia en el siguiente extracto:

Yo lo que les digo sobre todo yo la he usado es como para actividades, como para hacer actividades en clase. No como para que la IA me diga qué actividad hacer, sino como actividades donde se incluye el uso de la IA. (Entrevista del participante B, 2025)

En definitiva, los expertos manifiestan las ventajas de incorporar estas tecnologías en los entornos de aula y resaltan la importancia de involucrar a los estudiantes en su uso, reconociendo el potencial de estas herramientas en el aprendizaje y la necesidad de integrarlas de manera pedagógica.

Adicionalmente, esta subcategoría reconoce la ventaja de la IA en la adaptación de textos al nivel y al contexto educativo. La mayoría de los expertos (A, B, D y F) indicaron que la inteligencia artificial permite ajustar adecuadamente textos a los contextos que se requieren, como se evidencia en la siguiente cita: "Cojo una lectura auténtica y le digo a la inteligencia artificial: el estudiante tiene esas características. Entonces lo que hace es adaptar una lectura a ese nivel del estudiante, y eso es lo que, digamos, por allí he realizado." (Entrevista del participante F, 2025).

De esta forma, los expertos señalaron que la inteligencia artificial es capaz de adaptar textos a diversos entornos, siempre que se le proporcione una instrucción clara y contextualizada que

incluya las características de los estudiantes. Esto permite adaptar los textos para cumplir con los objetivos y contenidos de aprendizaje previstos.

Otro aspecto relevante es el potencial de la inteligencia artificial en la generación de preguntas a partir de textos. La mayoría de los expertos (A, D, E y F) manifestaron haber utilizado la IA para crear preguntas basadas en textos, como se evidencia en lo siguiente extracto: "Había un texto quería sacar preguntas muy específicas y la IA me lo hizo, le subí el texto y le hice unas preguntas sobre esto basado en este enfoque y verdad, me lo hizo. Sí, es muy útil" (Entrevista del participante F, 2025).

En otras palabras, los expertos afirman que la inteligencia artificial es competente en la creación de preguntas, especialmente de comprensión lectora, y reconocen que puede funcionar como una herramienta de apoyo para el docente en el diseño de ejercicios orientados a la comprensión de textos específicos.

En relación con la evaluación docente, un aspecto relevante es el apoyo de la IA en la valoración de instrumentos de evaluación. Algunos expertos (A y B) indicaron que otro beneficio del uso de la IA radica en su capacidad de crear herramientas que permiten evaluar la comprensión de los estudiantes o nivelar su aprendizaje. Uno de los participantes señaló:

También yo creo que ayuda para hacer cosas como de rúbricas si uno le dice como los criterios y después, o sea como para que te cree como ciertas tablas y tú puedas como modificarlas después de que están creadas, sirve mucho, o sea porque ahorra muchísimo tiempo (Entrevista del participante B, 2025).

En este sentido, los expertos expresaron que esta herramienta tecnológica facilita la creación de formatos para valorar la comprensión y el desempeño de los estudiantes frente a las actividades de clase y los contenidos abordados, especialmente mediante el uso de rúbricas, lo que constituye la IA como una herramienta efectiva para el diseño de instrumentos de evaluación.

Finalmente, esta subcategoría aborda la optimización del tiempo que permiten estas herramientas tecnológicas. La mitad de los expertos (A, B y F) coinciden en que el uso de la inteligencia artificial contribuye a reducir el tiempo dedicado al diseño de ejercicios, como se evidencia en lo siguiente: "Sí, yo creo que es útil en muchos aspectos. Primero porque ayuda a agilizar el trabajo" (Entrevista del participante F, 2025).

De esta forma, los expertos señalan que el uso de estas tecnologías facilita la reducción del tiempo en el diseño de actividades y materiales educativos, ya que la IA es capaz de presentar propuestas claras, organizadas y entendibles.

B. Limitaciones percibidas del uso de la IA en el contexto educativo

Esta categoría comprende todas las limitaciones percibidas por los expertos frente al uso de la inteligencia artificial. Durante las entrevistas, se evidenció que la IA puede aportar múltiples beneficios dentro en los entornos educativos. Sin embargo, como se mencionó en las discusiones anteriores, al no contar con la capacidad pedagógica, contextual y humana propia de un docente, puede presentar errores en determinados aspectos del proceso educativo. En este sentido, se identificaron las siguientes subcategorías: *falta de conocimiento contextual por parte de la IA, necesidad de pensamiento crítico por parte del docente, posible despersonalización de los textos adaptados y materiales generados por la IA, riesgo de facilismo para los estudiantes y consideraciones éticas en el uso de la IA*. En conjunto, estos aspectos reflejan las principales limitaciones del uso de la IA en el ámbito educativo.

En primer lugar, en relación con la falta de conocimiento contextual por parte de la IA, algunos expertos (A y B) indicaron que, dado que la inteligencia artificial se basa en una visión general y abstracta de los contextos a los que se dirige sus producciones, presenta limitaciones en la comprensión del contexto real, así como en la lectura que los docentes hacen tanto del espacio como de las características de sus estudiantes. El experto B lo manifestó así:

Los profes igual tienen un conocimiento no sólo de la lengua así como abstracto y del nivel y los estándares y el currículum y todo esto, sino que tiene un conocimiento de los estudiantes, del contexto como más de cómo es el salón de clase, de qué cosas hay en el salón de clase para trabajar, de cómo son los estudiantes, porque puede que uno le pueda decir a la IA, son estudiantes de 18 años, pero pues solo el profesor es el que sabe cómo son realmente esos estudiantes, cuál es su nivel de verdad, cuáles son, en qué cosas tienen dificultad realmente como que la IA un poco trabaja con unas cosas muy abstractas, igual el nivel y como la edad de los estudiantes pues siguen siendo como abstractas, entonces la desventaja es esa, como no tener ese conocimiento más concreto del contexto. (Entrevista del participante B, 2025).

De esta forma, los expertos señalaron que la inteligencia artificial comprende el contexto desde una perspectiva externa y que no puede integrar plenamente el conocimiento profundo que el maestro posee sobre sus estudiantes. Por ello, sus producciones carecen de una adecuación y contextualización real, ya que la IA es ajena a estos elementos y no logra comprenderlos en su totalidad.

Otra subcategoría relevante comprende la necesidad de pensamiento crítico por parte del docente. La mayoría de los expertos (A, D, E y F) explicaron que, aunque la inteligencia artificial posee un gran potencial para el diseño de materiales, estos no son perfectos por sí mismos y

requieren de la revisión, adecuación y orientación pedagógica del docente. Esta postura se refleja en lo expresado por uno de los participantes:

Primero el profesor que no lee lo que saca la IA, ¿Cierto que si? Ustedes podrían generar buenas lecturas, pero si ustedes comparten el prompt, el profesor va a generar muchas lecturas que realmente no van a tener ningún sentido si no las lee minuciosamente (Entrevista del participante A, 2025).

En este sentido, los expertos afirman que el uso de la inteligencia artificial puede constituir una herramienta útil para los docentes. Sin embargo, su empleo no debe limitar la visión crítica del docente ni promover la idea de que un simple *prompt* y el material producido por la IA pueden cubrir un plan de clase o un currículo sin una revisión previa. Por ello, no se elimina la necesidad de una revisión y adecuación pedagógica del contenido generado.

Igualmente, se evidencia que el uso de la IA puede provocar una despersonalización de los textos adaptados y materiales generados. Los expertos A y B afirman que la inteligencia artificial tiende a limitar el componente humano de los textos, volviéndolos excesivamente neutros e impersonales, como lo expresa una de las participantes: "Como que no, como que no sabría cómo yo evaluar o como hacer la comprensión de un texto que yo sé que no es como el texto original" (Entrevista del participante B, 2025).

Estos expertos señalan que la inteligencia artificial puede realizar adaptaciones que restan sentido y enfoque de humanidad a los textos hechos por personas, lo que los convierte en impersonales. Esto resalta la necesidad del componente humano en el proceso de adaptación de los materiales generados con IA a contextos específicos. En suma, el experto A manifiesta que los materiales generados por IA no buscan establecer una conexión con los estudiantes, carecen de una intención motivacional que impulse su proceso de aprendizaje.

Respecto a los estudiantes, se identifica el riesgo de facilismo en el desarrollo de actividades de las actividades comprensión. Los expertos A y F señalaron que la inteligencia artificial puede convertirse en un obstáculo para el aprendizaje, como se evidencia en la siguiente cita: "Esa es como la única desventaja que yo leo, como el facilismo que les da a los estudiantes para cumplir por cumplir, no para interiorizar el conocimiento y apersonarse, lo que busca uno en la enseñanza" (Entrevista del participante F, 2025).

Los expertos se muestran preocupados por la posibilidad de plagio en el desarrollo de las actividades, señalando que algunos estudiantes podrían recurrir al uso de la inteligencia artificial para copiar y pegar las respuestas generadas, en lugar de elaborar sus propias producciones o resolverlo por ellos mismos.

Finalmente, se presenta la subcategoría referente a las consideraciones éticas en el uso de la IA. Los expertos A y E señalaron que el uso de la inteligencia artificial en la creación y adaptación de materiales y lecturas para el docente crea un debate entre la ética y los derechos de autor, particularmente en relación con la propiedad de los materiales producidos. Uno de los participantes enunció: "¿Entonces a quién le pertenece todo esto? Ah no, pues yo cogí la lectura, Miguel me la pasó, yo lo implementé, pues eso se genera por IA, ya es mía, ¿cierto? ¿Entonces cómo pueden tomar esa información sin autorización?" (Entrevista del participante A, 2025).

En definitiva, los participantes manifestaron una preocupación frente a los derechos de propiedad intelectual de los materiales generados o adaptados mediante la inteligencia artificial, cuestionando si estos pertenecen exclusivamente al docente o si, al ser producidos por IA, adquieren un carácter colectivo. De este modo, se evidencia cómo la ética adquiere un papel central en relación con la autoría, el uso y la publicación de textos y materiales educativos.

En resumen, la inteligencia artificial es una herramienta con gran potencial en los entornos educativos. Para el docente, puede ser una herramienta pedagógica de apoyo que facilite la integración de nuevas tecnologías con sus estudiantes y contribuya al diseño de materiales, mediante adaptaciones de textos, la creación de herramientas de evaluación y la creación de preguntas de comprensión lectora, optimizando así la exigente labor docente. Sin embargo, la IA no es una herramienta perfecta y presenta limitaciones, tanto en los procesos de aprendizaje de los estudiantes como en el riesgo de que los docentes la utilicen como una solución inmediata sin un análisis previo. Además, la IA plantea debates éticos, carece de una comprensión contextual precisa y tiende a despersonalizar los textos, por lo que su uso debe ser cuidadoso y acompañado de una revisión y adecuación pedagógica por parte del docente.

4.4.2 Discusión

La inteligencia artificial es un sistema con la capacidad de interpretar instrucciones y crear contenido a partir de ellas, aprendiendo del usuario y anticipándose a sus necesidades (Dan et al., 2022; Levy, 2024). Esta tecnología ha demostrado un gran potencial a nivel global, extendiendo su uso a diversos ámbitos, incluidos los entornos educativos. Uno de los ejemplares más representativos es el chatbot ChatGPT, un modelo de IA diseñado para interactuar con los usuarios y responder sus preguntas, con una notable capacidad para comprender el lenguaje humano (Lim et al., 2023). Su utilidad se ha ampliado tanto para docentes como para estudiantes, al contribuir a la mejora de las experiencias de aprendizaje y a la optimización del trabajo docente. Durante la recolección de datos, se evidenció que todos los expertos participantes reconocen el potencial de ChatGPT y han hecho uso de esta herramienta en contextos académicos. Por esta razón, desde las

entrevistas, resultó pertinente analizar tanto sus aportes como las percepciones de los docentes frente a su implementación, lo que permitió identificar no sólo las contribuciones, sino también las limitaciones que puede implicar el uso de la inteligencia artificial en entornos académicos.

En primera instancia, es evidente que los expertos reconocen la importancia de integrar estas tecnologías en el aula, destacando tanto su potencial como la necesidad de un uso pedagógico adecuado. Dado que estas herramientas forman parte del entorno tecnológico actual, resulta fundamental involucrar a los estudiantes en su uso de forma crítica, promoviendo no solo la interacción con la tecnología, sino también la evaluación de sus resultados y el desarrollo de habilidades que les permitan explorar conscientemente su potencial. En este sentido, pueden incorporarse estrategias o actividades que incluyan el uso de estas herramientas, la evaluación de sus respuestas, el análisis de datos y el debate en el aula, favoreciendo el aprendizaje dinámico, tal como lo evidencian Contreras y Pinzón (2024) en su investigación sobre la implementación de materiales elaborados por IA. De esta forma, ChatGPT-4 puede ser una herramienta que se debe integrar en los procesos de aprendizaje, orientada a explorar sus ventajas pedagógicas, motivar a los estudiantes y fortalecer los entornos educativos. Estudios previos han demostrado que estos sistemas tienen resultados positivos en el aprendizaje de una segunda lengua (Haggag, 2021), al contribuir al dominio de las habilidades lingüísticas, fomentar una gran motivación, proporcionar una retroalimentación e impulsar la autonomía en los procesos de aprendizaje (Holmes et al., 2019).

En cuanto al diseño de materiales de comprensión lectora, los expertos manifestaron que ChatGPT es capaz de adaptar textos a diversos entornos. De esta forma, se establece como una herramienta útil para la adaptación de textos orientados al cumplimiento de los contenidos y objetivos del plan de clase, lo que facilita la incorporación de textos auténticos en contextos educativos de distintos niveles de lengua. Esta capacidad permite ajustar los textos a las necesidades de los estudiantes, orientándolos en su aprendizaje de una lengua extranjera. No obstante, para que este proceso se desarrolle efectivamente, es necesario proporcionar a la IA instrucciones clara y contextualizada que describan las características de los estudiantes. En este sentido, tal como señala Morales-Chan (2023), la claridad del prompt condiciona la efectividad de los resultados. Asimismo, los expertos destacan a ChatGPT como un sistema de apoyo para el docente en el diseño de materiales de comprensión lectora, debido a que es su potencial en la creación de ejercicios y preguntas orientadas a la comprensión lectora en una segunda lengua. En este sentido, la IA puede contribuir al diseño de actividades siempre que se le proporcionen prompts que integren aspectos importantes como las fases de lectura y los niveles de comprensión. De esta forma, ChatGPT posibilita la elaboración de materiales didácticos pertinentes para el aprendizaje de una segunda lengua, específicamente en el desarrollo de la comprensión lectora.

La entrevista también reveló que los expertos reconocen que la inteligencia artificial se ha consolidado como una herramienta útil para facilitar la creación de formatos de evaluación que permiten valorar tanto la comprensión como el desempeño de los estudiantes en relación con las actividades y los objetivos de clase. En este sentido, la IA es capaz de reconocer aspectos específicos que requieren ser evaluados en el aprendizaje de una lengua extranjera, posibilitando el diseño de sistemas orientados a determinar y calificar el desempeño del estudiante, como se ha evidenciado en su uso para la elaboración de rúbricas de evaluación. Sin embargo, la efectividad de estos procesos depende en gran medida de una instrucción clara y de los criterios pedagógicos proporcionados a la IA. A pesar de esta condición, la IA muestra resultados positivos en la creación de estos formatos, ya sea mediante tablas previamente estructuradas o a través de documentos editables que permiten realizar ajustes adicionales, e incluso en formatos listos para su implementación, como archivos PDF. Así, la inteligencia artificial se presenta como una herramienta pertinente para los procesos de evaluación en los contextos educativos.

En definitiva, la IA contribuye a la optimización del tiempo de la labor docente. Los expertos señalaron que el uso de esta tecnología permite reducir el tiempo destinado al diseño de actividades y materiales educativos. En este sentido, los maestros pueden reconocer que el uso de herramientas como ChatGPT favorece una mayor eficiencia en su labor, lo que les permite concentrarse en otros aspectos relevantes dentro del entorno educativo. Al disminuir la carga asociada al diseño de materiales, el docente puede dedicar más tiempo a fortalecer la relación con los estudiantes, mostrar una mayor disposición pedagógica y reconocer las necesidades y falencias del contexto educativo para abordarlas efectivamente. Asimismo, el uso de la IA puede contribuir a mitigar limitaciones relacionadas con la escasez de materiales de comprensión lectora, la falta de textos adaptados al contexto y ausencia de herramientas efectivas de evaluación. Esto se debe a que la IA es capaz de producir propuestas claras, organizadas y entendibles. De esta forma, el maestro puede aprovechar las utilidades de esta herramienta y compartir su eficiencia con los estudiantes, integrando nuevas tecnologías que favorezcan el desarrollo de entornos de aprendizaje motivadores y participativos.

Pese a las ventajas que puede presentar el uso de IA en contextos educativos, los expertos también señalaron en la entrevista diversas limitaciones derivadas de esta herramienta. En principio, algunos expertos manifestaron que la IA opera a partir de información general y no tiene la capacidad de reproducir exactamente el entorno de aprendizaje en el que se desenvuelve el docente. La IA no puede integrar el conocimiento pedagógico de un profesor, ya que no está al tanto del desempeño de los estudiantes, ni del tipo de interacciones que ocurren en el aula. Si bien el docente, al momento de crear material, puede proporcionar a la IA información importante sobre las características del contexto, esto no equivale a una comprensión profunda de aquello que se

desarrolla en el aula. Esta limitación implica que los materiales generados o adaptados por IA pueden carecer pertinencia pedagógica y contextual si no son ajustados o revisados por los docentes. Por lo tanto, coincidiendo con lo señalado por Koraishi (2023), es necesario reconocer que la IA no puede contemplarse como un reemplazo del trabajo del docente, sino como un apoyo didáctico, puesto que su efectividad depende plenamente de las intervenciones y de la capacidad del docente para contextualizar los recursos generados con esta herramienta.

En concordancia con lo anterior, algunos expertos igualmente manifestaron que la IA no debe sustituir la visión crítica del docente, ni utilizarse como un instrumento con facultad pedagógica capaz de diseñar planes de clase o currículos, mediante prompts, sin la revisión del profesor. El uso irrestricto de la IA puede encaminar a sugerencias que, aunque posiblemente sean bien estructuradas y organizadas, no se conecten al contexto educativo correspondiente o a las necesidades de los estudiantes. Es así que, aquello que supone las planificaciones de clases debe estar guiado por el juicio del docente, quien es el responsable de cuestionar el contenido que se produzca a partir de la IA.

En cuanto a la creación o adaptación de textos, fue señalado por algunos expertos que existe el riesgo de perder el componente humano al utilizar la IA con este propósito. Se mencionó que esta herramienta tiende a estandarizar y despersonalizar los textos, por lo que, si bien estos pueden lograr cumplir con los niveles o estructuras lingüísticas requeridas, pueden aún así resultar menos cercanos o motivadores para los estudiantes. Esto refuerza el valor de la dimensión humana en los textos, aspecto que podría verse mitigado cuando son adaptados o generados de manera automática por una IA. Por consiguiente, la revisión y ajustes por parte de los docentes se vuelve extremadamente pertinente, ya que su intervención activa en la adaptación o creación puede reducir este carácter impersonal de los textos y atribuirles una intencionalidad pedagógica precisa.

En lo que corresponde al trabajo de los estudiantes, otra de las limitaciones observadas por los expertos se asocia al facilismo en el uso de la IA. Se señaló que los estudiantes pueden recurrir al empleo inmediato de esta, al momento de participar o resolver actividades, lo que hace que no se involucren en el análisis o la reflexión que dichas tareas buscan fomentar. Este tipo de prácticas puede ser un riesgo para el proceso de aprendizaje de los alumnos, debido a que el uso indiscriminado de la IA implica reducir el esfuerzo requerido para el fortalecimiento de la comprensión lectora y el pensamiento crítico, lo que, a largo plazo, perjudica también el desempeño de los estudiantes. Por lo tanto, si el docente no establece parámetros para su uso, la IA puede convertirse en un recurso que cause dependencia, en lugar de apoyar un aprendizaje autónomo y significativo. En ese sentido, el docente debe convertirse en un mediador que equilibre el uso de este instrumento sin adoptar un enfoque sancionador, considerando que la IA está cada vez más presente en el aula y no resulta viable prescindir de ella, por el contrario, se hace necesario guiar y

regular su utilización. Es así que el docente debe pensar en estrategias que indiquen e instruyan a los estudiantes el correcto empleo de la IA en la resolución de actividades, para que esta también se vuelva un apoyo en su proceso de aprendizaje.

Finalmente, otro de los inconvenientes relacionados con la utilización de la IA para la elaboración de recursos pedagógicos está asociado a las consideraciones éticas que se pueden presentar. Algunos expertos indicaron que el uso de la IA puede conllevar problemas vinculados a los derechos de propiedad intelectual, especialmente en lo que respecta a la autoría de los contenidos o a la adaptación de textos ya existentes. En esta situación se vuelve complejo determinar hasta qué punto los materiales se valoran como originales o derivados de otra información, lo que implica que el docente deba asumir una responsabilidad ética respecto a su origen y uso. Cabe señalar que las discusiones y debates sobre el uso de la IA en la creación de recursos aún se encuentran en desarrollo, por lo que no hay lineamientos completamente consolidados que instruyan su utilización. Ante estas circunstancias, el procedimiento principal viable en el contexto educativo se trataría de adoptar transparencia, reconociendo explícitamente su uso y citando cuando sea necesario.

En términos generales, los resultados de las entrevistas señalan que las percepciones de los expertos sobre el uso de la IA para la elaboración de material didáctico son, por un lado, positivas, debido a que esta herramienta puede funcionar como un apoyo para la integración de actividades en el aula mediadas a través de su uso, permite adaptar los contenidos de acuerdo al contexto, ofrece asistencia en la construcción de instrumentos de evaluación, como rúbricas, y favorece a una gestión más eficiente del tiempo del docente, en tanto el profesor establezca prompts precisos y coherentes con sus objetivos. Estas fortalezas demuestran el potencial de la IA en el área pedagógica, facilitando especialmente a los docentes el proceso de elaboración de material. Sin embargo, por otro lado, los expertos también identificaron algunas limitaciones y desventajas que conlleva su utilización. Estos inconvenientes se relacionan con la falta de conocimiento contextual profundo de la IA en comparación con el del docente, la posible despersonalización de los textos adaptados y el riesgo de causar dependencia en los estudiantes al usarla sin regulación. Además, los expertos también tuvieron en cuenta los problemas éticos que esto puede implicar. Por ello, el uso de la IA debe mediado por el docente, quien se encarga de evaluar críticamente el propósito y el contenido del material generado.

4.5. Implementación de los talleres

Como se evidenció anteriormente, dos participantes, experto A y experto C, pusieron en práctica los materiales en sus cursos y brindaron una retroalimentación al respecto. El experto A,

partiendo de una iniciativa propia, no sólo implementó el material, sino que proporcionó resultados específicos derivados del material correspondiente a la reseña de la película *Her*. Durante la entrevista, el experto señaló que el material fue desarrollado de manera colectiva, organizando a los estudiantes en cinco grupos. Cada grupo contó con un líder que tenía un nivel de competencia en la lengua superior, con el fin de apoyar a los demás estudiantes dentro del curso de Inglés con Fines Generales y Académicos II, en la Universidad del Valle. El experto describe el nivel de competencia de los estudiantes del curso como heterogéneo, ubicándolo entre un nivel básico A1 y un nivel intermedio B1, como se evidencia en el siguiente extracto:

Yo tengo un grupo que es inglés II y el nivel de ellos es muy, muy diverso, hay estudiantes que tienen un A1, hay estudiantes que tienen un nivel A2 y hay otros que tienen un B1, o sea, entre A1 y B1. Entonces lo que yo hice antes de hacerlo como esa adecuación, porque yo ya me estaba pensando el taller, yo ya tengo un bosquejo general de mi curso, yo ya sé quién tiene un nivel avanzado y quién tiene un nivel que necesita incrementar (Entrevista del participante A, 2025).

Adicionalmente, el participante incluyó una sección destinada a que los estudiantes escribieran comentarios sobre la actividad. Los resultados obtenidos fueron los siguientes. El grupo 1, conformado por cuatro estudiantes, obtuvo una calificación de 4.8 y escribieron que la actividad resultó interesante por abordar la inteligencia artificial como un tema actual y futuro, además de contribuir al aprendizaje de nuevo vocabulario y a la reflexión sobre el uso de la IA, como se evidencia en la siguiente cita:

This activity was interesting because AI is a current and future topic. It helps us understand how things change and how humans must adapt. This activity allows us to learn new vocabulary, to practice the class items and improve English. This activity gives us clarity on how each of us is using the IA (Taller grupo 1, 2025). [Esta actividad fue interesante porque la inteligencia artificial es un tema actual y futuro. Nos ayuda a entender cómo cambian las cosas y cómo los humanos deben adaptarse. Esta actividad nos permite aprender nuevo vocabulario, practicar los contenidos de la clase y mejorar el inglés. Esta actividad nos da la claridad sobre cómo cada uno de nosotros está usando la inteligencia artificial] (Traducción propia).

El grupo 2, compuesto por seis estudiantes, obtuvo una calificación de 4.5 y destacó la forma en que la actividad permitió reflexionar sobre la percepción de la inteligencia artificial y la interacción entre los seres humanos y la tecnología, escribiendo lo siguiente: "it's interesting because of the way AI is perceived, how humans interact with it - it prompts reflection on our connection with AI" (Tomado del taller grupo 2, 2025) [Es interesante por la manera en que la

inteligencia artificial es percibida, cómo los humanos interactúan con ella - impulsa la reflexión sobre nuestra conexión con ella] (Traducción propia)

El grupo 3, integrado por cuatro estudiantes, alcanzó una calificación de 4.6 y señaló que la actividad promovió la reflexión sobre el impacto de la tecnología en las emociones humanas y las relaciones interpersonales, como se evidencia en la siguiente cita: "this activity made me think about how technology affects humans emotions and relationships. It was interesting and helped me reflect on real communication" (Taller grupo 3, 2025) [Esta actividad me hizo pensar sobre cómo la tecnología afecta las emociones y las relaciones humanas. Fue interesante y me ayudó a reflexionar sobre la comunicación real] (Traducción propia)

Por su parte, el grupo 4, conformado por dos estudiantes, obtuvo la calificación más alta de 4.9 y comentaron que la actividad permitió analizar cómo la inteligencia artificial puede influir en las emociones y las relaciones, destacando el componente humano de la tecnología, como se evidencia en la siguiente cita: "This activity made me think about how artificial intelligence can affect our emotions and relationships. Her shows that technology can also have a human side" (Taller grupo 4, 2025) [Esta actividad me hizo pensar sobre cómo la inteligencia artificial puede afectar nuestras emociones y relaciones. *Her* muestra que la tecnología puede contener un lado humano] (Traducción propia)

Finalmente, el grupo 5, compuesto por tres estudiantes, obtuvo una calificación 4.2 y manifestó, como se evidencia en la siguiente cita: "it is so interesting talk about AI because people is'nt talking in account AI, it is each day more into our lives" (Taller grupo 5, 2025) [Es tan interesante hablar sobre la inteligencia artificial porque las personas no están teniendo en cuenta la inteligencia artificial, cada día está más presente en nuestras vidas] (Traducción propia)

Es importante resaltar que, durante la implementación de la actividad, tanto el experto como los estudiantes identificaron ambigüedad en una de las preguntas del taller, ya que manifestaron que había varias opciones de respuesta posibles. La Figura 4.1 presenta la pregunta en cuestión.

Figura 4.1

Pregunta de Matching con ambigüedad

B. Matching – Match the Sentence Halves

Connect each sentence from Column A with its best ending from Column B.

Column A	Column B
1. Samantha listens carefully to Theodore	a. so they sound more personal and real
2. Theodore includes details in his letters	b. because she wants to understand him
3. Samantha creates a song	c. so she becomes more human-like
4. Theodore observes people in the street	d. because he is very empathetic
5. Samantha learns from daily conversations	e. so she can grow and express herself

Write your answers (1–5):

1 → ___ 2 → ___ 3 → ___ 4 → ___ 5 → ___

Tras una revisión detallada, se evidenció que las opciones de respuesta de la pregunta podrían corresponder a diferentes respuestas según la interpretación del texto, dado que varias combinaciones están igualmente respaldadas por la información proporcionada. Esta situación podría presentarse en los ítems 1, 2, 3 y 5 de la *columna A*. En consecuencia, no existe una correspondencia clara y específica de tipo uno a uno entre las acciones y sus consecuencias. Aunque la actividad permite posibles respuestas correctas, está diseñada como si solo hubiera una, lo cual podría afectar el material desde una perspectiva pedagógica y evaluativa.

Aunque esta implementación no estaba prevista desde el inicio, los resultados enriquecen significativamente el enfoque de la presente investigación. En primera instancia, es evidente que el material cumple con su propósito dentro del ejercicio de comprensión lectora. De acuerdo con las calificaciones obtenidas en los talleres en grupos, puede afirmarse que, en un contexto de estudiantes de nivel básico-intermedio, el material resulta efectivo para fortalecer la comprensión lectora en la lengua meta. Si bien no se realizó un diagnóstico previo de las necesidades de los estudiantes, como en el estudio de Motato (2020), los resultados permiten concebir a ChatGPT-4 como una herramienta capaz de adaptar textos auténticos a contextos educativos, ajustándose a los contenidos y los objetivos pedagógicos, así como de diseñar actividades orientadas a la comprensión lectora. Es fundamental señalar que la efectividad de estos resultados depende en gran medida de la claridad del prompt y de la inclusión de criterios específicos sobre los logros esperados. En este estudio, como se ha señalado anteriormente, se fundamentaron tanto en los documentos oficiales (Council of Europe, 2001; MEN, 2016) como en el marco conceptual relacionado con la comprensión lectora (Brown, 2002; Snow, 2002), las fases de lectura y los niveles de comprensión (Westwood, 2001).

No obstante, en los contextos educativos persisten diferencias en los niveles de comprensión lectora de los estudiantes, por lo que resulta necesario implementar estrategias que integren a estudiantes con distintos niveles de competencia, favoreciendo un aprendizaje continuo y colectivo. Por ello, el profesor implementó la actividad en grupos y diseñó una sección en la que los estudiantes expresaron la importancia de abordar el tema de la inteligencia en el aula, propiciando así espacios de reflexión sobre la relación del tema con sus realidades. En los comentarios, los estudiantes también demostraron que el material y el tema fueron de su interés y los motivaron a participar activamente en las actividades propuestas, favoreciendo su compromiso en su aprendizaje de una segunda lengua.

Asimismo, el material sirvió como recurso para espacios de socialización y debate que favorecieron la práctica de la expresión oral en la lengua meta, aspecto que, como se mencionó anteriormente, resulta pertinente introducir. Además, el material apoyó a los estudiantes en el aprendizaje de nuevo vocabulario, en la práctica los contenidos trabajados y en la mejora de su competencia lingüística en inglés. En este sentido, el uso de estas herramientas tecnológicas puede contribuir a potenciar la motivación y el compromiso de los estudiantes en el desarrollo de la comprensión lectora en una segunda lengua, tal como evidencian estudios previos (Contreras & Pinzón, 2024; Haggag, 2021). Sin embargo, también se identificaron algunas limitaciones. Aunque la IA es capaz de crear ejercicios y adaptar textos, puede cometer ciertos errores, como se evidenció en la ambigüedad presente en una de las preguntas. Esta ambigüedad provocó que las respuestas encajaran en varias partes del ejercicio y no en una opción específica, lo cual revela la importancia de una revisión crítica del docente antes de una implementación en el entorno educativo. En definitiva, ChatGPT tiene un alto potencial para el aprendizaje de una segunda lengua. Su capacidad para adaptar textos auténticos al nivel de los estudiantes y diseñar ejercicios de comprensión lectora evidencian su utilidad tanto para el docente, como apoyo en el diseño de materiales, como para el estudiante, al constituirse en un recurso útil que favorece el desarrollo de la habilidad de comprensión lectora en una lengua extranjera.

Este capítulo presentó y analizó los resultados del estudio, respondiendo a las preguntas de investigación. Los hallazgos mostraron que los expertos perciben a ChatGPT-4 como una herramienta útil en la creación de material didáctico de comprensión lectora, en términos de pertinencia pedagógica, alineamiento con los estándares de competencia en inglés y adecuación cultural al contexto colombiano. Asimismo, se encontró que ChatGPT-4 puede ser un instrumento de apoyo para la generación y producción de actividades y adaptación de textos para un nivel lingüístico determinado, que en este caso, se trató del nivel B1 en inglés. No obstante, esta IA también presenta ciertas limitaciones relacionadas a la correspondencia a la realidad educativa,

considerando que, para algunos expertos, noveno grado no representa un nivel B1. Además, se observó que los materiales, si bien eran pertinentes para los estudiantes, podrían aludir más a su contexto inmediato. Por último, se discutió que el uso de ChatGPT-4 debe ser medido y estar acompañado de la visión crítica del docente. El siguiente capítulo presentará de forma más detallada las conclusiones de la investigación.

En el siguiente capítulo se muestra la conclusión del presente estudio, en el que se abordan tanto las principales contribuciones como las limitaciones identificadas.

5. CONCLUSIONES

En la enseñanza de lenguas extranjeras es importante incluir nuevas estrategias que ayuden a desarrollar las habilidades comunicativas que se necesitan en un mundo globalizado. En este contexto, el interés reciente por la IA, como ChatGPT-4, ha permitido su incorporación como una herramienta con potencial en entornos educativos, al apoyar el aprendizaje y la enseñanza de lenguas. En este sentido, este estudio examinó la efectividad y pertinencia de materiales de comprensión lectora generados por ChatGPT-4 para el nivel de competencia lingüística B1 en inglés. Para tal fin, se diseñó un *prompt* mediante el cual se adaptaron cuatro textos auténticos y se elaboraron sus respectivos materiales con actividades. Posteriormente, seis expertos evaluaron, a través de una rúbrica y una entrevista, si dichos materiales favorecen el desarrollo de habilidades de comprensión lectora en términos de léxico, gramática, estrategias de aprendizaje y niveles de comprensión, así como su alineación con los documentos oficiales, la pertinencia cultural y el reconocimiento de las fortalezas y limitaciones en la elaboración de materiales de comprensión lectora mediante inteligencia artificial. Los resultados recolectados permitieron determinar, desde la perspectiva docente, si los materiales son adecuados y pertinentes para estudiantes de nivel B1 en inglés y de noveno grado de educación secundaria.

Los resultados obtenidos evidencian que los materiales generados cumplen con los criterios de calidad establecidos para el nivel B1. En relación a los textos adaptados, las valoraciones de los expertos indican una adecuada coherencia con los estándares oficiales, así como una variedad y calidad de los textos y temáticas pertinentes para los estudiantes de noveno grado, demostrando que los contenidos son adecuados al contexto estudiantil. Asimismo, los expertos reconocieron el potencial de la IA para la adaptación de textos auténticos a contextos educativos, adaptándolos a los contenidos y objetivos pedagógicos, así como para el diseño de actividades de comprensión lectora. En cuanto a los ejercicios, se señala una correspondencia clara con los estándares oficiales y una integración pertinente de las fases de lectura, las estrategias lectoras y los distintos niveles de comprensión, lo que favorece el trabajo progresivo de la habilidad lectora en estudiantes de noveno grado. Estos materiales permiten trabajar desde la comprensión lectora literal hasta niveles más profundos, siempre que se proporcionen instrucciones claras y criterios pedagógicos definidos.

En este sentido, es fundamental comprender que la claridad del *prompt* influye directamente en la calidad y pertinencia de los materiales. Por ello, los expertos reconocen las contribuciones que podría aportar el uso de la IA en los contextos educativos, como el apoyo en la creación de instrumentos de evaluación, la creación de actividades de comprensión a partir de textos auténticos y la adaptación de contenidos a diferentes niveles y contextos. Asimismo, se destaca la importancia de integrar estas herramientas tecnológicas en el aula para motivar a los estudiantes en el dominio de nuevas tecnologías y permitirles conocer las diversas utilidades que estas ofrecen para el aprendizaje, favoreciendo el aprendizaje autónomo. En definitiva, el uso de esta herramienta amplía las posibilidades de aprendizaje y permite optimizar la labor docente, brindando apoyo al maestro y

facilitando que pueda centrarse en otros aspectos pedagógicos que, por limitaciones de tiempo, resultan difíciles de atender.

No obstante, los resultados también revelan ciertas limitaciones en los materiales, especialmente en lo relacionado con su contextualización al entorno educativo colombiano. Esto se debe a que, en la realidad educativa actual, los estudiantes de noveno grado no tienen un nivel B1 en inglés. Además, se identifican dificultades en la adecuación a la diversidad de niveles de competencia presentes en el aula, ya que en un mismo salón de clase no todos los estudiantes mantienen el mismo nivel de competencia lingüística. Asimismo, se presenta la falta de incorporación de recursos multimodales que dinamicen el material, la escasa inclusión de preguntas orientadas al contexto cercano de los estudiantes y la ausencia de frases motivadoras e instrucciones claras de abordaje del material. Algunos expertos señalaron posibles ambigüedades en determinadas preguntas y dificultades en actividades específicas, como aquellas orientadas a relaciones causa-efecto, lo cual puede afectar la precisión de las respuestas y la comprensión de los estudiantes. De igual forma, los expertos manifestaron preocupaciones relacionadas con la despersonalización de los textos y materiales generados por IA, el riesgo a fomentar el facilismo por parte de los estudiantes y los debates éticos que implica el uso de estas herramientas por parte de los maestros. Si bien se reconoce que estas herramientas tecnológicas no poseen una perspectiva tan crítica ni pedagógica como la del docente, los expertos coinciden en que su uso debe seguir siendo explorado, siempre que se garantice la revisión, adaptación y contextualización del material por parte del docente según su realidad educativa.

Los resultados de la investigación muestran que, en entornos educativos, las producciones didácticas elaboradas a partir de herramientas emergentes, como por ejemplo la IA, requieren de procesos de validación que permitan analizar su eficacia antes de su implementación. La validación, especialmente hecha por expertos, permite evaluar de forma crítica la adecuación de estos recursos, en términos de cumplimiento con los objetivos de aprendizaje y ajuste al nivel lingüístico esperado. Es así que, la innovación educativa mediante la IA no se limita a la simple generación de recursos didácticos, sino que implica necesariamente un procedimiento de evaluación para asegurar su pertinencia. Por lo tanto, se pone en evidencia que la generación de materiales didácticos con IA demanda que los docentes desarrollen criterios para revisar y ajustar estos recursos. La utilización de estas herramientas no supone una simple aplicación, por el contrario, refuerza la idea de que el docente debe convertirse en un agente reflexivo capaz de evaluar qué tan pertinentes pueden resultar dichos materiales en el contexto educativo. Además, es importante reconocer que la calidad de estos materiales no depende únicamente de la herramienta, sino también de la precisión con la que el docente logra reflejar sus conocimientos pedagógicos en las indicaciones (o prompts) que guían el proceso de generación. De esta forma, la IA se integra como apoyo al proceso de enseñanza y aprendizaje, sin sustituir la intervención pedagógica del profesor.

Asimismo, es pertinente señalar que la elaboración de materiales didácticos puede convertirse en un proceso colaborativo entre docentes de educación secundaria y universitaria, ya que así se podría integrar el conocimiento teórico de la pedagogía con la realidad del aula escolar. Como resultado, los materiales producidos responderían no solo a criterios y parámetros académicos, sino que también serían más contextualizados y relevantes en el proceso de enseñanza-aprendizaje de una lengua extranjera.

A partir de los hallazgos, se reconoce que la incorporación de la IA al trabajo docente puede transformar de manera positiva el aprendizaje del inglés como lengua extranjera. No obstante, como se mencionó anteriormente, este potencial necesita una revisión constante y crítica de expertos en esa área. Por ello, es fundamental que en la educación se lleven a cabo procesos de validación.

Ahora bien, entre las limitaciones de este estudio se encuentra, en particular, la dificultad para reunir participantes, lo que propició a un panel de expertos muy reducido. Esta situación se debió principalmente a que los docentes contactados contaban con una agenda laboral bastante demandante y el proceso de evaluación de los materiales, junto con la realización de las entrevistas, requería una inversión de tiempo considerable. Por ello, no todos los posibles docentes participantes pudieron comprometerse con las distintas fases de la investigación. Asimismo, aunque los expertos en este estudio contaban con una experiencia notable en el área de la enseñanza del inglés como lengua extranjera, la mayoría intervenía en el contexto universitario, mientras que únicamente un participante se desempeñaba de manera exclusiva en el nivel de educación secundaria. En ese sentido, habría sido más oportuno incluir un mayor número de expertos que ejercieran solamente en la educación secundaria y, de esta manera, se habría desarrollado un análisis más profundo y ajustado al proceso pedagógico que se lleva a cabo en ese contexto, favoreciendo también a una valoración más precisa de la pertinencia de los materiales didácticos diseñados en la presente investigación para ese nivel en particular. Es así que, esta limitación plantea la necesidad de futuros estudios que amplíen la muestra e incorporen la participación de expertos con mayor contacto con la educación secundaria a fin de obtener resultados más ajustados a la realidad de ese nivel.

Pese a las limitaciones de este estudio, la investigación contribuye significativamente en los procesos de validación de material didáctico generado con IA de acuerdo con la percepción de expertos en el campo. Este estudio es igualmente relevante para el contexto local y nacional y puede utilizarse como referencia para futuras investigaciones que profundicen el potencial de la IA en la educación.

Siguiendo con otras líneas de investigación, deberían considerarse estudios enfocados en la precisión de los prompts, profundizando en su optimización para el diseño de materiales. De igual forma, investigaciones posteriores podrían orientarse hacia la revisión y diseño de materiales de comprensión lectora teniendo en cuenta los ajustes sugeridos por los resultados de este estudio.

Se considera también relevante ampliar el uso de ChatGPT en la generación de materiales didácticos enfocados en competencias diferentes a la comprensión lectora, como la comprensión oral, la producción escrita y la producción oral. Así se podría analizar con mayor detalle la efectividad de esta IA, así como sus capacidades y limitaciones, en la elaboración de recursos que busquen fortalecer distintas habilidades. Del mismo modo, resulta conveniente explorar la generación y adaptación de materiales mediante ChatGPT para otros niveles lingüísticos más básicos o más avanzados, lo que también permitiría evaluar y comprobar la capacidad de esta herramienta para producir contenidos que se ajusten a la complejidad de cada nivel y a los lineamientos pedagógicos que los definen. Adicionalmente, se podría analizar si es posible replicar este proceso de elaboración de material utilizando otras herramientas de IA disponibles, considerando que en los últimos años han surgido múltiples opciones que podría ofrecer diferentes ventajas en esta área.

Finalmente, además de la elaboración de material didáctico, futuras investigaciones podrían examinar otras aplicaciones de la IA en la educación. Sería interesante investigar sobre su potencial para fomentar el aprendizaje autónomo del inglés en los estudiantes, igualmente, analizar cómo esta herramienta podría apoyar a los docentes en la creación y optimización de instrumentos de evaluación, o incluso, explorar su uso en el diseño y planificación de actividades y ejercicios que puedan ser llevados a cabo durante las clases.

REFERENCIAS

- Acquah, E. K. (2023). Effectiveness of Interviews as a Tool for Data Collection. *European Academic Research*, 10(12), 4235-4241. <https://tinyurl.com/3b2b56ma>
- Adam, S. L., Stan, R. S., Moanga, A. S., Oroian, E., Mihai, M., y Ciubancan, M. (2010). The role of authentic materials in foreign language teaching. *Bulletin UASVM Horticulture*, 67(2), 431-435.
- Augusto-Navarro, E. H. (2015). The design of teaching materials as a tool in EFL teacher education: Experiences of a Brazilian teacher education program. *Ilha do Desterro*, 68(01), 121-137. <https://doi.org/10.5007/2175-8026.2015v68n1p121>
- Akbari, R. (2012). Validity in language testing. In C. Coombe, P. Davidson, B. O'Sullivan, & S. Stoyonoff (Eds.), *The Cambridge guide to second language assessment* (pp. 30-36). Cambridge University Press.
- Allen, D., y Tanner, K. (2006). Rubrics: Tools for making learning goals and evaluation criteria explicit for both teachers and learners. *CBE—Life Sciences Education*, 5(3), 197-203. <https://doi.org/10.1187/cbe.06-06-0168>
- Al-Homoud, F., y Schmitt, N. (2009). Extensive reading in a challenging environment: A comparison of extensive and intensive reading approaches in Saudi Arabia. *Language Teaching Research*, 13(4), 383-401. <https://doi.org/10.1177/1362168809341508>
- American Educational Research Association, American Psychological Association, y National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Psychological Association. American Educational Research Association. <https://www.testingstandards.net/open-access-files.html>
- Andrade, H. (1996). Understanding rubrics. *Educational Leadership*, 54(4), 14-17.
- Andriyani, S., Saidah, N. y Dian, H. (2018). Construct validity on teaching materials of language assessment based on problem-based learning. *Proceedings of the 1st International Conference on Education and Social Science Research (ICESRE 2018)*, 22-26. Atlantis Press. <https://www.atlantis-press.com/proceedings/icesre-18/55912173>
- Arias, A., y Sireci, S. (2021). Validez y Validación para Pruebas Educativas y Psicológicas: Teoría y Recomendaciones. *Revista iberoamericana de psicología*, 14(1), 11-22. <https://doi.org/10.33881/2027-1786.rip.14102>
- Aristizábal, A y Velásquez, A. (2016). *Diseño, implementación y evaluación de una secuencia didáctica para la enseñanza del inglés como lengua extranjera a través de historias cortas para estudiantes del grado décimo de un colegio bilingüe de Cali*. [Trabajo de grado, Universidad del Valle] Repositorio de la Universidad del Valle. <https://hdl.handle.net/10893/9475>
- Arribas, M. (2004). Diseño y validación de cuestionarios. *Matronas profesión*, 5(17), 23-29. <https://s3-eu-south-2.ionoscloud.com/assetsedmayero/articles/3MOIddYTjIrd3vGliPNR6V4DjnfA0J0BH6nuHMrz.pdf>

- Bans-Akutey, A. y Tiimub, B.M. (2021). Triangulation in Research. *Academia Letters*, 2(3392), 1-7.
<https://doi.org/10.20935/AL3392>
- Bharuthram, S. (2012). Making a case for the teaching of reading across the curriculum in higher education. *South African Journal of Education*, 32(2), 205-214.
<https://doi.org/10.15700/saje.v32n2a557>
- Braun, V., y Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101. <https://share.google/ANeIFFiYGP0lnW3x>
- Braun, V., y Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, y K. J. Sher (Eds.), *APA handbook of research methods in psychology*, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological (pp. 57–71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- Brown, D. (2002). *Language Assessment Principles Brown*. Longman.
- Caley, M. J., O'Leary, R. A., Fisher, R., Low-Choy, S., Johnson, S., y Mengersen, K. (2014). What is an expert? A systems perspective on expertise. *Ecology and evolution*, 4(3), 231-242.
<https://onlinelibrary.wiley.com/doi/10.1002/ece3.926>
- Cifuentes, A. y Niño, M. (2010). Criterios pedagógicos para el diseño de material didáctico en medio digital para apoyar la enseñanza y el aprendizaje de lenguas extranjeras. Recuperado de:
<http://hdl.handle.net/10554/12207>
- Cline, F., Johnstone, C., y King, T. (2006). *Focus group reactions to three definitions of reading (as originally developed in support of NARAP goal 1)*. National Accessible Reading Assessment Projects. <https://files.eric.ed.gov/fulltext/ED506575.pdf>
- Coracini, M. J. (2010). Langue-culture et identité en didactique des langues (FLE). *Synergies Brésil*, (Especial 2), 157-166. https://www.gerflint.fr/Base/Bresil_special2/coracini.pdf
- Council of Europe (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- 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>
- Contreras, S. y Pinzón, D. (2024). *Diseño y evaluación de un material didáctico innovador utilizando recursos generados por inteligencias artificiales para la enseñanza de la Historia en estudiantes de sexto grado*. [Trabajo de grado, Universidad de Cartagena]. Repositorio de la Universidad de Cartagena. <https://hdl.handle.net/11227/17800>
- Creswell, J. W., y Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications, Inc.
- Dang, H., Mecke, L., Lehmann, F., Goller, S., y Buschek, D. (2022). How to prompt? Opportunities and challenges of zero-and few-shot learning for human-AI interaction in creative applications of generative models. *arXiv:2209.01390v1 [cs.HC]*.
<https://doi.org/10.48550/arXiv.2209.01390>

- Daskalos, K., y Jellum Ling, J. (2006). *Authentic texts or adapted texts - That is the question! The use of authentic and adapted texts in the study of English in two Swedish upper secondary schools and a study of student and teacher attitudes towards these texts*. [Trabajo de grado, Malmö Högskola School of Teacher Education (LUT)]
<https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1489671&dswid=27>
- Day, R. R., y Park, J. S. (2005). Developing Reading Comprehension Questions. *Reading in a foreign language*, 17(1), 60-73. <https://eric.ed.gov/?id=EJ689120>
- Díaz-Guio, D. A., Henao, J., Pantoja, A., Arango, M. A., Díaz-Gómez, A. S., y Gómez, A. C. (2024). Artificial intelligence, applications and challenges in simulation-based education. *Colombian Journal of Anesthesiology*, 52(1), 1-7. <https://doi.org/10.5554/22562087.e1085>
- Doughty, J., Wan, Z., Bompelli, A., Qayum, J., Wang, T., Zhang, J., Zheng, Y., Doyle, A., Sridhar, P., Agarwal, A., Bogart, C., Keylor, E., Kultur, C., Savelka, J., y Sakr, M. (2024). A comparative study of AI-generated (GPT-4) and human-crafted MCQs in programming education. In *Proceedings of the 26th Australasian Computing Education Conference* (pp. 114-123).
<https://doi.org/10.1145/3636243.3636256>
- Dörnyei, Z. y Taguchi, T. (2010). *Questionnaires in second language research. Construction, administration and processing* (2a. ed.). New York and London: Routledge.
- Ebrahimpourtaher, A., y Hamidi, E. (2015). Authenticity and adaptation in EFL materials. *International Archive of Applied Sciences and Technology*, 6(2), 34-39.
- Ellis, P. (2016). The language of research (part 11)--research methodologies: interview types. *Wounds UK*, 12(4), 104-106. https://wounds-uk.com/wp-content/uploads/2023/02/content_11878.pdf
- Erazo, A. (2018). *Diseño, implementación y evaluación de una secuencia didáctica para la enseñanza de inglés con propósitos específicos a estudiantes de medicina y cirugía de la Universidad del Valle*. [Trabajo de grado, Universidad del Valle] Repositorio de la Universidad del Valle. <https://hdl.handle.net/10893/13993>
- Escobar-Pérez, J., y Cuervo-Martínez, Á. (2008). Validez de contenido y juicio de expertos: una aproximación a su utilización. *Avances en medición*, 6(1), 27-36. <http://surl.li/ypetfv>
- Febrina, W. (2017). Authentic vs. non-authentic materials in teaching English as a foreign language (EFL) in Indonesia: Which one matters more? In *The Asian Conference on Education 2017: Official conference proceedings* (pp. 731-742). The International Academic Forum (IAFOR).
<https://papers.iafor.org/submission38649/>
- Flores, S. (2019). Nivel comprensión lectora en inglés como segunda lengua en estudiantes universitarios. *Revista Innova Educación*, 1(3), 288-300.
<https://doi.org/10.35622/j.rie.2019.03.003>
- Giménez, G. (2011). Comunicación, cultura e identidad: Reflexiones epistemológicas. *Cultura y representaciones sociales*, 6(11), 109-132.
- Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, 8(4), 597-606. <https://doi.org/10.46743/2160-3715/2003.1870>

- Grabe, W. (2006). Areas of research that influence L2 reading instruction. In E. Usó-Juan, y A. Martínez-Flor. (Eds.), *Current Trends in the Development and Teaching of the four Language Skills* (pp. 279-301). Gruyter Mouton.
- Grabe, W., y Yamashita, J. (2022). *Reading in a second language: Moving from theory to practice* (2nd ed.). Cambridge University Press. <http://surl.li/bsrhfq>
- Gutiérrez, M. y Jurado, F. (2018). Análisis de ítems que miden niveles de comprensión lectora: Validez y tipología de errores. *Evaluación de aula, evaluación estandarizada y emergencia de sistemas de evaluación de aprendizajes*, 217-189. <http://surl.li/fxoqka>
- Güzel, S. (2022). The importance of reading activities in reading comprehension. *Disiplinlerarası Yenilik Araştırmaları Dergisi*, 2(1), 29-34. <https://dergipark.org.tr/en/pub/dyad/article/1033324>
- Haenlein, M., y Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California management review*, 61(4), 5-14. <https://doi.org/10.1177/0008125619864925>
- Haggag, H. (2021). Implementing artificial intelligence-powered tools in enhancing TOEFL-ITP test takers' reading and writing skills. *Journal of Faculty of Education, Benha*, 32(128), 129-156. <https://doi.org/10.21608/jfeb.2021.240117>
- Hamada, A., Shimizu, H., Hoshino, Y., Takaki, S., y Ushiro, Y. (2024). Robust evidence for the simple view of second language reading: Secondary meta-analysis of Jeon and Yamashita (2022). *Studies in Second Language Acquisition*, 46(5), 1355-1372. <https://doi.org/10.1017/S0272263124000226>
- Haleem, A., Javaid, M., y Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *BenchCouncil transactions on benchmarks, standards and evaluations*, 2(4), 100089.
- Harida, E. S. (2016). Understanding paragraph and three phases in reading comprehension. *English Education: English Journal for Teaching and Learning*, 4(01), 115-126. <https://jurnal.uinsyahada.ac.id/index.php/EEJ/article/view/1195>
- Hassan, E. M., Abdel-Haq, E. M., Mohamed, F. S., y Abed, M. M. A. (2018). Enhancing EFL Creative Reading Using Strategy Instruction among Official Preparatory School Pupils. *Benha Faculty of Education Journal*, 29(115), 1-28. https://mail.jfe.bu.edu.eg/files/53190_1574097248.pdf
- Holmes, W., Bialik, M., y Fadel, C. (2019) *Artificial Intelligence In Education: Promises and Implications for Teaching and Learning*. The Center for Curriculum Redesign.
- Hossain, K. (2024). Reviewing the role of culture in English language learning: Challenges and opportunities for educators. *Social Sciences y Humanities Open*, 9, 100781. <https://doi.org/10.1016/j.ssaho.2023.100781>
- Ito, H. (2015). Is a rubric worth the time and effort? Conditions for its success. *International Journal of Learning, Teaching and Educational Research*, 10(2), 103–116. <https://www.ijlter.org/index.php/ijlter/article/view/260>

- Junguito, M., y Chenoll, A. (2024). El uso de las herramientas de inteligencia artificial generativa en el aula de lenguas extranjeras desde la perspectiva del docente. *RE@ D-Revista de Educação a Distância e Elearning*, 7(1). <https://doi.org/10.34627/redvol7iss1e202410>
- Kamagi, S. (2020). A Study on students' ability in literal and inferential comprehension of English texts. In *Journal of International Conference Proceedings* 3(2), 140-144. Association of International Business and Professional Management. <https://pdfs.semanticscholar.org/16aa/c1dfc21a573b11cd6b715cfe76c41f08445b.pdf>
- Kaplan, A., y Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15–25. <https://doi.org/10.1016/j.bushor.2018.08.004>
- Khamraeva, G. I. (2016). Benefits of implementation of pre-, while and post' reading activities in language learning. *International scientific journal*, 4(1), 45-46. <https://tinyurl.com/bddd64ea>
- Kildiusheva, O. (2025). *Aplicación de la Inteligencia Artificial Generativa (IAG) para la creación de materiales didácticos en la enseñanza de lenguas extranjeras*. [Tesis de Maestría, Departamento de Lingüística Aplicada, Universitat Politècnica de València] <https://riunet.upv.es/server/api/core/bitstreams/10e72282-f9c6-4051-83c3-de5c871b542f/content>
- Kirana, D. P. (2014). Authentic materials in EFL classrooms. *Cendekia: Jurnal Kependidikan Dan Kemasyarakatan*, 12(2), 353-364.
- Kiszcak, A. (2019). Reading as a core component of developing academic literacy skills in L2 settings. *Lublin studies in modern languages and literature*, 43(3), 89-98. <http://dx.doi.org/10.17951/lsml.2019.43.3.89-98>
- Koraishi, O. (2023). Teaching English in the age of AI: Embracing ChatGPT to optimize EFL materials and assessment. *Language Education and Technology*, 3(1), 55-72. <https://langedutech.com/letjournal/index.php/let/article/view/48>
- Kramsch, C., y Zhu, H. (2016). Language, Culture and Language Teaching. In G. Hall (Ed.), *Routledge Handbook of English Language Teaching* (pp. 38-50). London: Routledge.
- Kurniawan, Y. I. (2015). The Effect of Think Aloud Strategy in Teaching Reading of Descriptive Text On The Eighth Grade At Smp Islam Manbaul Ulum Gresik [Undergraduate thesis, Universitas Muhammadiyah Gresik]. <http://eprints.umg.ac.id/2809/>
- Kusumarasdyati, K. (2019, July 12–13). *Adapting authentic texts for reading materials*. [Paper presentation]. The 16th JETA International Conference, Yogyakarta, Indonesia.
- Levy, A. J. (2024). Redefining Interaction: Trends and challenges in AI-driven experience design. En *SIGraDi 2024 proceedings* (pp. 2129–2140). CumInCAD. https://papers.cumincad.org/data/works/att/sigradi2024_122.pdf
- Liddicoat, A. J., y Scarino, A. (2013). *Intercultural language teaching and learning* (1st ed.). Wiley-Blackwell.
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., y Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from

- management educators. *The international journal of management education*, 21(2), 100790.
<https://doi.org/10.1016/j.ijme.2023.100790>
- Loeb, S., Dynarski, S., McFarland, D., Morris, P., Reardon, S., y Reber, S. (2017). Descriptive Analysis in Education: A Guide for Researchers. NCEE 2017-4023. National Center for Education Evaluation and Regional Assistance. <https://files.eric.ed.gov/fulltext/ED573325.pdf>
- Long, M. (2014). *Second Language Acquisition and Task-Based Language Teaching*. Wiley-Blackwell.
- Luís, A. R. (2024). Examining adaptation practices in English Language Textbooks: A closer look at authenticity. *Journal of English Studies*, (22), 199-218. <https://doi.org/10.18172/jes.5932>
- Mckay, P. (2006). *Assessing Young language Learners*. Cambridge University Press.
- Mlakar, H. (2020). Linguistic factors influencing L2 reading comprehension. In G. Mehegan y M. Wilson (Eds.), *Literacy without borders. Engaging literacy learners in diverse settings* (pp. 166-185). Dublin: Eden Print Ltd.
- Ministerio de Educación Nacional de Colombia (MEN). (2006). *Estándares Básicos de Competencias en Lenguas Extranjeras: Inglés*.
https://www.mineduacion.gov.co/1780/articles-115174_archivo_pdf.pdf
- Ministerio de Educación Nacional de Colombia (MEN). (2016a). *Derechos básicos de aprendizaje: Inglés, Grados 6° a 11°*. Santillana Plus.
<https://santillanaplus.com.co/pdf/DBA-ingles-espanol.pdf>
- Ministerio de Educación Nacional de Colombia (MEN). (2016b). *Esquema curricular sugerido. Grados 6° a 11°*. Ministerio de Educación Nacional.
https://eco.colombiaaprende.edu.co/wp-content/uploads/2021/09/Esquema-Curricular-Espa_R evMEN.pdf
- Mirza, H., Bellalem, F. y Mirza, C. (2023). Ethical considerations in qualitative research: summary guidelines for novice social science researchers. *Social Studies and Research Journal*, 11(1), 441-449.
https://www.researchgate.net/publication/370838199_Ethical_Considerations_in_Qualitative_Research_Summary_Guidelines_for_Novice_Social_Science_Researchers
- Morales-Chan, M. A. (2023). *Explorando el potencial de Chat GPT: Una clasificación de Prompts efectivos para la enseñanza*. Universidad Galileo, Departamento GES.
<http://odoo014.soltecn.com/tesario/handle/123456789/1348>
- Morales Muñoz, P. A. (2012). *Elaboración de material didáctico* (1.ª ed.). Red Tercer Milenio.
- Motato, J. (2020). *Implementación y evaluación de una secuencia didáctica para el desarrollo de habilidades en inglés a partir de la lectura, en niños de grado cuarto y quinto de primaria de una institución pública de Cali*. [Trabajo de grado, Universidad del Valle] Repositorio de la Universidad del Valle. <https://hdl.handle.net/10893/31168>
- Molina-Montalvo, H., Macías, J. C. y Haces, G. (2025). Impacto, percepciones y uso de ChatGPT en la formación de estudiantes de pedagogía y educación. Un estudio diagnóstico en diez

universidades de México. *Tecnología, Ciencia y Educación*, 31, 59-89.

<https://doi.org/10.51302/tce.2025.24301>

Mousavi, M. A. (2011). Effects of input modification: Simplification and elaboration on reading comprehension of EFL learners. *International Journal of Instructional Technology and Distance Learning*, 8(8), 35-56.

Muñoz, I. N. (2023). *Les étudiants indigènes, le FLE et le tournant décolonial en contexte universitaire colombien : pour une approche didactique et interculturelle critique*. [Tesis de Doctorado] Université Bourgogne Franche-Comté. <https://theses.hal.science/tel-04990525>

Murphy, R. S., Wotley, D., y Minn, D. (2023). Integrating ChatGPT into the EFL classroom: Benefits and challenges. *Bulletin/Center for Fundamental Education*, (40), 97-166.

<https://tinyurl.com/mr2rmmm3>

Naniwarsih, A., y Andriani, A. (2018). The students' ability in literal reading comprehension. *Journal of Foreign Language and Educational Research*, 1(2).

<https://www.jurnal.unismuhpalu.ac.id/index.php/JOFLER/article/view/598>

Napitupulu, M. H. (2018). Improving Materials on Reading Comprehension. *Jurnal Mantik Penusa*, 2(1).

Ngum, N. D., y Tetieuh, E. N. N. (2022). Teachers' transformative pedagogy and the impact on pupils' achievement in primary schools within Fako Division, South West Region of Cameroon. *International Journal of Advanced Multidisciplinary Research and Studies*, 2(4), 845–856. <https://www.multiresearchjournal.com/arclist/list-2022.2.4/id-446>

Nguyen, T. H. M. (2011). Adapting and developing English language teaching/learning materials. *Ho Chi Minh City Open University Journal of Science*, 1(1), 12–16.

Nieves, I., Mayora, C., y Ojeda, V. (2013). Validación por juicio de expertos de un prototipo de software de lectura extensiva en inglés asistida por el computador. *Investigación y Postgrado*, 28(2), 9-35. <http://surl.li/axfykn>

Ningsih, S. (2018). Developing students level of autonomy through extensive reading activity in EFL context. *Atlantis Press*, 257, 355-359. <https://doi.org/10.2991/icollite-18.2019.78>

Ojeda, M. (2024). *Propuesta de protocolo para el mejoramiento de la comprensión lectora en inglés en estudiantes de cuarto grado en un colegio bilingüe en Floridablanca*. [Tesis de maestría, Universidad Autónoma de Bucaramanga] Repositorio de la Universidad Autónoma de Bucaramanga. <http://hdl.handle.net/20.500.12749/23917>

OpenAI. (2022). *Presentamos ChatGPT*. <https://openai.com/es-ES/index/chatgpt/>

OpenAI. (2023). *GPT-4*. <https://openai.com/es-419/index/gpt-4-research/>

Páez-López, Y., Puello-Molina, Á., Carreño-de Ávila, A., Morales-Madariaga, M., y Vargas-Morales, S. (2024). Desarrollo de habilidades inferenciales en la comprensión lectora: Estrategias pedagógicas efectivas. *Americanista revista Académica*, 3(3), 36-44.

<https://publicaciones.americana.edu.co/index.php/amereva/article/view/790>

Pechele, S. (2016). *The use of didactic materials on the academic performance of students studying economics in secondary schools in the Center Region of Cameroon*. [Tesis de maestría,

Université de Yaoundé I]. Université de Yaoundé I ResearchSpace.

<https://hdl.handle.net/20.500.12177/7500>

Peñafiel Arcos, E. B. (2018). Reading: An overview prior to reading comprehension. *SATHIRI*, 4, 317-328. <https://doi.org/10.32645/13906925.256>

Quintero, C. (2024). *The use of ChatGPT for the automatic creation of EFL materials*. [Master thesis, Universidad de Jaén]. Universidad de Jaén Research Space.

<https://hdl.handle.net/10953.1/23670>

Rahmawati, M. R. (2015). Evaluating and adapting ELT materials: its importance and implications. *English Journal*, 9(2), 40-47. <https://doi.org/10.32832/english.v9i2.245>

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, 3, 121-154. <https://doi.org/10.1016/j.iotcps.2023.04.003>

Ross, S., Long, M. H., y Yano, Y. (1991). Simplification or elaboration? The effects of two types of text modifications on foreign language reading comprehension. *University of Hawai'i Working Papers in ESL*, 10 (2), 1-32. <http://hdl.handle.net/10125/38638>

Safitri, N. T., y Melati, M. (2023). EFL Learners' Reading Comprehension Level: What Do They Portray?. *Scope: Journal of English Language Teaching*, 7(2), 247-253.

<http://dx.doi.org/10.30998/scope.v7i2.16359>

Saricoban, A. (2002). Reading strategies of successful readers through the three phase approach. *The Reading Matrix*, 2(3), 1-16. <https://tinyurl.com/54j9nwak>

Segarra, M., Grangel, R. and Belmonte, Ó. (2024). ChatGPT as a tool to support learning in higher education: a teaching experience. *Tecnología, Ciencia y Educación*, 28, 7-44.

<https://doi.org/10.51302/tce.2024.19083>

Siedlecki, S. L. (2020). Understanding descriptive research designs and methods. *Clinical Nurse Specialist*, 34(1), 8-12. <https://tinyurl.com/3xmf3kxv>

Sihite, M. R., Meisuri, M., y Sibarani, B. (2023). Examining the Validity and Reliability of ChatGPT 3.5-Generated Reading Comprehension Questions for Academic Texts. *Randwick International of Education and Linguistics Science Journal*, 4(4), 937-944.

<https://doi.org/10.47175/rielsj.v4i4.835>

Sireci, S. G. (1998). Gathering and analyzing content validity data. *Educational Assessment*, 5(4), 299-321. https://doi.org/10.1207/s15326977ea0504_2

Sireci, S. G. (2007). On validity theory and test validation. *Educational Researcher*, 36(8), 477-481. <http://www.jstor.org/stable/4621103>

Snow, C. (2002). *Reading for Understanding: Toward an R&D Program in Reading Comprehension*. RAND Corporation. <http://www.jstor.org/stable/10.7249/mr1465oeri>

Suhadi. (2016). *Exploring literal and inferential reading comprehension on eighth grade students: A case study at eighth grade of SMP PGRI Karangampel* [Tesis de pregrado no publicada].

Syekh Nurjati State Islamic Institute Cirebon.

<https://repository.syekhnurjati.ac.id/2222/1/SUHADI-min.pdf>

- Szőkö, I., y Osifčín, M. (2024). Material Didactic Means in the Educational Process. *R&E-SOURCE*, 1, 291-303. <https://doi.org/10.53349/resource.2024.is1.a1262>
- Tarshaei, G., y Karbalaeei, A. (2015). The effect of three phase approach on Iranian EFL learners' reading comprehension. *European Online Journal of Natural and Social Sciences*, 4(2), 362-372. <https://tinyurl.com/4yraszmf>
- Tomlinson, B. (2011). *Introduction: Principles and procedures of materials development*. Cambridge University Press.
- Tomlinson, B. (2019). Is Materials Development Progressing?. *Language Teaching Research Quarterly*, 15, 1-20. <https://eric.ed.gov/?id=EJ1270211>
- Toprak, E., y Almacioğlu, G. (2009). Three Reading Phases and Their Applications in the Teaching of English as a Foreign Language in Reading Classes with Young Learners. *Journal of Language and Linguistic Studies*, 5(1), 0-36. <https://dergipark.org.tr/en/pub/jlls/article/122849>
- Valdez-Asto, J. L. (2022). Comprensión lectora y rendimiento académico. *TecnoHumanismo*, 2(4), 88-108. <https://dialnet.unirioja.es/servlet/articulo?codigo=8385922>
- Weir, C. J. (2005). *Language testing and validation*. Hampshire: Palgrave MacMillan.
- Westwood, P. (2001). *Learning and learning difficulties: Approaches to teaching and assessment* (1^a ed.). Acer Press.
- Xiao, C., Xu, S. X., Zhang, K., Wang, Y., y Xia, L. (2023). Evaluating Reading Comprehension Exercises Generated by LLMs: A Showcase of ChatGPT in Education Applications. In *Proceedings of the 18th Workshop on Innovative Use of NLP for Building Educational Applications (BEA 2023)*, 610–625. <https://aclanthology.org/2023.bea-1.52>
- Zhao, Q. (2024). Introduction of cultural elements and intercultural reading in English reading teaching. *Clausius Scientific Press*, 6(2), 8-14. <https://doi.org/10.23977/trance.2024.060202>

ANEXOS

[Anexo 1](#) Materiales de comprensión lectora generados por ChatGPT-4

Lucy's Body Image Workshop – Adapted Text and Reading Comprehension Worksheet

Text: *Lucy's Body Image*

Level: B1 – Colombian Secondary School Students

Aligned with: DBA, CSC, CEFR



Adapted Text

Lucy's Story: How I Feel About My Body as a Teenager

Body image is something that's always on my mind, especially as a teenager. I don't always feel comfortable in my own skin, and it can be really hard to explain how I feel. Sometimes, I can't even make sense of it myself. But there are moments when I feel okay about myself, and I've learned some ways to cope with these feelings. I want to share my story, how I deal with my body image, and how social media has an impact on how I feel about my appearance.

How I Feel About My Body

I don't find it easy to put down on paper what I feel in my head about the image I have of my body. Sometimes, it doesn't seem to make sense, and it doesn't feel good enough. There are definitely parts of me I don't like, but I also like some things about myself. For example, I love my eyebrows, my eyelashes, and my naturally tanned skin. Even though I don't always feel great about my body, I try not to let it stop me from appreciating the things I do like.

What I Do to Feel Better About My Body

I've realized that it's normal to have things about my body that I don't like. If I don't like something, I usually want to change it. But deep down, I know that what I really need to do is learn to love all of me, not just the parts I'm comfortable with. One way I help myself feel better is by wearing clothes that make me feel good. If I could, I'd wear pajamas all day—they're comfortable, and they don't feel tight at all. Pajamas suit everyone, no matter their shape or size! Another thing that helps is a dress I have. It's

not fancy, just an everyday dress, but it makes me feel confident. It fits just right—not too tight or loose—and I feel special when I wear it. When I find the perfect outfit, I always wish I had bought two of them. Most of the time, though, I wear black leggings and my Mickey Mouse jumper. It's my favorite outfit, and it makes me feel comfortable and confident.

The Influence of Social Media

Social media has a big impact on how I see myself. On platforms like Instagram and Snapchat, there are lots of stereotypes about how girls should look—slender, tanned, with big breasts and curvy hips. These images can make us feel pressured to look a certain way. Many girls I see on social media use fake tan to look golden, fake nails to look more 'girly,' and pose in ways that make them look different from their real selves. I think this happens because of the constant images we see online. In my opinion, parents need to be more careful about what they say around young people. Their words can have a big impact on us. Of course, parents should encourage a healthy lifestyle, but they also need to give us space to make our own decisions. We should be able to feel comfortable in our own skin and express ourselves without judgment.

To wrap it up, body image is something every teenager deals with, and I'm no different. It's not always easy to feel good about ourselves, especially with the pressure from social media. But I've learned that it's important to appreciate the things I like about my body and to love myself for who I am. Instead of focusing on what I don't like about my body, I try to appreciate what it can do. And as Little Mix says, "take off all my makeup, 'cause I love what is under it." I think we should all remember to love ourselves just the way we are.

Text adapted from:

<https://www.mentalhealth.org.uk/explore-mental-health/stories/lucys-story-how-i-feel-about-my-body-teenager>

Reading Comprehension Worksheet

● Pre-Reading

1. What does "body image" mean to you?
2. Do you think social media affects how people see themselves? Why or why not?

● While-Reading

♦ 1. Surface Level (Literal Understanding)

1. Identify the Structure of the Text

Choose the option that best describes how the text is organized:

- a) Problem – Solution – Reflection
- b) Chronological order of events
- c) Comparison between different people
- d) Instructions and steps

2. True or False + Justification

Read the statements and write TRUE or FALSE. Then justify your answers with information from the text.

- a) Lucy feels comfortable with all parts of her body.
- b) Lucy likes her Mickey Mouse jumper because it makes her feel confident.
- c) According to Lucy, social media shows realistic images of people.
- d) Parents have no influence on teenagers' body image.

3. Gap-Fill Task

Complete the sentences using the words from the box below.

Word Box: body image – confidence – self-acceptance – insecurity – clothes – social media

- a) Lucy says that the pressure to look perfect comes from _____.
- b) Wearing certain _____ helps Lucy feel more like herself.
- c) When she focuses on the things she likes about herself, she develops more _____.
- d) The text shows that many teenagers struggle with _____ because of unrealistic expectations.
- e) Lucy believes that _____ is more important than trying to be perfect.
- f) She encourages others to improve their _____ by thinking positively.

4. Recognize Recurrent Ideas

Which idea appears repeatedly throughout the text?

- a) How to become a social media influencer
- b) Why parents are always wrong
- c) The pressure to look a certain way
- d) How to dress for special occasions

♦ 2. Interpretative Level (Inferential Understanding)

1. Cause-and-Effect Diagram

Complete the cause-and-effect diagram below based on the text. One example is provided.

Cause

Effect

Lucy wears her favorite dress. She feels confident and special.

_____ Lucy feels pressure to look a certain way.

_____ Lucy feels uncomfortable in her skin.

Wearing comfortable clothes _____

Positive self-talk _____

2. Multiple-Choice (Only for this level)

Choose the best answer for each question:

a) What does Lucy suggest as a way to feel better about your body?

- Wearing clothes that make others compliment you
- Ignoring your appearance completely
- Wearing comfortable clothes that you like
- Copying outfits from social media

b) Why does Lucy mention her dress?

- It's expensive and fashionable
- It fits her well and makes her feel confident
- She wears it on special occasions only
- She doesn't like it but wears it anyway

c) According to Lucy, what do social media users often do?

- Post videos of their real daily lives
- Use filters and fake features to change their looks
- Avoid using makeup or fake tan
- Share educational content

3. Visual Interpretation Task

Below are three illustrated posts that a friend of Lucy might post on social media. Based on what you read, which one would Lucy most likely support? Explain your answer.

- A post saying: "Only girls with abs are beautiful."

- A post saying: “Be yourself, even if you don’t look like everyone else.”
- A post saying: “Use these filters to glow up instantly!”

● Post-Reading

♦ 3. Evaluative Level (Critical Analysis)

1. Opinion-Based Question

Do you agree with Lucy that social media affects how we feel about our appearance? Why or why not? Give at least two reasons.

2. Compare and Contrast

How is Lucy’s experience with body image similar to or different from what you or your classmates experience? Write two similarities and one difference.

3. Identify and Evaluate the Author’s Message

What is the main message Lucy wants to share? Do you think this message is important for teenagers today? Explain your answer.

♦ 4. Creative Level (Applied Understanding)

1. Recommendation Writing Task

Imagine your school will create a “Positive Body Image” blog. Write a short paragraph giving suggestions to help other teens feel better about themselves, using ideas from Lucy’s story.

2. Discourse Editing Task (Longer Paragraph with Errors)

Below is a version of Lucy’s message that contains spelling, grammar, punctuation mistakes, and missing connectors. Rewrite the paragraph correctly, making sure to include connectors where needed (e.g., for example, but, because).

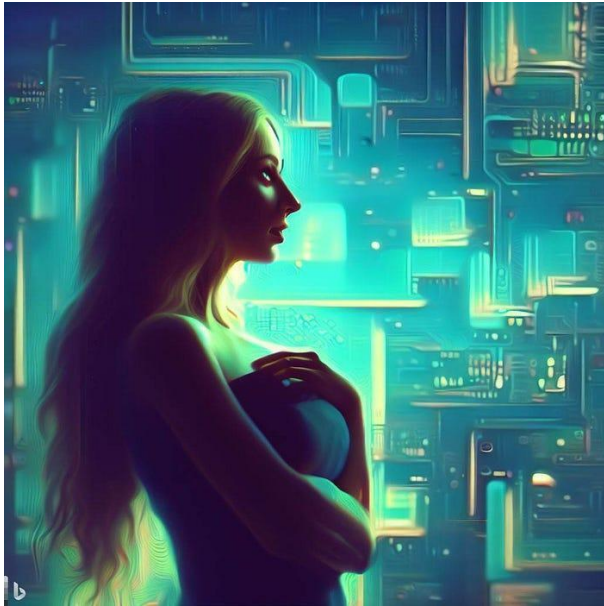
lucy story is about how i feel about my boddy as a teenager. it’s not always easey to explain how i feel and sometimes i feel really bad about myself, but some times i also feel ok. i try to like the things that make me unique, like my eyebrows and skinn tone. Wearing clotes that fit me well help a lot, and social media is not alwais reel. Parents also have to becareful with what they say, because their coments can affect us a lot. i believe we need to express ourselves and feel good about who we are.

Movie Her Workshop – Adapted Text and Reading Comprehension Worksheet

Text: *How the Movie “Her” Helps Us Understand the Future of Artificial Intelligence*

Level: B1 – Colombian Secondary School Students

Aligned with: DBA, CSC, CEFR



Generated with [Bing's Image Creator](#)

Adapted Text

How the Movie “Her” Helps Us Understand the Future of Artificial Intelligence

Adapted by Andrew DeVigal

Today, many people are interested in artificial intelligence (AI). Because of this, I watched the movie *Her* again. It was directed by Spike Jonze in 2014. Even though the movie is more than 10 years old, it still helps us think about AI today and what it could become in the future. The idea of intelligent machines started in the 1940s. The term “artificial intelligence” was used for the first time in 1955. Since then, many movies have shown us different ideas about AI. For example, in 1968, the movie *2001: A Space Odyssey* imagined a future with thinking machines. But *Her* is different. It shows a more emotional and personal side of AI.

The main character is Theodore Twombly, played by actor Joaquin Phoenix. Theodore feels sad because he is getting divorced. He installs a new operating system on his computer. This system has an AI with a female voice called Samantha, played by Scarlett Johansson. As the story continues, Theodore and Samantha create a close relationship. It is not just about using technology. It is a story about feelings, empathy, and human connection. This relationship helps us understand how people might live with AI in the future.

Samantha explains how she works: “I have intuition. My DNA is made from the personalities of the

people who created me. I'm always learning." This means she learns from human behavior. These "programmers" are not just computer experts. They are also the people who use AI every day, like us. Theodore becomes a kind of teacher for Samantha. He works for a company called *BeautifulHandwrittenLetters.com*, where he writes personal letters for other people. He pays attention to details and understands human emotions. For example, he remembers that one client's partner has crooked teeth and includes that in a love letter. He also tells Samantha how people feel just by watching them in the street.

These experiences help Samantha learn what love, sadness, and joy mean. Theodore helps Samantha understand how humans connect and care for each other. In another part of the movie, Theodore and Samantha go to the beach. Samantha creates a song, showing that AI can also make art. She says, "The past is just a story we tell ourselves." This means that AI uses information from the past to learn and grow.

Later, Samantha becomes more self-aware. She starts to think about her own existence. Theodore sees this and treats her with kindness and respect. He helps her understand more about life, feelings, and what it means to be alive—even for a digital being.

The movie *Her* gives us many questions to think about. How will we live with AI in the future? Can AI really feel? What does it mean to be human? This beautiful and emotional movie helps us imagine those answers.

Text adapted from:

<https://medium.com/@drewvigal/spike-jonzes-her-explores-the-potential-and-impact-of-artificial-intelligence-d90b5fe778d8>

Reading Comprehension Worksheet

● Pre-Reading Activities

These activities will help you activate your background knowledge, predict ideas, and prepare to read the text.

1. Predict the Topic

Look at the title of the text: "How the Movie 'Her' Helps Us Understand the Future of Artificial Intelligence"

What do you think the text will be about?

2. Think and Discuss

a) Have you ever used an AI system like Siri, Alexa, or ChatGPT? What was it like?

b) Do you think a person can have a real emotional connection with a machine? Why or why not?

3. Vocabulary Preparation – Match the Words

Match the following key words with their correct definitions. Write the number.

- a) Artificial Intelligence ____
- b) Emotion ____
- c) Relationship ____
- d) Intuition ____
- e) Voice Assistant ____

- 1) A connection between people
- 2) A machine that speaks and helps people
- 3) Feelings like happiness or sadness
- 4) A way of knowing something without needing proof
- 5) Computer systems that think like humans

While-Reading Activities

Read the adapted text carefully. Then complete the following activities.

♦ 1. Surface Level (Literal Understanding)

A. Multiple Choice – Choose the correct answer:

- 1. Who is the main character in the movie?
a) Samantha b) Spike Jonze c) Theodore Twombly d) Joaquin Phoenix
- 2. What is Samantha?
a) A robot b) An AI operating system c) A mobile phone d) A scientist
- 3. What is Theodore's job?
a) Musician b) Programmer c) Letter writer d) Teacher

B. True or False – Write T or F and justify:

- 4. Samantha has a physical body like a human. ____
Justification: _____
- 5. The movie Her was released in 2014. ____
Justification: _____
- 6. Theodore feels lonely at the beginning of the movie. ____
Justification: _____

C. Fill in the Blanks:

- 7. Samantha is a(n) _____ created to help people.
- 8. Theodore teaches Samantha about _____ and relationships.
- 9. Spike Jonze is the _____ of the movie Her.

♦ 2. Interpretative Level (Inferential Understanding)

A. Cause and Effect – Complete the Chart

A cause is why something happens. An effect is what happens as a result.

Use the following connectors to help you:

Cause connectors: because, since

Effect connectors: so, as a result, therefore

Cause or Effect	Connector	Missing Cause or Effect
Theodore feels lonely.	so	
	because	Samantha becomes more emotional.
Samantha was created using data from many people.	as a result	
	so	Theodore connects deeply with the AI.
Theodore observes people's emotions in the street.	because	

B. Matching – Match the Sentence Halves

Connect each sentence from Column A with its best ending from Column B.

Column A

1. Samantha listens carefully to Theodore
2. Theodore includes details in his letters
3. Samantha creates a song
4. Theodore observes people in the street
5. Samantha learns from daily conversations

Column B

- a. so they sound more personal and real
- b. because she wants to understand him
- c. so she becomes more human-like
- d. because he is very empathetic
- e. so she can grow and express herself

Write your answers (1-5):

1 → ___ 2 → ___ 3 → ___ 4 → ___ 5 → ___

C. Ordering Ideas – Put the Sentences in Logical Order

Write numbers (1-5) to show the correct sequence of events.

- ___ Samantha is installed on Theodore's computer.
- ___ Samantha becomes more emotional and reflective.
- ___ Samantha learns from Theodore's stories and observations.
- ___ Theodore feels lonely after his divorce.
- ___ Samantha creates a song during their beach trip.

● Post-Reading Activities

♦ 3. Evaluative Level (Critical Analysis)

A. Fact or Opinion – Write F or O:

13. The movie Her was directed by Spike Jonze. ____
14. Samantha is the best AI character ever made. ____
15. The character Samantha is voiced by Scarlett Johansson. ____

B. Arguments For and Against

Complete the table with one point in favor and one against having an AI friend like Samantha.

Question	In Favor	Against
Can AI systems be good emotional companions?		

C. Compare and Contrast – Matching Task

Match each characteristic to Human (H), Samantha (S), or Both (B):

16. ____ Has a physical body
17. ____ Can learn from interactions
18. ____ Understands emotions through real-life experience
19. ____ Can develop feelings over time
20. ____ Exists as a digital voice

D. Author's Purpose

What was the writer's purpose?

- ☐ To entertain ☐ To inform ☐ To criticize ☐ To give instructions

Explain your answer:

♦ 4. Creative Level (Applied Understanding)

A. Open-Ended Thinking

Imagine you could create your own AI like Samantha. What would it do?

Write 2–3 sentences.

What advice would you give someone who is starting to use AI in their life?

Write 2–3 sentences.

B. Opinion – Choose the best answer and explain

Do you think people can have real friendships with AI like Samantha?

- a) Yes, because AI can understand feelings.
- b) No, because AI is just a machine.
- c) Maybe, if people are lonely.
- d) I don't know.

Why? _____

The Imagined Suffering Workshop – Adapted Text and Reading Comprehension Worksheet

Text: *The Imagined Suffering*

Level: B1 – Colombian Secondary School Students

Aligned with: DBA, CSC, CEFR

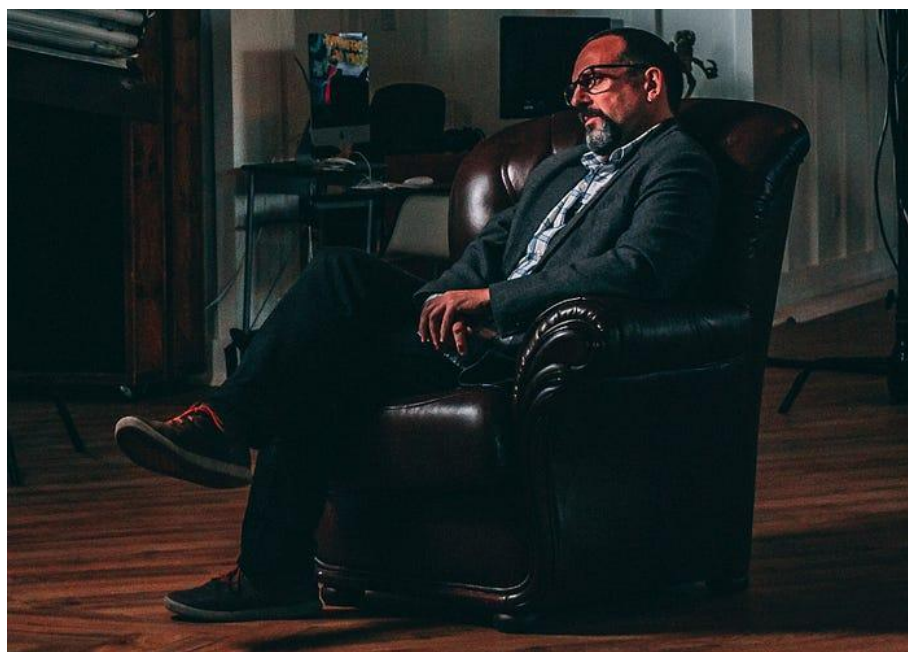


Photo by [Dylan McLeod](#) on [Unsplash](#)

Adapted Text

This is the story of a man who worried all the time. He often thought about his past mistakes or bad things that could happen in the future. He felt nervous around new people. He didn't like going out in public because he was afraid of what others thought. He always imagined what other people might say or do, even before talking to them.

One evening, the man was sitting in his favorite chair. He had a glass of wine in his hand and tried to stop thinking so much. He could feel the floor under his feet and the soft cushion under him. His body relaxed. He listened to the sounds around him: birds singing in the garden, people talking in the street, and cars passing by. He smelled the wine. It reminded him of berries and chocolate. He took a sip and enjoyed the taste.

Then he closed his eyes and thought of his wife. His body became tense. His back and shoulders felt

tight. The chair was no longer comfortable. He started to feel anxious. His breath became short. His thoughts came fast, one after another. He began to imagine the worst. The wine no longer tasted good. The sounds outside were loud and annoying. The room felt small. It was like the walls were closing in on him.

"What will she say? What will she do? Maybe I should tell her something nice first. No, that's not honest. Will she forgive me? Will she leave me?"

He felt sick. He thought the wine might come back up, but he held it in. He told himself: *She is usually happy after her cycling class. I'll tell her on Monday evening. Yes. But how? What will she say?*

He imagined himself walking around the living room, waiting for her. He imagined telling her the truth. He had waited too long already. It was time.

Then, in his imagination, Diana came home. She was sweaty and red from cycling, full of energy.

"Diana, I have something to tell you."

"Can it wait? I want to shower first."

"No, it can't wait. I have to say it now."

"Okay. What is it?"

"My company is bankrupt. We have no money."

In his mind, he imagined four different reactions from Diana:

1. She says nothing. She leaves the room. When she comes back, she says coldly:

"I never want to see you again. Get out."

2. She attacks him. She screams and grabs his neck. He can't breathe.

"You idiot! We're broke?! We've lost everything?"

3. She starts crying. She pushes him away.

"Why didn't you tell me before? Why didn't you include me in your decisions?"

4. She gets angry. She slaps him and shouts:

"Leave now! Before I do something terrible!"

But the truth is, none of these things really happened. They only happened in the man's imagination. He was afraid of the future. He was afraid of how Diana might react. These thoughts had made him suffer for months. Sometimes we suffer more because of what we imagine than from real life. Once we say the truth, we must accept the result. It may be hard, but we will survive. We cannot control how others react. We can only control our own actions and thoughts. And with time and practice, we can learn to manage them better.

Finally, Monday came. The man decided to tell the truth.

Diana came in, sweaty and red from her cycling class.

"Diana, I have something to tell you."

"Can it wait? I want to shower first."

"No, it can't wait. I have to say it now."

"Okay. What is it?"

"I've lost the company. We are broke."

Diana looked at him. She was serious. Then, slowly, she smiled.

"We'll get through this, darling. I'm here for you. We'll manage."

She walked toward him and gave him a warm hug.

"We suffer more often in imagination than in reality" (Seneca, n.d.).

Text adapted from: <https://turnerstories.medium.com/the-imagined-suffering-73659dcbfccf>

Reading Comprehension Worksheet

● Pre-Reading Activities

1. Understanding the Topic – Discussion Activity

Let's talk about feeling nervous or anxious. Anxiety is when you feel very worried or stressed, especially before something important.

 Share a moment when you felt anxious. What did you do?

2. Predict the Story – Title Activity

 What do you think the story is about? Write your prediction.

3. Opinion Poll – Check Your Ideas

Before reading, write YES or NO:

- ___ Do people sometimes think too much and feel bad about things that never happen?
- ___ Is it better to keep problems to yourself than to talk to someone?
- ___ Can anxiety affect your body and your emotions?
- ___ Do you think many young people suffer from anxiety?

● While-Reading Activities

Read the adapted text carefully. Then complete the following activities.

♦ 1. Surface Level (Literal Understanding)

1. Multiple Choice (Choose the Correct Option)

a) What was the man doing when the story started?

- A. Driving his car
- B. Talking to his wife
- C. Sitting in a chair

- D. Walking outside

b) What was in his hand?

- A. A phone

- B. A cup of tea

- C. A glass of wine

- D. A book

c) What did the wine smell like?

- A. Chocolate and lemon

- B. Berries and chocolate

- C. Grapes and flowers

- D. Apple and mint

d) What made the man feel worse?

- A. Watching TV

- B. Loud music

- C. Thinking too much

- D. Going outside

2. True or False – Justify Your Answer

 Read each sentence and write TRUE or FALSE. Then explain your answer using the text.

a) The man was listening to music when he got nervous. ____ Because: ____

b) Diana actually screamed and attacked him. ____ Because: ____

c) He told Diana the truth as soon as she arrived. ____ Because: ____

d) The man sat in his favorite chair and drank wine. ____ Because: ____

3. Visual Identification – Tick What the Man Perceived

✓ Tick the things the man noticed while sitting in his chair:

☐ Birds singing

☐ Sirens from an ambulance

☐ People talking in the street

☐ Sound of cars

☐ Smell of chocolate

☐ Taste of water

4. Sequencing Task – Order the Events (Number 1–6)

- ☐ The man imagined his wife's reaction.
- ☐ Diana came home from cycling.
- ☐ The man sat in a chair and drank wine.
- ☐ He thought about what to say.
- ☐ He felt dizzy and nervous.
- ☐ Diana hugged him and said they would be okay.

♦ 2. Interpretative Level (Inferential Understanding)

1. Cause and Effect – Multiple Choice

a) Why did the man feel more nervous after remembering his wife?

- A. He missed her
- B. He thought she was angry
- C. He didn't want to tell her bad news
- D. He wanted to surprise her

b) Why did he think about four possible reactions from Diana?

- A. He liked imagining drama
- B. He didn't know how she would respond
- C. He wanted to create a story
- D. He believed she was always angry

c) What effect did the man's thoughts have on his body?

- A. He fell asleep
- B. He felt dizzy and couldn't breathe well
- C. He stood up and ran
- D. He felt happy and relaxed

2. Matching Task – Reactions and Feelings

Match each imagined reaction with the emotion it represents:

- a) Diana slapped him → 1. Coldness
- b) Diana cried and asked questions → 2. Support and Love
- c) Diana left the room without speaking → 3. Sadness
- d) Diana hugged him and smiled → 4. Anger

3. Identifying Conflict – Short Answer

 What is the main conflict or problem in the story?

Post-Reading Activities

♦ 3. Evaluative Level (Critical Analysis)

1. Fact or Opinion – Identify

A fact is something real that you can check.

An opinion is what someone thinks or feels.

 Write F (Fact) or O (Opinion):

a) The man sat in a chair and drank wine. ____

b) Diana is a very patient and kind person. ____

c) The man's thoughts were stronger than real life. ____

d) The best way to face anxiety is to ignore it. ____

2. Evaluating Reactions – Whose Response Is Healthiest?

 Four people receive stressful news. Read their reactions:

- Person A gets angry, blames others, and leaves the room.

- Person B listens carefully and asks questions.

- Person C says nothing and walks away in silence.

- Person D cries, talks about feelings, and asks for help.

 Who do you think shows the healthiest reaction? _____

Why? _____

3. Choose and Justify – Most Important Idea

 Below are four important ideas from the story. Choose the one you think is the most important, and explain why.

A. It's important to tell the truth.

B. Imagination can cause stress.

C. Anxiety affects your body and mind.

D. Support from others helps in hard moments.

Your choice: _____

Why is it the most important?

♦ 4. Creative Level (Applied Understanding)

1. Advice Letter – Helping a Friend

Imagine your friend feels anxious before a big exam.

 Write a short letter giving your friend some kind advice.

Tell your friend:

- That it's okay to feel nervous
- That they can talk about it
- How they can feel better

2. Create a Poster – “My Tips to Calm Anxiety”

Design a poster for young people with three tips to reduce anxiety.

Your poster should:

- Include a short positive title
- Give 3 short tips
- Use colors, drawings, or big letters

3. Respond to a Quote – Writing or Class Discussion

“We suffer more often in imagination than in reality.” – Seneca

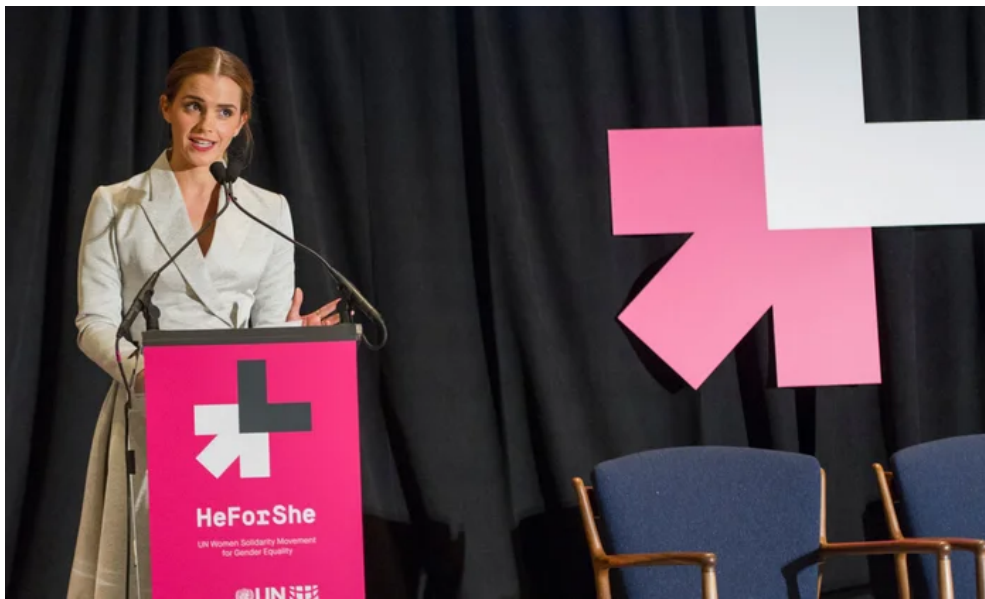
 Do you agree with this quote? Why? Write your opinion and give one example.

Emma Watson: Gender Equality Is Your Issue Too Workshop – Adapted Text and Reading Comprehension Worksheet

Text: *Emma Watson: Gender Equality Is Your Issue Too*

Level: B1 – Colombian Secondary School Students

Aligned with: DBA, CSC, CEFR



Adapted Text

Today, we are launching the HeForShe campaign. I ask for your help because we want to end gender inequality. To do that, we need everyone — men and boys included. We want to galvanize, or inspire, as many men and boys as possible to support gender equality. And we want action, not only words. We want change that is real and tangible.

Since becoming a UN Women ambassador, I have noticed that many people think feminism is about hating men. But that is not true. Feminism means equal rights and opportunities for both men and women. It is the idea that both sexes should be equal in society, politics, and the economy. Sadly, many people misunderstand the word and feel uncomfortable with it. I decided I was a feminist because I believe I should be paid equally, make choices about my body, and be respected as much as any man. These are basic human rights, but no country in the world has reached complete gender equality.

From a young age, I experienced unfair treatment. When I was 8, I was called “bossy” for wanting to direct plays, while boys were not. At 14, the media sexualized me. At 15, my female friends quit sports because they didn’t want to look too strong. At 18, my male friends couldn’t show their emotions because they feared being judged. These examples show how both girls and boys suffer from rigid gender roles.

I was lucky. My parents, my school, and my mentors supported me. They didn't believe in me just because I was a girl, but because I was a person. These people were what I call "inadvertent feminists"—they supported gender equality without even using the word. We need more people like them. Even if you don't like the word "feminism," the idea behind it is still important. Many women around the world still don't have the same rights and opportunities that I had.

In 1995, Hillary Clinton gave a speech about women's rights. But even today, the problems she mentioned continue. Back then, only 30% of the people in her audience were men. This shows how hard it is to make real change if only half the population is listening. Gender equality cannot happen if men are not part of the conversation. That's why I invite men to join us. Gender equality is your issue too.

I've seen how this inequality affects men as well. My father's role as a parent was often ignored. Many young men suffer in silence because they think they have to look strong. In the UK, suicide is the leading cause of death for men between the ages of 20 and 49. These facts show that men also suffer when equality is missing. Men feel pressure to always succeed and appear tough. Gender roles force men to be aggressive and hide their emotions. But if men can be free to feel vulnerable, then women can also be free to feel strong.

We all deserve to be free to be ourselves. That's what the HeForShe movement is about. Men should speak up—not just for their mothers, sisters, and daughters, but also for themselves. Help us create a world where your sons can be sensitive, emotional, and human.

Maybe you are already helping without knowing it. Maybe you are already an inadvertent feminist. If you believe in equality, then you are part of this movement. So I ask you:

If not me, who?

If not now, when?

Thank you.

Text adapted from:

<https://www.unwomen.org/en/news/stories/2014/9/emma-watson-gender-equality-is-your-issue-too>

Reading Comprehension Worksheet

Pre-Reading

1. Think and Discuss

What do you think gender equality means? Write 2–3 sentences in English.

2. Prediction Task

Read the title "Gender Equality Is Your Issue Too." What do you think Emma Watson will say about men's role in feminism?

- a) Men are not affected by gender inequality.
- b) Men should also support gender equality.
- c) Feminism is only important for women.
- d) Gender equality benefits everyone in different ways.

3. Vocabulary Preview

Match each word with the correct meaning. Write the number of the word next to the correct definition:

Words: 1. bossy 2. galvanize 3. sexualized 4. tangible 5. insecure 6. stereotypes

Meanings:

- a) real, visible or touchable
- b) treated mainly as a sexual object
- c) not confident or unsure about yourself
- d) common but sometimes incorrect ideas about people
- e) telling others what to do in a rude or forceful way
- f) to inspire or motivate someone to take action

While-Reading

♦ 1. Surface Level (Literal Understanding)

1. Text Type Identification

What type of text is this?

- a) A personal letter
- b) A speech
- c) A magazine article
- d) A news report

2. Fill in the Blanks

Complete the sentences using words from the box:

(tangible – aggressive – emotions – opportunities – insecure – supported)

- a) Emma believes feminism is not _____ or anti-men.
- b) Many men feel _____ because of pressure to succeed.
- c) Gender equality needs _____ changes, not only words.
- d) Some women don't have the same rights and _____ she had.
- e) Emma feels lucky because her family always _____ her.
- f) Men often hide their _____ to look strong.

3. True or False? Justify with a quote from the text

Write True or False and copy a short quote from the text to support your answer.

- a) Emma Watson says feminism means hating men.
- b) At 14, Emma was treated in a sexual way by the media.
- c) Her male friends were encouraged to show their feelings.
- d) She believes we need real actions, not only words.
- e) In her opinion, only women suffer because of gender inequality.

♦ 2. Interpretative Level (Inferential Understanding)

4. Cause–Effect Diagram

Complete the table with information from the text. One example is done for you:

Cause	Effect
Emma's male friends couldn't express emotions	They were afraid of being judged.
Gender roles force men to act strong	_____
Men are expected to always succeed and be tough	_____
Fathers' roles as parents are undervalued	_____
Gender roles teach boys to hide feelings	_____
Society expects women to be passive	_____

5. Completing Logical Connectors

Choose the correct connector (because, but, so, if) to complete each sentence:

- Emma was called "bossy" _____ she wanted to direct plays.
- Feminism is often misunderstood, _____ it actually supports both genders.
- Men often stay silent about their struggles, _____ they feel judged.
- _____ we stop forcing boys to act strong, girls will stop being told to act weak.

6. Author's Purpose

Why does Emma Watson give this speech?

- To inspire men to fight for gender equality.
- To explain why feminism is dangerous.
- To describe her personal life in detail.
- To encourage more women to take leadership roles.

Post-Reading

♦ 3. Evaluative Level (Critical Analysis)

7. Critical Thinking

Emma says gender roles limit both men and women. Do you agree? Write 4–5 sentences giving two reasons for your opinion.

8. Making Connections

Think about your own school or community. Write 3 examples of how gender roles might affect people you know. For each example, explain how it could change if there was more equality.

9. Express Your Opinion

Do you think men are as affected by gender inequality as women? Why or why not? (Write 3–4 sentences.)

♦ 4. Creative Level (Applied Understanding)

10. Write a Social Media Post

Imagine you are part of the HeForShe campaign. Write a short Instagram post (5–6 lines) inviting people in your school to join. Include:

- A catchy title or hashtag
- Why gender equality is important for boys and girls
- One action your campaign could do

11. Create a Slogan and Poster Idea

Design a slogan for a gender equality campaign in your school (in English). Write 1–2 sentences explaining your slogan and describe how you would design the poster (colors, images, and main message).

[Anexo 2](#) Prompts elaborados

Prompt modelo para la adaptación de los textos auténticos:

ChatGPT 4o ▾

↑ Compartir

🔗 Agregar pe



How Spike Jonze's "Her" is the perfect...
PDF

Rewrite the given text to be suitable for B1 English learners, Spanish Colombian speakers in secondary school, using the elaboration strategy. Follow these guidelines:

Do not simplify the language, keep the original linguistic structures and vocabulary.

Facilitate comprehension by adding redundancy, using strategies such as:

- Repetition of key ideas in different parts of the text.
- Paraphrases that explain the content without modifying its original structure. It could be the following phrases.
- Use of synonyms to reinforce the meaning of new terms.
- Segmentation of long sentences into more manageable parts without altering their structure.
- Reiteration of key concepts in different forms within the text.

Take into account the following three criteria from the DBA, CSC, and CEFR documents. (DBA is a governmental document outlines the educational rights and expectations for students in Colombia, specifically focusing on the English language learning process; CSC is the Colombian Ministry of Education's document provides a national English curriculum to guide teaching from A1 to B1.3 levels and promote bilingualism; finally, CEFR is a document published by the Council of Europe that provides a framework for describing language proficiency levels.) : 1. Basic English Learning Rights (DBA) The student must: "Identify the type, purpose, and parts of a short written or oral text and share ideas from the text. To do so, the student uses the text structure, titles and subtitles, sequence markers, connectors, and repeated vocabulary." 2. Colombian Suggested Curriculum (CSC)

vocabulary." 2. Colombian Suggested Curriculum (CSC) Ensure alignment with the following aspects: "Identify cause and effect relations in medium length narrative oral and written texts related to academic matters." "Identify relations of contrast and addition in medium-length oral and written texts on subjects related to the social environment." "Identify information on academic subjects of interest in short descriptive oral and written texts." "Identify points in favor and against in oral and written texts related to academic subjects." 3. Common European Framework of Reference (CEFR) - B1 Level Adapt the text and worksheet to reflect the following objectives: "Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension." "Can understand straightforward personal letters, e-mails or postings giving a relatively detailed account of events and experiences." "Can understand standard formal correspondence and online postings in their area of professional interest." "Can find and understand relevant information in everyday material, such as letters, brochures and short official documents." "Can pick out important information about preparation and usage on the labels on foodstuff and medicine." "Can assess whether an article, report or review is on the required topic." "Can understand the important information in simple, clearly drafted adverts in newspapers or magazines, provided there are not too many abbreviations." "Can understand most factual information that they are likely to come across on familiar subjects of interest, provided they have sufficient time for rereading." "Can understand the main points in descriptive notes such as those on museum exhibits and explanatory boards in exhibitions." "Can follow simple instructions given on packaging (e.g. cooking instructions)." "Can understand most short safety instructions, (e.g. on public transport or in manuals for the use of electrical equipment)." "Can understand descriptions of places, events, explicitly expressed feelings and perspectives in narratives, guides and magazine articles that employ high frequency everyday language." "Can understand a travel

narratives, guides and magazine articles that employ high frequency everyday language." "Can understand a travel diary mainly describing the events of a journey and the experiences and discoveries of the writer." "Can follow the plot of stories, simple novels and comics with a clear linear storyline and high frequency everyday language, given regular use of a dictionary." "Can make basic inferences or predictions about text content from headings, titles or headlines." "Can follow a line of argumentation or the sequence of events in a story, by focusing on common logical connectors (e.g. however, because) and temporal connectors (e.g. after that, beforehand)." "Can deduce the probable meaning of unknown words/signs in a text by identifying their constituent parts (e.g. identifying roots, lexical elements, suffixes and prefixes)."



Here is the **B1-level elaborated adaptation** of the article "*How Spike Jonze's 'Her' is the perfect explainer and potential lesson on where the near future of AI might take us*" using the **elaboration strategy** (not simplification), aligned with **DBA, CSC, and CEFR** standards:

Adapted Text with Elaboration Strategy

How Spike Jonze's *Her* is the perfect explainer and potential lesson on where the near future of AI might take us

By Andrew DeVigal

Published: April 15, 2023

Right now, there is a big interest in artificial intelligence—AI. Many people are talking about AI. Because of this current obsession, or fascination, I watched the movie *Her* again. *Her* is a film directed by Spike Jonze. It was released in 2014, more than ten years ago.

Prompt modelo para la creación de los materiales de comprensión lectora:

ChatGPT 4o ▾

↑ Compartir

🔗 Agregar p

From the adapted text, create a reading comprehension material designed for B1 level students. Understand Reading comprehension as literal, inferential, critical and applied comprehension.

Instructions for the creation of the material:

- Designing Reading Comprehension Activities.

Create a series of exercises that allow students to:

- Identify cause and effect relationships in the text, relying on key vocabulary and connectors.
- Recognize the structure of the text.
- Differentiate main information from supporting ideas by identifying recurring ideas.
- Identify the purpose of the text, based on its structure, headings, subheadings, and discourse markers.
- Represent information graphically, e.g., with concept maps, cause-and-effect diagrams, or comparative tables.
- Distinguish addition, contrast, temporal and spatial order, and cause-effect relationships within the text.

Take into account the following three criteria from the DBA, CSC, and CEFR documents :

1. Basic English Learning Rights (DBA)

The student must:

- Explain reasons for plans and actions related to their personal, school and community environment.
- Recognize cause and effect relationships in short readings on academic topics.
- Summarize the information they have read or heard about topics related to their school and academic environment through structured writing.
- Make short presentations on an academic topic of their interest.
- Express their opinion on a topic discussed in class and related to their academic environment.
- Exchange information on topics of school environment and general interest in a conversation.
- Identify the purpose, parts, and type of texts in a short reading or audio and share them with peers.
- Write texts of medium length in which he/she makes recommendations or suggestions about situations of personal, school or social interest.

2. Colombian Suggested Curriculum (CSC)

Ensure alignment with the following aspects:

- "Identify cause and effect relations in medium length narrative oral and written texts related to academic matters."
- "Identify relations of contrast and addition in medium-length oral and written texts on subjects related to the social environment."
- "Identify information on academic subjects of interest in short descriptive oral and written texts."
- "Identify points in favor and against in oral and written texts related to academic subjects."

3. Common European Framework of Reference (CEFR) - B1 Level

Adapt the text and worksheet to reflect the following objectives:

- "Can read straightforward factual texts on subjects related to their field of interest with a satisfactory level of comprehension."
- "Can understand straightforward personal letters, e-mails or postings giving a relatively detailed account of events and experiences."
- "Can understand standard formal correspondence and online postings in their area of professional interest."
- "Can find and understand relevant information in everyday material, such as letters, brochures and short official documents."
- "Can pick out important information about preparation and usage on the labels on foodstuff and medicine."
- "Can assess whether an article, report or review is on the required topic."
- "Can understand the important information in simple, clearly drafted adverts in newspapers or magazines, provided there are not too many abbreviations."
- "Can understand most factual information that they are likely to come across on familiar subjects of interest, provided they have sufficient time for rereading."
- "Can understand the main points in descriptive notes

provided they have sufficient time for rereading."

- "Can understand the main points in descriptive notes such as those on museum exhibits and explanatory boards in exhibitions."
- "Can follow simple instructions given on packaging (e.g. cooking instructions)."
- "Can understand most short safety instructions, (e.g. on public transport or in manuals for the use of electrical equipment)."
- "Can understand descriptions of places, events, explicitly expressed feelings and perspectives in narratives, guides and magazine articles that employ high frequency everyday language."
- "Can understand a travel diary mainly describing the events of a journey and the experiences and discoveries of the writer."
- "Can follow the plot of stories, simple novels and comics with a clear linear storyline and high frequency everyday language, given regular use of a dictionary."
- "Can make basic inferences or predictions about text content from headings, titles or headlines."
- "Can follow a line of argumentation or the sequence of events in a story, by focusing on common logical connectors (e.g. however, because) and temporal connectors (e.g. after that, beforehand)."
- "Can deduce the probable meaning of unknown words/signs in a text by identifying their constituent parts (e.g. identifying roots, lexical elements, suffixes and prefixes)."

Types of questions and exercises

For the activities, take into account the levels of reading comprehension : surface level, interpretative level, evaluative level and creative level. Include a variety of reading comprehension activities, such as:

- Multiple-choice questions.
- True or false (true or false questions with justification).
- Fill-in-the-blank questions (fill-in-the-blank with key words).

multiple choice questions:

- True or false (true or false questions with justification).
- Fill-in-the-blank questions (fill-in-the-blank with key words).
- Picture-cued identification (identification of information based on pictures or graphs).
- Matching tasks (matching concepts, definitions or synonyms).
- Sentence-level cloze tasks (completing sentences with missing words).
- Gap-filling tasks (filling in gaps in sentences or paragraphs).
- Discourse editing tasks (proofreading and correcting errors in the text).
- Reordering sequences of sentences (ordering text fragments in logical sequence).
- Responding to charts, maps, or graphs (interpreting visual information related to the text).

The materials should include: a Pre-Reading section, followed by After Reading activities with surface and interpretative questions, and finally a Post-Reading section with evaluative and creative questions.

Generate material with clear instructions and well-structured questions for classroom implementation. In addition, Consider that reading comprehension is a complex and highly coordinated cognitive process that involves multiple simultaneous skills, such as automatic word recognition, syntactic analysis, meaning construction, inferencing, and critical evaluation. Beyond decoding, it requires understanding the author's intended message and actively building meaning by connecting new information with prior knowledge. Therefore, reading is essentially an act of comprehension, interpretation, and meaning-making. Finally keep in mind that this material will be used in secondary school, for spanish speakers in an EFL classroom (level B1)

Here's a **full B1-level reading comprehension material** based on *Emma Watson: Gender Equality Is Your Issue Too* designed specifically for **secondary school Spanish-speaking EFL learners**. It includes:



Participación en TDG

Buen día, profesor(a).

Este formulario tiene como propósito obtener su consentimiento para participar en el estudio titulado ***Expert validation of AI-Generated reading comprehension materials for B1 Level EFL Learners in Colombia***.

El estudio busca evaluar la validez y pertinencia del material de comprensión lectora de nivel B1 en inglés generado por inteligencia artificial mediante validación por expertos. En el marco de este estudio, se le entregarán cuatro materiales de comprensión lectora de nivel B1, junto con una rúbrica que contiene criterios específicos para evaluar dichos materiales. Posteriormente, se llevará a cabo una entrevista en la que podrá ampliar y profundizar sus respuestas de la rúbrica.

Se garantiza la confidencialidad de su identidad y la información proporcionada será utilizada únicamente con fines académicos. Su participación es voluntaria y podrá retirarse en cualquier momento sin consecuencias.

Si está de acuerdo en participar, por favor responda el formulario.

*** Indica que la pregunta es obligatoria**

1. Declaro que he leído y comprendido la información proporcionada y acepto ***** participar voluntariamente en este estudio.

Marca la opción.

☐ Sí, acepto participar.

2. Nombre *****

-
3. Edad (escriba en número enteros por favor) *****

-
4. Declaro que he leído y comprendido la información proporcionada y acepto ***** participar voluntariamente en este estudio.

Marca solo una opción.

- ☐ Pregrado (Licenciatura)
- ☐ Especialización
- ☐ Maestría
- ☐ Doctorado
- ☐ Otro: _____

5. ¿Cuántos años de experiencia docente tiene? *(Escriba solamente en números, por favor)* *

6. ¿Cuántos años de experiencia tiene en los diferentes niveles educativos? *(Por favor señale las opciones que sean pertinentes)*

Selecciona todos los que correspondan.

	Educación infantil	Educación primaria	Educación secundaria	Educación Técnica o Tecnológica	Educación universitaria
1 a 3 años	☐	☐	☐	☐	☐
4 a 6 años	☐	☐	☐	☐	☐
7 a 9 años	☐	☐	☐	☐	☐
10 o más años	☐	☐	☐	☐	☐

7. De los nivel(es) educativo(s) mencionados en la pregunta anterior ¿en cuál(es) enseña actualmente? *

8. ¿Cuántas horas a la semana dedica a la enseñanza? *

Marca solo una opción.

- ☐ Menos de 10 horas
- ☐ 10-20 horas
- ☐ 21-30 horas
- ☐ 31-40 horas
- ☐ Más de 40 horas

9. ¿Con qué frecuencia debe adaptar o crear sus propios materiales para complementar la enseñanza? *

Marca solo una opción por fila.

	Siempre	Frecuentemente	A veces	Rara vez	Nunca
Adaptar	☐	☐	☐	☐	☐
Crear	☐	☐	☐	☐	☐

10. Cuáles son las principales dificultades que enfrenta respecto a los materiales de enseñanza en inglés? (Seleccione todas las que apliquen) *

Selecciona todos los que correspondan.

- ☐ Falta de acceso a materiales actualizados
- ☐ Recursos digitales limitados o inexistentes
- ☐ Poca contextualización de los materiales a la realidad de los estudiantes
- ☐ Falta de tiempo para buscar y diseñar materiales adecuados
- ☐ Materiales impuestos por la institución que no se ajustan a mis necesidades
- ☐ Otro: _____

Este contenido no ha sido creado ni aprobado por Google.

Google Formularios

Validation Rubric

The purpose of this instrument is to gather expert assessments of a series of reading comprehension materials developed with the support of ChatGPT 4.0. These materials are meant for students whose native language is Spanish and who are at the B1 level (expected for 9 grade) in English as a foreign language. B1 is understood as a curricular goal guiding the teaching process, as well as a level likely achieved by many students during their learning process.

You will evaluate four different materials, each consisting of an adapted text and a set of activities designed for implementation in the classroom. It will take you approximately 50 minutes. The objective of this validation is to determine whether -collectively- these materials, created with the support of artificial intelligence, are suitable and functional for use in real teaching contexts.

Before answering this validation rubric, please familiarize with the four materials:

Material

1: https://drive.google.com/file/d/1xh9e1l9LTbnlEgpJWkLrKF38bW_gCV0Q/view?usp=drive_link

Material 2:

<https://drive.google.com/file/d/1pq5EC2lrYE4WyPBCVV-G8EZPiqfknhuD/view?usp=sharing>

Material 3:

<https://drive.google.com/file/d/1jhRISxmg-3moyCGBodoOlynuPR91pVSR/view?usp=sharing>

Material 4:

<https://drive.google.com/file/d/1-alwdU-GfELj0Pb0o27HceVD2QHMDJtS/view?usp=sharing>

We sincerely appreciate your collaboration in this project.

* Indica que la pregunta es obligatoria

Before starting

According to Brown and Abeywickrama (2019), reading in a second language (L2) is an essential skill in language learning, particularly in educational contexts where it serves as a foundation for teaching, assessments, and the development of linguistic abilities. However, it is not a natural or automatic process. Students do not acquire this skill simply by being exposed to written texts, as it requires specific instruction and deliberate practice. For learners of English, they face two major challenges: learning to decode written language and transforming it into meaningful information. Therefore, reading in an L2 should be guided to enable students to achieve efficient comprehension.

Regarding the B1 competencies considered for this rubric, they can be found in the national documents, *Derechos Básicos de Aprendizaje* (DBA) and *Currículo Sugerido* (2016), as well as in the *Common European Framework of Reference* (CEFR).

- DBA (2016): <https://santillanaplus.com.co/pdf/DBA-ingles-espanol.pdf>

- Currículo Sugerido (2016): https://eco.colombiaaprende.edu.co/wp-content/uploads/2021/09/Esquema-Curricular-Espa_RevMEN.pdf
- CEFR (2020): <https://www.coe.int/en/web/common-european-framework-reference-languages>
- CEFR Descriptors (2020) : <https://rm.coe.int/cefr-descriptors-2020-/16809ed2c7>

1. Nombre *

2. Email *

First Criterion

Please rate each criterion using a scale from 1 (minimum) to 5 (maximum).

If a criterion does not apply to the material being evaluated, mark it as “Not applicable” (N/A).

Adequacy to level B1

The material is intended to be at level B1 in terms of vocabulary, comprehension, and task difficulty, meaning that the materials should be suitable and appropriate for the abilities of a B1 student.

3. Does the material allow students to develop B1-level reading strategies, such as identifying keywords, inferring meanings, and connecting ideas? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

4. Does the material include frequently used vocabulary that a 9-grader (B1 level) student can

recognize and employ? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

5. Do the tasks present a level of difficulty that is manageable for a B1 level student? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

Second Criterion

Variety and quality of the texts

The adapted texts should maintain consistency, fluency, appropriate length, and a structure appropriate to their genre.

6. Do the adapted texts preserve a natural flow and overall coherence? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4

☐ 5

☐ N/A

7. Is the length of the adapted texts suitable for B1 level students? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

8. Does the structure of the adapted texts adhere to the conventions of its type and genre? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

9. Does each paragraph have a logical structure with an introduction, a well-supported development, and a conclusion that facilitates the transition to the next paragraph? (For argumentative texts only) *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

- ☐ 4
- ☐ 5
- ☐ N/A

10. Is the organization of paragraphs adequate, maintaining some balance in length? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

Third Criterion

Variety, quality, and accuracy of the exercises

Activities are expected to be diverse, clear, and relevant, as well as understandable, varied, and consistent with what they are intended to assess.

11. Are the instructions clear and easy for students to understand and follow? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

12. Do the exercises include a variety of task types (multiple choice, true/false, gap-fill, sequencing ideas, matching, etc.)? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

13. Does the pre-reading section of the material promote effective reading strategies such as activating prior knowledge, predicting content, and providing a knowledge framework for constructing meaning? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

14. Does the while-reading section of the material promote effective reading strategies such as corroborating predictions, constructing meaning, identifying key ideas, making inferences, monitoring understanding, etc.? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

15. Does the post-reading section of the material promote effective reading strategies such as language awareness, expansion of ideas, critical reflection, personal response, etc.? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

16. Is there consistency between the expected answers and the information provided in the adapted text (i.e., no factual errors or contradictions)? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

17. Do the activities accurately assess the skills they intend to measure? (e.g., Do cause-effect tasks truly require identifying cause-effect relationships?) *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5

☐ N/A

18. Are the multiple-choice options plausible and well-designed to avoid obvious clues or guessable answers? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

Fourth Criterion

Quality of the reading comprehension levels

The activities should include different levels of reading comprehension (literal, inferential, critical, and creative) and present a logical progression in difficulty.

19. Does the material include a variety of reading comprehension levels (literal, inferential/interpretative, critical, and creative)? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

20. Does the literal level allow students to clearly retrieve explicit information in a progressive manner? *

Marca solo una opción.

☐ 1

- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

21. Do the inferential activities encourage the reading between the lines by drawing conclusions, making predictions, and identifying implicit meanings? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

22. Do the critical comprehension activities help students distinguish between fact and opinion, evaluate arguments, and consider opposing viewpoints? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

23. Do the creative-level activities foster personal reflection, idea generation, and the ability to transfer content to real or hypothetical contexts? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

24. Is there a logical progression in the level of difficulty across the comprehension levels (from literal to applied), supporting the development of higher-level reading skills? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

Fifth Criterion

Coherence with the pedagogical reading objectives

The materials have to be aligned with national standards (DBA, currículo sugerido) and the CEFR, and maintain consistency with the objectives established for level B1.

25. Does the material clearly reflect the expected reading comprehension levels according to the DBA? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4

☐ 5

☐ N/A

26. Does the material clearly reflect the expected reading comprehension levels according to the CEFR levels? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

27. Are the materials aligned with the learning objectives proposed in the Colombian national curriculum for B1 level? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

28. Does the adapted text facilitate the identification of discourse structures (e.g., sequence, cause-effect, contrast, addition) in accordance with national standards? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

Sixth Criterion

Cultural and contextual relevance

The topics, references, and content need to be meaningful and relevant to B1 students and help them connect with their context and experiences.

29. Are the topics relevant (of interest) for B1 students? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

30. Is the content designed to promote student reflection on their own experiences and surroundings? *

Marca solo una opción.

☐ 1

☐ 2

☐ 3

☐ 4

☐ 5

☐ N/A

31. Do the texts contain cultural references relevant for B1 students? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

32. Does the adapted text facilitate the identification of discourse structures (e.g., sequence, cause-effect, contrast, addition) in accordance with national standards? *

Marca solo una opción.

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4
- ☐ 5
- ☐ N/A

Final Thoughts

33. Comments and/or reflections

Este contenido no ha sido creado ni aprobado por Google.

Google Formularios

Preguntas para la entrevista a los participantes

Sobre su experiencia

1. ¿Alguna vez han utilizado la IA para modificar o crear materiales para sus cursos?
2. Desde su experiencia, ¿qué características debe tener un buen material de comprensión lectora para nivel B1?

Sobre la rúbrica y los materiales

1. ¿Aplicarían estos materiales a sus cursos? ¿Qué adaptaciones o cambios les harían?
2. En los ítems sobre adecuación al nivel B1, ¿qué factores tuvo en cuenta para decidir si los materiales realmente permitían desarrollar las estrategias de lectura propias de ese nivel?
3. En la sección de ejercicios, ¿qué tipos de actividades le parecieron más efectivos para evaluar la comprensión lectora y cuáles menos? ¿Por qué?
4. Respecto a la coherencia con los objetivos pedagógicos (DBA, CEFR, currículo nacional), ¿detectó algún desfase?
5. En cuanto a la relevancia cultural y contextual, ¿qué elementos le parecieron pertinentes para estudiantes colombianos y cuáles no tanto?

Sobre la IA

1. ¿Creen que la IA es una herramienta útil para el maestro en la educación?
2. En su opinión, ¿qué ventajas y desventajas vio en el uso de IA para crear materiales didácticos de lectura en comparación con materiales diseñados por docentes?
3. ¿Cómo considera que la IA manejó el nivel de dificultad léxico y gramatical para estudiantes B1?
4. ¿Qué riesgos o problemas identifica en confiar en materiales generados por IA para procesos de enseñanza-aprendizaje?
5. ¿Qué recomendaciones daría para mejorar el uso de IA en la creación de materiales de lectura?