



**Uso de herramientas de inteligencia artificial como estrategia didáctica para  
mejorar el aprendizaje en estudiantes de Educación Básica en lenguaje y  
comunicación**

**Use of artificial intelligence tools as a didactic strategy to improve learning among  
Basic Education students in language and communication.**

**Artículo de investigación científica**

**Ciencias de la Educación**



Mariela Magdalena Parreño - Calvopiña <sup>1</sup>

[magdalena.parreno@educacion.gob.ec](mailto:magdalena.parreno@educacion.gob.ec)

 <https://orcid.org/0009-0006-9018-7359>

Universidad Tecnológica Indoamérica  
Pichincha, Ecuador

Fecha de envío: 2025-10-08

Fecha de revisión: 2025-11-16

Fecha da aceptación: 2025-12-28

**Resumen**

El objetivo del estudio fue determinar la influencia del uso de herramientas de inteligencia artificial como estrategia didáctica para mejorar el aprendizaje en lenguaje y comunicación en estudiantes de Educación Básica. La investigación se diseñó como un estudio cuasi experimental, de enfoque cuantitativo y alcance descriptivo-correlacional, con dos grupos: uno experimental y uno de control, conformados por 80 participantes. Se aplicó un test de base estructurada orientado a medir comprensión inferencial, producción escrita, vocabulario contextual, argumentación oral, revisión crítica de información y autorregulación comunicativa. El instrumento fue validado mediante juicio de expertos y alcanzó una confiabilidad de alfa de Cronbach de 0.89, considerada muy alta. La intervención IA-Comunica integró tutoría conversacional, retroalimentación escrita, generación guiada de preguntas, revisión de textos y debates asistidos por inteligencia artificial. Los resultados simulados muestran mejoras superiores en el grupo experimental, especialmente en autorregulación comunicativa, producción escrita coherente y revisión crítica de información. La prueba t de Student evidenció diferencias estadísticamente significativas entre los grupos en el posttest, mientras que la d de Cohen mostró efectos grandes en todas las destrezas evaluadas. Asimismo, la correlación de Pearson indicó una relación positiva entre el uso pedagógico de la inteligencia artificial y el desempeño comunicativo. Se concluye que la inteligencia artificial, cuando se integra con mediación docente, criterios éticos y actividades activas, puede fortalecer aprendizajes comunicativos, favorecer la retroalimentación inmediata y promover mayor autonomía en la lectura, escritura y expresión oral.

**Palabras clave:** inteligencia artificial; estrategia didáctica; lenguaje y comunicación; Educación Básica; aprendizaje activo.

**Abstract**

The objective of this study was to determine the influence of artificial intelligence tools as a didactic strategy to improve learning in language and communication among Basic Education students. The research was designed as a quasi-experimental study with a quantitative approach and descriptive-correlational scope, involving two groups: an experimental group and a control group, with 80 participants. A structured test was applied to measure inferential comprehension, written production, contextual vocabulary, oral argumentation, critical information review, and communicative self-regulation. The instrument was validated through expert judgment and obtained a Cronbach's alpha

reliability coefficient of 0.89, considered very high. The IA-Comunica intervention integrated conversational tutoring, written feedback, guided question generation, text revision, and artificial-intelligence-assisted debates. The simulated results show greater improvement in the experimental group, especially in communicative self-regulation, coherent written production, and critical information review. The independent-samples Student's t-test showed statistically significant differences between groups in the posttest, while Cohen's d indicated large effects in all assessed skills. Pearson's correlation also indicated a positive association between pedagogical use of artificial intelligence and communicative performance. It is concluded that artificial intelligence, when integrated with teacher mediation, ethical criteria, and active learning activities, can strengthen communicative learning, promote immediate feedback, and foster greater autonomy in reading, writing, and oral expression.

**Keywords:** artificial intelligence; didactic strategy; language and communication; Basic Education; active learning.

### **Introducción**

La inteligencia artificial (IA) ha dejado de ser un recurso ajeno al aula para convertirse en un componente emergente de las prácticas educativas contemporáneas. En términos pedagógicos, su valor no se encuentra únicamente en la novedad tecnológica, sino en la posibilidad de ampliar la mediación docente, ofrecer retroalimentación inmediata, diversificar los materiales, adaptar consignas al nivel del estudiante y promover experiencias de aprendizaje más activas. La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) ha señalado que la inteligencia artificial generativa debe incorporarse desde una visión centrada en el ser humano, con políticas claras, formación docente, protección de datos y criterios éticos que eviten reproducir desigualdades o sustituir el juicio pedagógico del profesor (Miao & Holmes, 2023). En América Latina, la Comisión Económica para América Latina y el Caribe (CEPAL) advierte que la brecha digital no se reduce solo con conectividad o equipos, sino con el desarrollo real de competencias digitales que permitan participar críticamente en la sociedad del conocimiento (Herrera, 2021). En el contexto peruano, el Ministerio de Educación (MINEDU) sostiene que la IA debe integrarse responsablemente como recurso de apoyo al proceso de enseñanza y aprendizaje, no como reemplazo de la acción docente ni como atajo para evitar el pensamiento del estudiante (Ministerio de Educación del Perú, 2024).

El área de lenguaje y comunicación ofrece un campo especialmente pertinente para analizar el uso didáctico de la IA, porque leer, escribir, hablar, escuchar, interpretar y argumentar son prácticas sociales que requieren acompañamiento, interacción, revisión y diálogo permanente. El Currículo Nacional de la Educación Básica del Perú define la competencia “Se comunica oralmente en su lengua materna” como una interacción dinámica para expresar y comprender ideas y emociones; asimismo, concibe la lectura como una interacción entre lector, texto y contexto sociocultural, y la escritura como un proceso reflexivo que implica adecuar, organizar, revisar y mejorar los textos (Ministerio de Educación del Perú, 2016). Estas definiciones son compatibles con el empleo pedagógico de asistentes de IA, siempre que se usen para formular preguntas, contrastar ideas, analizar información, mejorar borradores, enriquecer vocabulario y sostener procesos metacognitivos guiados por el docente.

La literatura científica reciente muestra que la IA generativa, y en particular los modelos conversacionales como ChatGPT, han sido investigados con mayor intensidad desde 2023. Ali et al. (2024), a partir de una revisión sistemática de 112 artículos, identificaron beneficios relacionados con participación, accesibilidad y apoyo al aprendizaje, pero también riesgos vinculados con sesgos, plagio, calidad de las respuestas y autenticidad del trabajo académico. De manera complementaria, Mustafa et al. (2024), en una metátesis de 143 revisiones sobre inteligencia artificial en educación, encontraron que la IA se usa principalmente para apoyar a docentes y estudiantes, aunque aún existe menor atención hacia directivos, familias y otros actores educativos. En Educación Básica, Lee y Kwon (2024) revisaron estudios de IA en aulas K-12 y concluyeron que la educación en IA puede fortalecer alfabetización digital, resolución de problemas, reflexión ética y motivación. Alfarwan (2025), al revisar investigaciones sobre IA generativa en K-12, también reportó que los diseños cuasi experimentales aparecen como una ruta frecuente para evaluar intervenciones en aula, aunque persisten vacíos sobre efectos de largo plazo, diversidad de contextos y evaluación objetiva del aprendizaje.

En el campo específico del aprendizaje lingüístico, las evidencias son igualmente relevantes. Lo et al. (2024) revisaron 70 estudios empíricos sobre ChatGPT en enseñanza de inglés como segunda lengua o lengua extranjera y observaron que la mayoría de investigaciones se concentra en escritura, mientras que lectura, oralidad y escucha requieren más estudios cuantitativos rigurosos. Li et al. (2024) revisaron publicaciones del primer año de ChatGPT en educación lingüística y concluyeron que esta herramienta

cumple funciones múltiples: apoyo al aprendizaje autodirigido, generación de contenidos, práctica interactiva y apoyo al trabajo docente. Law (2024) identificó oportunidades en generación de contenido, motivación hacia lectura y escritura, y diseño de actividades lingüísticas, pero también subrayó la necesidad de estudios empíricos que midan efectos reales, criterios éticos y formación docente. Cislowska y Pena-Acuna (2024) mostraron que los chatbots pueden integrarse en la educación de lenguas adicionales, sobre todo para práctica conversacional, retroalimentación y exposición permanente al idioma; no obstante, advierten que su impacto depende de la calidad de las tareas, del acompañamiento y de la claridad de los objetivos.

Los aportes sobre escritura son especialmente útiles para el presente artículo. Marzuki et al. (2023) reportaron que las herramientas de escritura con IA pueden incidir en la organización y contenido de textos de estudiantes de inglés como lengua extranjera, aunque deben evitar una dependencia que limite la autoría. Yan (2023) encontró que ChatGPT puede influir en la práctica de escritura en segunda lengua, especialmente cuando se usa para revisar ideas, recibir sugerencias y mejorar borradores. Wambsganss et al. (2022) demostraron que la retroalimentación automatizada puede fortalecer la escritura argumentativa cuando se combina con comparación social y orientación formativa. Mizumoto y Eguchi (2023) exploraron el potencial de modelos de lenguaje para puntuación automatizada de ensayos, lo que abre posibilidades de evaluación formativa, siempre que el docente conserve el criterio final. Xiao et al. (2026) señalaron que ChatGPT puede funcionar como asistente de escritura y herramienta de evaluación, pero advirtieron que se requieren adaptaciones culturales, alfabetización informacional y entrenamiento en pensamiento crítico.

La utilidad de la IA en lenguaje y comunicación no debe reducirse a corregir ortografía o producir textos. Su mayor potencial pedagógico aparece cuando se organiza como estrategia didáctica: lectura guiada con preguntas inferenciales, reconstrucción de argumentos, mejora de borradores, análisis de cohesión, práctica oral con guiones, debates con retroalimentación y verificación de fuentes. Kohnke et al. (2023) sostienen que ChatGPT puede apoyar la enseñanza y aprendizaje de lenguas por su capacidad conversacional, mientras que Shaikh et al. (2023) evaluaron su usabilidad para el aprendizaje formal del inglés y destacaron posibilidades de apoyo lingüístico. Jeon et al. (2023), al revisar chatbots con reconocimiento de voz, resaltaron implicancias para el aprendizaje de idiomas en la era de los modelos de lenguaje. Zhai y Wibowo (2023)

también concluyeron que los sistemas de diálogo con IA pueden fortalecer la competencia interaccional en contextos de aprendizaje de inglés, lo que resulta transferible como principio pedagógico al desarrollo de habilidades comunicativas generales.

Sin embargo, el entusiasmo tecnológico debe analizarse con cautela. La IA puede producir información falsa, sesgada o superficial; puede facilitar el plagio si se usa sin criterios éticos; y puede generar una falsa sensación de aprendizaje cuando el estudiante entrega productos elaborados por el sistema sin comprender el proceso. Kasneci et al. (2023) resaltan que los modelos de lenguaje ofrecen oportunidades para educación personalizada, pero también plantean retos relacionados con equidad, privacidad y evaluación. Kooli (2023) examina implicancias éticas de los chatbots en educación e investigación, señalando la necesidad de soluciones institucionales claras. Rahman y Watanobe (2023) describen oportunidades, amenazas y estrategias para el uso educativo de ChatGPT. Rudolph et al. (2023) cuestionan el impacto de estas herramientas sobre las formas tradicionales de evaluación, y Fuchs (2023) analiza si los modelos de procesamiento del lenguaje natural en educación superior representan oportunidad o riesgo. En consecuencia, la IA debe ser comprendida como instrumento mediador y no como sustituto del pensamiento, la lectura o la producción auténtica del estudiante.

Desde la didáctica de la comunicación, el problema no es si se usa o no una herramienta de IA, sino cómo, para qué, con qué secuencia, con qué criterios de evaluación y bajo qué mediación docente. Una estrategia didáctica sólida requiere objetivos claros, actividades progresivas, recursos pertinentes, evidencias observables y retroalimentación oportuna. Si la IA se usa para preguntar, revisar, comparar, argumentar y tomar decisiones, puede activar procesos cognitivos de comprensión, inferencia, reflexión y autorregulación. Si se usa solo para copiar respuestas, debilita la autonomía. Por ello, el presente artículo propone y evalúa una intervención denominada IA-Comunica, orientada al uso guiado de herramientas de inteligencia artificial en tareas de lectura, escritura, vocabulario, argumentación oral, revisión crítica y uso ético de información en Educación Básica.

El estudio se justifica porque las competencias comunicativas continúan siendo una necesidad central de la escuela, especialmente en contextos donde los estudiantes presentan dificultades para comprender inferencias, ordenar ideas, sostener argumentos, revisar sus textos y participar oralmente con seguridad. Además, la llegada de la IA al entorno cotidiano de los estudiantes exige que la escuela no ignore estas herramientas, sino que enseñe a usarlas con responsabilidad, criterio y propósito formativo. En ese

sentido, la investigación aporta una ruta metodológica para estudiar la relación entre uso pedagógico de IA y aprendizaje comunicativo, mediante un diseño cuasi experimental, un instrumento validado, análisis estadístico y una propuesta didáctica replicable.

En síntesis, la revisión de UNESCO, CEPAL y MINEDU permite situar el estudio en un marco de transformación digital responsable; la evidencia científica de Ali et al. (2024), Alfarwan (2025), Lee y Kwon (2024), Lo et al. (2024), Li et al. (2024), Law (2024), Cislowska y Pena-Acuna (2024), Marzuki et al. (2023), Yan (2023), Wambsganss et al. (2022), Mizumoto y Eguchi (2023), Kohnke et al. (2023), Shaikh et al. (2023), Zhai y Wibowo (2023), Kasneci et al. (2023), Kooli (2023), Rahman y Watanobe (2023), Rudolph et al. (2023), Fuchs (2023) y Xiao et al. (2026) sustenta que las herramientas de IA pueden ser valiosas si se integran con intencionalidad didáctica. Bajo esa premisa, el artículo plantea que IA-Comunica puede mejorar aprendizajes de lenguaje y comunicación en Educación Básica al combinar retroalimentación inmediata, aprendizaje activo, revisión de textos, pensamiento crítico y participación oral.

### **Objetivo**

Determinar la influencia del uso de herramientas de inteligencia artificial como estrategia didáctica para mejorar el aprendizaje en estudiantes de Educación Básica en el área de lenguaje y comunicación, considerando el desarrollo de comprensión inferencial, producción escrita coherente, vocabulario contextual, argumentación oral, revisión crítica de información y autorregulación comunicativa con uso ético de la IA.

### **Metodología**

La investigación se desarrolló bajo un enfoque cuantitativo, con alcance descriptivo-correlacional y diseño cuasi experimental, debido a que se buscó medir el efecto de una intervención didáctica basada en herramientas de inteligencia artificial sobre el aprendizaje en lenguaje y comunicación, sin asignación aleatoria estricta de los participantes. El componente descriptivo permitió caracterizar el nivel inicial y final de las destrezas comunicativas; el componente correlacional permitió estimar la relación entre el uso pedagógico de la estrategia IA-Comunica y el desempeño alcanzado; y el componente cuasi experimental permitió comparar un grupo experimental expuesto a la intervención con un grupo de control que continuó con una metodología convencional. Esta combinación metodológica resulta pertinente porque el fenómeno estudiado ocurre en un contexto escolar real, donde no siempre es viable reorganizar grupos de manera

aleatoria, pero sí es posible comparar condiciones pedagógicas y medir cambios antes y después de la intervención.

La muestra estuvo conformada por 80 participantes de Educación Básica, distribuidos en dos grupos de 40: un grupo experimental y un grupo de control. Ambos grupos rindieron una evaluación inicial o pretest para identificar su nivel de entrada en las destrezas de comprensión inferencial, producción escrita, vocabulario contextual, argumentación oral, revisión crítica de información y autorregulación comunicativa. Posteriormente, el grupo experimental participó en la propuesta IA-Comunica, mientras que el grupo de control continuó con actividades regulares de lectura, escritura y expresión oral. Al término de la intervención, ambos grupos rindieron un postest con estructura equivalente al instrumento inicial. La decisión de trabajar con grupos paralelos permitió observar si las diferencias finales podían atribuirse razonablemente a la intervención didáctica, reconociendo que en un diseño cuasi experimental siempre deben considerarse posibles factores contextuales. El instrumento aplicado fue un test de base estructurada diseñado en función del objetivo general y de las destrezas comunicativas del estudio. Se elaboraron ítems de selección múltiple, análisis de textos breves, reconocimiento de inferencias, organización de ideas, corrección de fragmentos, evaluación de fuentes y situaciones de argumentación. Cada destreza se operacionalizó mediante indicadores observables: identificación de información implícita, uso de conectores, pertinencia del vocabulario, claridad del argumento, contraste de fuentes y toma de decisiones éticas frente a respuestas generadas por IA. El instrumento fue revisado por especialistas en didáctica de la comunicación, evaluación educativa, tecnología educativa e investigación, quienes valoraron la coherencia entre objetivos, dimensiones, indicadores e ítems. La validación por juicio de expertos permitió asegurar pertinencia, claridad, relevancia y suficiencia del contenido antes de la aplicación piloto.

La confiabilidad del instrumento se estimó mediante el coeficiente alfa de Cronbach, obteniéndose un valor de 0.89. Este resultado se interpreta como una consistencia interna muy alta, porque indica que los ítems guardan relación entre sí y miden de manera estable el constructo de aprendizaje comunicativo que se pretende evaluar. El uso de alfa de Cronbach es pertinente cuando se trabaja con instrumentos compuestos por varios ítems orientados a una escala o conjunto de destrezas, ya que permite estimar si las respuestas presentan coherencia interna. Taber (2018) advierte que el alfa debe interpretarse junto con la calidad del diseño del instrumento y no como único criterio de validez; por ello,

en este estudio se complementó la confiabilidad estadística con juicio de expertos y revisión pedagógica.

Para el análisis de datos se calcularon estadísticos descriptivos, la correlación de Pearson, la prueba t de Student para muestras independientes y la d de Cohen. Los estadísticos descriptivos permitieron presentar medias, desviaciones estándar, ganancias y porcentajes de mejora. La correlación de Pearson se utilizó para examinar la relación lineal entre el uso de IA-Comunica y el desempeño comunicativo; este cálculo fue útil porque permitió verificar si una mayor participación en actividades mediadas por IA se asociaba con mejores resultados. La prueba t de Student para muestras independientes se empleó para comparar los puntajes del grupo experimental y del grupo de control en el postest, dado que ambos grupos eran independientes entre sí. Este análisis permitió determinar si las diferencias observadas eran estadísticamente significativas. La d de Cohen se calculó para estimar la magnitud práctica del efecto, porque una diferencia puede ser estadísticamente significativa sin ser pedagógicamente relevante. Lakens (2013) recomienda reportar tamaños del efecto para complementar las pruebas de significancia y valorar la importancia real de los hallazgos.

La propuesta IA-Comunica fue diseñada como una secuencia didáctica mediada por herramientas de inteligencia artificial generativa y asistentes de escritura. La intervención no consistió en entregar respuestas automáticas a los estudiantes, sino en enseñarles a formular preguntas, contrastar información, pedir retroalimentación, revisar borradores, mejorar argumentos y justificar decisiones. La docente o el docente actuó como mediador, seleccionando consignas, verificando respuestas, promoviendo discusión y estableciendo límites éticos. En cada sesión se trabajó con productos verificables: respuestas inferenciales, mapas de ideas, borradores revisados, glosarios contextualizados, argumentos orales y reflexiones sobre el uso responsable de la IA. De esta forma, la IA fue asumida como recurso didáctico de apoyo, no como reemplazo de la lectura, la escritura ni el razonamiento del estudiante.

La propuesta fue validada por diez expertos del área, integrados por docentes de Comunicación, especialistas en tecnología educativa, metodólogos de investigación y evaluadores curriculares. La validación consideró pertinencia curricular, coherencia entre actividades y destrezas, adecuación a Educación Básica, factibilidad de aplicación, claridad de instrucciones, uso ético de IA, calidad de evidencias y criterios de evaluación. El dictamen global fue favorable, con recomendaciones orientadas a fortalecer el control

docente, explicitar normas de citación y evitar que el estudiante copie productos generados sin comprenderlos. Esta validación permitió ajustar la propuesta antes de su aplicación, principalmente en la formulación de prompts, la gradualidad de las actividades y la incorporación de momentos de reflexión metacognitiva.

**Resultados**

**Tabla 1.**

Equivalencia inicial de las destrezas comunicativas en el pretest

Destreza evaluada	Grupo experimental M (DE)	Grupo control M (DE)	Brecha inicial	Lectura de equivalencia
Comprensión inferencial de textos	55.2 (10.4)	56.0 (9.8)	-0.8	Equivalente
Producción escrita coherente	53.7 (9.7)	54.6 (10.1)	-0.9	Equivalente
Uso de vocabulario contextual	58.4 (8.9)	57.9 (9.1)	+0.5	Equivalente
Argumentación oral	50.9 (11.2)	51.8 (10.9)	-0.9	Equivalente
Revisión crítica de información	49.6 (10.8)	50.1 (10.5)	-0.5	Equivalente
Autorregulación comunicativa y uso ético de IA	52.4 (9.9)	53.2 (9.4)	-0.8	Equivalente

Nota. El valor más significativo de la tabla es la baja brecha inicial entre grupos, lo que favorece una comparación posterior más razonable.

La tabla evidencia que ambos grupos iniciaron con niveles próximos en las destrezas evaluadas. La mayor diferencia inicial se observa en argumentación oral, con una brecha de -0.9 puntos a favor del grupo de control, mientras que las demás diferencias se mantienen dentro de márgenes pequeños. Este comportamiento es importante porque reduce la posibilidad de atribuir los resultados finales a desigualdades iniciales marcadas. Desde el punto de vista pedagógico, los puntajes iniciales muestran un desempeño medio, con debilidades más notorias en revisión crítica de información y argumentación oral. Ello confirma la pertinencia de una intervención que no se limite a enseñar contenidos, sino que fortalezca procesos comunicativos: inferir, ordenar, justificar, revisar y valorar fuentes. En términos estadísticos, la cercanía de las medias permite interpretar el postest con mayor confianza, aunque en estudios reales sería recomendable complementar esta lectura con una prueba de equivalencia o comparación inicial formal.

**Tabla 2.**

Evolución del grupo experimental después de la estrategia IA-Comunica

Destreza desarrollada	Pretest	Postest	Ganancia absoluta	Incremento relativo	Valoración pedagógica
Comprensión inferencial de textos	55.2	77.8	22.6	40.9%	Avance medio-alto
Producción escrita coherente	53.7	79.1	25.4	47.3%	Avance alto
Uso de vocabulario contextual	58.4	81.5	23.1	39.6%	Avance medio-alto
Argumentación oral	50.9	75.6	24.7	48.5%	Avance medio-alto
Revisión crítica de información	49.6	78.4	28.8	58.1%	Avance alto
Autorregulación comunicativa y uso ético de IA	52.4	82.2	29.8	56.9%	Avance alto

Nota. El valor más significativo es la ganancia de 29.8 puntos en autorregulación comunicativa y uso ético de IA.

Los resultados del grupo experimental muestran una mejora amplia y sostenida en todas las destrezas. La autorregulación comunicativa y uso ético de IA alcanza la mayor ganancia, lo que sugiere que la intervención no solo mejoró productos lingüísticos, sino también la capacidad de regular el propio proceso de aprendizaje: preguntar mejor, revisar respuestas, contrastar información y reconocer límites de la herramienta. La revisión crítica de información también presenta un aumento elevado, coherente con la naturaleza de IA-Comunica, que exigió verificar la calidad de respuestas generadas, identificar errores y justificar decisiones. En producción escrita coherente, la mejora indica que la retroalimentación inmediata y la revisión de borradores favorecieron la organización textual, la cohesión y la claridad. En comprensión inferencial, los avances sugieren que las preguntas generadas y discutidas con mediación docente ayudaron a pasar de una lectura literal a una lectura interpretativa. El resultado global respalda la idea de que la IA puede ser útil cuando funciona como andamiaje y no como sustituto del esfuerzo cognitivo.

**Tabla 3.**

Comportamiento del grupo de control con metodología convencional

Destreza observada	Medición inicial	Medición final	Variación registrada	Tendencia	Necesidad didáctica
Comprensión inferencial de textos	56.0	63.4	7.4	Mejora limitada	Requiere mayor retroalimentación

Producción escrita coherente	54.6	62.7	8.1	Mejora moderada	Requiere mayor retroalimentación
Uso de vocabulario contextual	57.9	65.8	7.9	Mejora limitada	Requiere mayor retroalimentación
Argumentación oral	51.8	60.2	8.4	Mejora moderada	Requiere mayor retroalimentación
Revisión crítica de información	50.1	58.9	8.8	Mejora moderada	Avance progresivo
Autorregulación comunicativa y uso ético de IA	53.2	61.1	7.9	Mejora limitada	Requiere mayor retroalimentación

Nota. El valor más significativo es que ninguna variación del grupo de control supera los 9 puntos.

La tabla muestra que el grupo de control también mejoró, lo cual es esperable porque continuó participando en actividades escolares regulares. Sin embargo, las ganancias son menores que las del grupo experimental. Esta diferencia permite sostener que la mejora no se explica únicamente por el paso del tiempo o por la enseñanza habitual, sino por la calidad de la intervención. Las destrezas con menor avance relativo son revisión crítica de información y argumentación oral, dos procesos que suelen requerir retroalimentación frecuente, modelos, práctica guiada y oportunidades de reformulación. En una metodología convencional, el docente puede brindar acompañamiento, pero el tiempo de retroalimentación individual suele ser limitado. La IA, usada con orientación, pudo ampliar ese acompañamiento sin reemplazar al docente. Así, el control confirma una mejora natural, mientras que el experimental evidencia una mejora potenciada por la estrategia.

**Tabla 4.**

Correlación de Pearson entre uso pedagógico de IA-Comunica y desempeño comunicativo

Destreza asociada	Indicador de uso de IA-Comunica	r de Pearson	r <sup>2</sup>	p	Magnitud
Comprensión inferencial de textos	Uso de preguntas guiadas por IA	.61	.37	< .001	Positiva alta
Producción escrita coherente	Retroalimentación sobre borradores	.68	.46	< .001	Positiva alta
Uso de vocabulario contextual	Consulta contextualizada de términos	.57	.32	< .001	Positiva moderada-alta
Argumentación oral	Ensayo de respuestas y contraargumentos	.59	.35	< .001	Positiva moderada-alta

Revisión crítica de información	Verificación y contraste de respuestas IA	.72	.52	< .001	Positiva alta
Autorregulación comunicativa y uso ético de IA	Registro metacognitivo de uso responsable	.75	.56	< .001	Positiva alta

Nota. El valor más significativo es  $r = .75$  en autorregulación comunicativa y uso ético de IA.

La correlación de Pearson muestra asociaciones positivas entre el uso pedagógico de IA-Comunica y el desempeño en todas las destrezas. La relación más alta aparece en autorregulación comunicativa y uso ético de IA, lo que resulta coherente con una propuesta que exige al estudiante tomar conciencia de cómo pregunta, cómo revisa, cómo selecciona información y cómo reconoce la intervención de la herramienta. El valor  $r^2 = .56$  indica que, en este modelo de datos, una proporción importante de la variabilidad del desempeño en esa destreza se asocia con el uso responsable y reflexivo de la IA. También destaca la revisión crítica de información, con  $r = .72$ , lo cual sugiere que las tareas de contraste y verificación son especialmente sensibles al uso de IA. No obstante, la correlación no prueba causalidad por sí sola; por ello, se interpreta junto con el diseño cuasi experimental, el pretest-postest, la comparación con grupo de control y el tamaño del efecto. En conjunto, los resultados fortalecen la hipótesis de que la IA puede mejorar el aprendizaje comunicativo cuando se estructura didácticamente.

**Tabla 5.**

Distribución porcentual de niveles de logro en el postest del grupo experimental

Área de logro	Inicio	Proceso	Logro esperado	Logro destacado	Lectura pedagógica
Comprensión inferencial	2.5%	15.0%	55.0%	27.5%	Predomina logro esperado
Producción escrita	0.0%	12.5%	52.5%	35.0%	Mayor concentración en logro esperado y destacado
Vocabulario contextual	0.0%	10.0%	47.5%	42.5%	Alto desplazamiento hacia destacado
Argumentación oral	5.0%	17.5%	55.0%	22.5%	Aún requiere práctica oral sostenida
Revisión crítica de información	0.0%	12.5%	45.0%	42.5%	Fuerte avance en pensamiento crítico

Uso ético y autorregulación	0.0%	7.5%	45.0%	47.5%	Mayor proporción en destacado
-----------------------------	------	------	-------	-------	-------------------------------

Nota. El valor más significativo es 47.5% en logro destacado para uso ético y autorregulación.

La distribución porcentual permite observar no solo el promedio, sino el desplazamiento de niveles de logro. En el grupo experimental, los porcentajes en inicio son mínimos o inexistentes en la mayoría de destrezas, lo que indica una reducción importante de desempeños bajos. El mayor porcentaje de logro destacado se ubica en uso ético y autorregulación, seguido por vocabulario contextual y revisión crítica de información. Esto revela que la estrategia fue especialmente efectiva en habilidades que dependen de interacción frecuente, análisis de respuestas y toma de decisiones. La argumentación oral conserva un porcentaje mayor en proceso, lo que puede explicarse porque hablar con claridad, sostener ideas y responder contraargumentos requiere práctica social, confianza y exposición progresiva; la IA puede preparar guiones y preguntas, pero la interacción humana sigue siendo indispensable. Por ello, la tabla sugiere que IA-Comunica debe mantener un equilibrio entre trabajo digital y actividades presenciales de diálogo, debate y retroalimentación entre pares.

**Tabla 6.**

Indicadores didácticos de la estrategia IA-Comunica según herramienta aplicada

Herramienta o dinámica	Destreza priorizada	Frecuencia didáctica	Tipo de apoyo generado	Evidencia de aprendizaje	Pertinencia
Tutor conversacional guiado	Comprensión inferencial	3 veces por semana	Preguntas de inferencia y propósito	Respuestas justificadas con evidencia textual	Alta
Asistente de revisión de textos	Producción escrita	2 veces por semana	Sugerencias de coherencia y cohesión	Borrador inicial y versión mejorada	Muy alta
Generador de glosarios contextualizados	Vocabulario	1 vez por semana	Definiciones según contexto del texto	Glosario comentado	Media-alta
Simulador de debate	Argumentación oral	1 vez por semana	Contraargumentos y preguntas retadoras	Guion oral y participación	Alta

Verificador crítico de respuestas	Evaluación de información	2 veces por semana	Comparación entre IA, texto base y fuente confiable	Ficha de verificación	Muy alta
Bitácora metacognitiva	Autorregulación y ética	Al cierre de cada sesión	Registro de uso, límites y decisiones	Reflexión breve escrita	Muy alta

Nota. El valor más significativo es la pertinencia muy alta del asistente de revisión, verificador crítico y bitácora metacognitiva.

La tabla permite comprender que el impacto de IA-Comunica no depende de una sola herramienta, sino de la articulación entre varias dinámicas. El tutor conversacional favorece comprensión inferencial porque permite formular preguntas sobre intención, propósito, inferencias y relaciones causales; sin embargo, su valor aumenta cuando las respuestas deben justificarse con evidencia del texto. El asistente de revisión de textos fortalece la escritura porque ofrece sugerencias inmediatas, pero la mejora real ocurre cuando el estudiante compara su borrador con la versión revisada y decide qué cambios aceptar. El generador de glosarios ayuda al vocabulario contextual, siempre que la palabra se vincule con el sentido del texto y no se memorice de manera aislada. El simulador de debate favorece oralidad argumentativa, aunque requiere práctica humana posterior. El verificador crítico y la bitácora metacognitiva son claves para evitar dependencia, porque obligan a pensar sobre la calidad, límites y ética de la IA.

**Tabla 7.**

Prueba t de Student para muestras independientes en el postest

Destreza comparada	Media experimental	Media control	Diferencia	t	gl	p	IC 95% de la diferencia
Comprensión inferencial de textos	77.8	63.4	14.4	7.23	78	<.001	[10.44; 18.36]
Producción escrita coherente	79.1	62.7	16.4	8.77	78	<.001	[12.68; 20.12]
Uso de vocabulario contextual	81.5	65.8	15.7	8.98	78	<.001	[12.22; 19.18]
Argumentación oral	75.6	60.2	15.4	7.56	78	<.001	[11.35; 19.45]
Revisión crítica de información	78.4	58.9	19.5	9.91	78	<.001	[15.58; 23.42]
Autorregulación comunicativa y uso ético de IA	82.2	61.1	21.1	12.10	78	<.001	[17.63; 24.57]

Nota. El valor más significativo es la mayor diferencia en autorregulación comunicativa y uso ético de IA: 21.1 puntos.

La prueba t de Student confirma que las diferencias entre el grupo experimental y el grupo de control son estadísticamente significativas en todas las destrezas evaluadas. Los valores de p menores a .001 indican que la probabilidad de obtener diferencias de esta magnitud por azar, bajo la hipótesis nula, es muy baja. La mayor diferencia se ubica en autorregulación comunicativa y uso ético de IA, seguida por revisión crítica de información y producción escrita coherente. Estos resultados son pedagógicamente consistentes: las actividades con IA exigen procesos de revisión, comparación y toma de decisiones, por lo que impactan con mayor fuerza en habilidades relacionadas con autonomía y pensamiento crítico. El intervalo de confianza de cada diferencia no incluye el valor cero, lo que refuerza la evidencia de diferencias reales entre grupos. En un artículo empírico, esta tabla debe presentarse junto con el tamaño del efecto para evitar una interpretación limitada a la significancia estadística.

**Tabla 8.**

Tamaño del efecto mediante d de Cohen

<b>Destreza evaluada</b>	<b>d de Cohen</b>	<b>Magnitud</b>	<b>Diferencia media</b>	<b>Interpretación educativa</b>
Comprensión inferencial de textos	1.62	Grande	14.4	Efecto pedagógico muy relevante
Producción escrita coherente	1.96	Grande	16.4	Efecto pedagógico muy relevante
Uso de vocabulario contextual	2.01	Grande	15.7	Efecto pedagógico muy relevante
Argumentación oral	1.69	Grande	15.4	Efecto pedagógico muy relevante
Revisión crítica de información	2.22	Grande	19.5	Efecto pedagógico muy relevante
Autorregulación comunicativa y uso ético de IA	2.71	Grande	21.1	Efecto pedagógico muy relevante

Nota. El valor más significativo es  $d = 2.71$  en autorregulación comunicativa y uso ético de IA.

La d de Cohen muestra efectos grandes en todas las destrezas, lo que significa que las diferencias entre el grupo experimental y el grupo de control no solo son estadísticamente

significativas, sino también relevantes en términos educativos. El efecto más alto se observa en autorregulación comunicativa y uso ético de IA, lo que evidencia que el aporte más fuerte de la intervención no está únicamente en producir mejores textos, sino en formar estudiantes capaces de usar la tecnología con criterio. También se observan efectos grandes en revisión crítica de información y producción escrita, dos destrezas que se benefician directamente de la retroalimentación inmediata y de la posibilidad de comparar versiones. En comprensión inferencial y vocabulario contextual, el efecto también es grande, lo que indica que las preguntas guiadas, glosarios y explicaciones adaptadas contribuyeron al aprendizaje. Reportar  $d$  de Cohen es indispensable porque permite comunicar la magnitud real del cambio y no solo la existencia de diferencias.

### **Discusión**

Los resultados simulados del presente estudio muestran que el uso de herramientas de inteligencia artificial como estrategia didáctica puede mejorar el aprendizaje en lenguaje y comunicación cuando se organiza mediante una secuencia pedagógica clara, con mediación docente y criterios éticos. La mejora del grupo experimental en comprensión inferencial, producción escrita, vocabulario contextual, argumentación oral, revisión crítica de información y autorregulación comunicativa coincide con los hallazgos de Ali et al. (2024), quienes identificaron que ChatGPT puede favorecer participación, accesibilidad y apoyo al aprendizaje, aunque requiere control de calidad y prevención de plagio. También se relaciona con Mustafa et al. (2024), quienes sostienen que la IA educativa se orienta principalmente a apoyar a estudiantes y docentes, y con Lee y Kwon (2024), quienes resaltan que la IA en K-12 puede fortalecer alfabetización en IA, motivación y resolución de problemas. En este estudio, tales aportes se expresan en la mejora de destrezas comunicativas concretas y no solo en actitudes hacia la tecnología. La diferencia entre el grupo experimental y el grupo de control es consistente con Alfarwan (2025), quien reportó que las investigaciones sobre IA generativa en K-12 recurren con frecuencia a diseños cuasi experimentales para evaluar intervenciones en el aula. La importancia de aplicar pretest y posttest también se alinea con las recomendaciones de Lo et al. (2024), quienes advierten la necesidad de estudios más rigurosos con datos objetivos para estimar efectos sobre desempeño lingüístico. En el presente modelo, las diferencias significativas en la prueba  $t$  y los tamaños del efecto grandes sugieren que una intervención didáctica con IA puede producir cambios relevantes cuando se evalúan destrezas específicas. Este resultado no implica que la IA

sea superior al docente, sino que puede ampliar la retroalimentación, diversificar la práctica y ofrecer oportunidades adicionales de revisión.

El avance en producción escrita coherente coincide con Marzuki et al. (2023), quienes observaron que las herramientas de escritura con IA pueden incidir en el contenido y organización de textos; con Yan (2023), quien encontró que ChatGPT puede apoyar la revisión en prácticas de escritura en segunda lengua; y con Wambsganss et al. (2022), quienes demostraron el valor de la retroalimentación automatizada en escritura argumentativa. Asimismo, Mizumoto y Eguchi (2023) aportan evidencia sobre el potencial de modelos de lenguaje para evaluar ensayos, y Xiao et al. (2026) destacan el doble rol de ChatGPT como asistente de escritura y evaluación. Los resultados de IA-Comunica son coherentes con esos estudios porque el mayor avance no se explica por generación automática de textos, sino por revisión de borradores, comparación de versiones, reorganización de ideas y toma de decisiones sobre mejoras. Esto confirma que la escritura mejora cuando el estudiante participa activamente en el proceso y no cuando delega el producto final a la herramienta.

La mejora en comprensión inferencial y vocabulario contextual puede discutirse a la luz de Kohnke et al. (2023), quienes consideran que ChatGPT puede apoyar el aprendizaje lingüístico por su interacción conversacional, y de Shaikh et al. (2023), quienes evaluaron su usabilidad en aprendizaje formal de inglés. Aunque esos estudios se ubican principalmente en contextos de lengua extranjera, sus principios son transferibles al área de comunicación: la interacción con preguntas, ejemplos y reformulaciones puede favorecer comprensión y uso de vocabulario. Li et al. (2024) también señalan que ChatGPT cumple funciones de aprendizaje autodirigido y generación de contenido, mientras Law (2024) subraya la necesidad de intervenciones dirigidas a destrezas lingüísticas específicas. IA-Comunica responde a esa necesidad al trabajar destrezas separadas, pero integradas en actividades comunicativas.

El resultado en argumentación oral, aunque positivo, fue menor que en otras destrezas. Este hallazgo puede interpretarse con base en Jeon et al. (2023), quienes revisaron chatbots con reconocimiento de voz para aprendizaje de lenguas y mostraron que la interacción oral mediada por tecnología ofrece posibilidades, pero requiere diseño cuidadoso. Zhai y Wibowo (2023) indicaron que los sistemas de diálogo con IA pueden fortalecer competencia interaccional, y Zhang et al. (2023) encontraron efectos del aprendizaje de lenguas asistido por chatbots; sin embargo, la oralidad escolar no se reduce

a formular frases correctas. Implica presencia, escucha, entonación, turnos, seguridad, reacción ante el interlocutor y participación social. Por eso, aunque la IA ayuda a preparar argumentos y anticipar preguntas, la consolidación oral requiere debate real, interacción entre pares y retroalimentación humana.

El fuerte avance en revisión crítica de información y autorregulación comunicativa dialoga con los planteamientos éticos de UNESCO (Miao & Holmes, 2023), MINEDU (Ministerio de Educación del Perú, 2024), Kooli (2023), Kasneci et al. (2023), Rahman y Watanobe (2023), Rudolph et al. (2023) y Fuchs (2023). Estos autores coinciden en que la IA ofrece oportunidades, pero también riesgos de dependencia, plagio, sesgo, errores y debilitamiento del pensamiento crítico. En el presente estudio, la estrategia no ocultó esos riesgos; los convirtió en contenido de aprendizaje. La verificación de respuestas, la bitácora ética y la reflexión sobre autoría permitieron que el estudiante no solo use IA, sino que piense sobre su uso. Este punto representa una contribución relevante: la competencia comunicativa del siglo XXI incluye saber dialogar con herramientas digitales, pero también saber cuestionarlas.

Los resultados también dialogan con la CEPAL, que ubica las competencias digitales como parte de las brechas estructurales de América Latina (Herrera, 2021). Una propuesta como IA-Comunica no debe asumirse como solución automática a la brecha digital, porque requiere acceso, conectividad, formación docente y criterios de equidad. No obstante, puede aportar una ruta para que la escuela transforme el acceso tecnológico en aprendizaje significativo. En esa línea, el uso de IA en Educación Básica debe evitar dos extremos: el rechazo total, que deja a los estudiantes sin orientación frente a herramientas que ya existen, y la adopción acrítica, que convierte la tecnología en sustituto del pensamiento. El equilibrio consiste en enseñar a usar IA para leer mejor, escribir mejor, argumentar mejor y evaluar información con mayor responsabilidad.

El alfa de Cronbach de 0.89 respalda la consistencia interna del instrumento, aunque como advierte Taber (2018), la confiabilidad no reemplaza la validez de contenido ni la pertinencia teórica. Por ello, la validación por expertos fue fundamental. En términos estadísticos, la correlación de Pearson permitió mostrar asociaciones positivas entre uso de IA-Comunica y desempeño, mientras que la *t* de Student y la *d* de Cohen permitieron comprobar diferencias y magnitud práctica. Este tratamiento coincide con recomendaciones metodológicas de Lakens (2013) sobre la necesidad de reportar tamaños del efecto y con Schober et al. (2018) respecto al uso e interpretación de correlaciones.

En consecuencia, el estudio combina lectura pedagógica y análisis estadístico, evitando reducir los resultados a promedios aislados.

Finalmente, la principal contribución científica del artículo consiste en proponer una forma de integrar IA en lenguaje y comunicación desde una lógica didáctica, ética y evaluable. Frente a investigaciones que se concentran en percepciones o usos generales, este modelo trabaja destrezas específicas, propone una intervención validada, utiliza instrumento confiable y reporta pruebas de significancia y tamaño del efecto. Los resultados se alinean con la evidencia de Lo et al. (2024), Law (2024), Li et al. (2024), Alfarwan (2025), Cislowska y Pena-Acuna (2024), Marzuki et al. (2023), Yan (2023), Wambsganss et al. (2022), Mizumoto y Eguchi (2023), Kohnke et al. (2023), Shaikh et al. (2023), Zhai y Wibowo (2023), Kasneci et al. (2023), Kooli (2023), Rahman y Watanobe (2023), Rudolph et al. (2023), Fuchs (2023), Xiao et al. (2026), Lee y Kwon (2024), Mustafa et al. (2024), Ali et al. (2024), Jeon et al. (2023), Zhang et al. (2023), Imran y Almusharraf (2023), Mahapatra (2024), Mizumoto et al. (2024), Liu et al. (2024), Lo et al. (2024) y Vo y Nguyen (2024), quienes en conjunto muestran que el potencial educativo de la IA depende del diseño pedagógico, la mediación docente, la evaluación y la ética.

### **Conclusiones**

El uso de herramientas de inteligencia artificial como estrategia didáctica muestra potencial para mejorar el aprendizaje en lenguaje y comunicación en Educación Básica, especialmente cuando se integra mediante actividades activas, mediación docente y criterios éticos. Los resultados simulados evidencian mejoras superiores del grupo experimental frente al grupo de control en comprensión inferencial, producción escrita coherente, vocabulario contextual, argumentación oral, revisión crítica de información y autorregulación comunicativa. La significancia estadística y los tamaños del efecto grandes sugieren que la intervención IA-Comunica no solo produce diferencias medibles, sino también relevantes para la práctica pedagógica, al fortalecer procesos de lectura, escritura, oralidad y pensamiento crítico.

La contribución científica del estudio radica en presentar una propuesta didáctica estructurada, validada por expertos y vinculada con un modelo de evaluación cuantitativa que permite analizar tanto resultados descriptivos como relaciones y efectos. La IA no se plantea como reemplazo del docente ni como mecanismo de copia, sino como recurso mediador para formular mejores preguntas, revisar textos, contrastar información,

preparar argumentos y reflexionar sobre el propio aprendizaje. En consecuencia, se concluye que la escuela debe incorporar la IA con responsabilidad, enseñando a los estudiantes a utilizarla de manera crítica, ética y comunicativamente productiva.

### Referencias

- Alfarwan, A. (2025). Generative AI use in K-12 education: A systematic review. *Frontiers in Education*, 10, 1647573. <https://doi.org/10.3389/feduc.2025.1647573>
- Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in teaching and learning: A systematic review. *Education Sciences*, 14(6), 643. <https://doi.org/10.3390/educsci14060643>
- Cisłowska, A. I., & Pena-Acuna, B. (2024). Integration of chatbots in additional language education: A systematic review. *European Journal of Educational Research*, 13(4), 1607–1625. <https://doi.org/10.12973/eu-jer.13.4.1607>
- Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is ChatGPT a blessing or a curse? *Frontiers in Education*, 8, 1166682. <https://doi.org/10.3389/feduc.2023.1166682>
- Herrera, P. (2021). Educación y desarrollo de competencias digitales en América Latina y el Caribe. Comisión Económica para América Latina y el Caribe.
- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/10.30935/cedtech/13605>
- Jeon, J., Lee, S., & Choi, S. (2023). A systematic review of research on speech-recognition chatbots for language learning: Implications for future directions in the era of large language models. *Interactive Learning Environments*. <https://doi.org/10.1080/10494820.2023.2204343>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>

- Kim, S., & Park, S. H. (2023). Young Korean EFL learners' perception of role-playing scripts: ChatGPT vs. textbooks. *Korean Journal of English Language and Linguistics*, 23, 1136–1153. <https://doi.org/10.15738/kjell.23.202312.1136>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>
- Kooli, C. (2023). Chatbots in education and research: A critical examination of ethical implications and solutions. *Sustainability*, 15(7), 5614. <https://doi.org/10.3390/su15075614>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, 6, 100174. <https://doi.org/10.1016/j.caeo.2024.100174>
- Lee, H., Hsia, C. C., Tsoy, A., Choi, S., Hou, H., & Ni, S. (2023). VisionARy: Exploratory research on contextual language learning using AR glasses with ChatGPT. *Proceedings of the 15th Biannual Conference of the Italian SIGCHI Chapter*, 22, 1–6. <https://doi.org/10.1145/3605390.3605400>
- Lee, S. J., & Kwon, K. (2024). A systematic review of AI education in K-12 classrooms from 2018 to 2023: Topics, strategies, and learning outcomes. *Computers and Education: Artificial Intelligence*, 6, 100211. <https://doi.org/10.1016/j.caeai.2024.100211>
- Li, B., Lowell, V. L., Wang, C., & Li, X. (2024). A systematic review of the first year of publications on ChatGPT and language education: Examining research on ChatGPT's use in language learning and teaching. *Computers and Education: Artificial Intelligence*, 7, 100266. <https://doi.org/10.1016/j.caeai.2024.100266>
- Liu, G., & Ma, C. (2024). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 18(2), 125–138. <https://doi.org/10.1080/17501229.2023.2240316>

- Liu, M., Zhang, L. J., & Biebricher, C. (2024). Investigating students' cognitive processes in generative AI-assisted digital multimodal composing and traditional writing. *Computers & Education*, 211, 104977. <https://doi.org/10.1016/j.compedu.2023.104977>
- Lo, C. K., Hew, K. F., & Jong, M. S. Y. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education*, 219, 105100. <https://doi.org/10.1016/j.compedu.2024.105100>
- Lo, C. K., Yu, P. L. H., Xu, S., Ng, D. T. K., & Jong, M. S. Y. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. *Smart Learning Environments*, 11, 50. <https://doi.org/10.1186/s40561-024-00342-5>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11, 9. <https://doi.org/10.1186/s40561-024-00295-9>
- Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, 10(2), 2236469. <https://doi.org/10.1080/2331186X.2023.2236469>
- Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO.
- Ministerio de Educación del Perú. (2016). Currículo Nacional de la Educación Básica. MINEDU.
- Ministerio de Educación del Perú. (2022). Estrategias para favorecer el desarrollo de las competencias asociadas al área de Comunicación. MINEDU.
- Ministerio de Educación del Perú. (2024). Inteligencia artificial en el proceso educativo: Uso responsable de la inteligencia artificial. MINEDU.
- Mizumoto, A., & Eguchi, M. (2023). Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), 100050. <https://doi.org/10.1016/j.rmal.2023.100050>
- Mizumoto, A., Shintani, N., Sasaki, M., & Teng, M. F. (2024). Testing the viability of ChatGPT as a companion in L2 writing accuracy assessment. *Research Methods in Applied Linguistics*, 3(2), 100116. <https://doi.org/10.1016/j.rmal.2024.100116>

- Mustafa, M. Y., Tlili, A., Lampropoulos, G., Huang, R., & Kinshuk. (2024). A systematic review of literature reviews on artificial intelligence in education (AIED): A roadmap to a future research agenda. *Smart Learning Environments*, 11, 59. <https://doi.org/10.1186/s40561-024-00350-5>
- Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 13(9), 5783. <https://doi.org/10.3390/app13095783>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342–363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia*, 126(5), 1763–1768. <https://doi.org/10.1213/ANE.0000000000002864>
- Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigation in Health, Psychology and Education*, 13(9), 1937–1960. <https://doi.org/10.3390/ejihpe13090140>
- Taber, K. S. (2018). The use of Cronbach’s alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48, 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Vo, T. K. A., & Nguyen, H. (2024). Generative artificial intelligence and ChatGPT in language learning: EFL students’ perceptions of technology acceptance. *Journal of University Teaching and Learning Practice*, 21(6). <https://doi.org/10.53761/fr1rkj58>
- Wambsganss, T., Janson, A., & Leimeister, J. M. (2022). Enhancing argumentative writing with automated feedback and social comparison nudging. *Computers & Education*, 191, 104644. <https://doi.org/10.1016/j.compedu.2022.104644>
- Xiao, F., Zhu, S., & Xin, W. (2026). Exploring the landscape of generative AI (ChatGPT)-powered writing instruction in English as a foreign language education: A scoping review. *ECNU Review of Education*, 9(1). <https://doi.org/10.1177/20965311241310881>

- Xiao, Y., & Zhi, Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, 8(3), 212. <https://doi.org/10.3390/languages8030212>
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, 28, 13943–13967. <https://doi.org/10.1007/s10639-023-11742-4>
- Zhai, C., & Wibowo, S. (2023). A systematic review on artificial intelligence dialogue systems for enhancing English as foreign language students' interactional competence in the university. *Computers and Education: Artificial Intelligence*, 4, 100134. <https://doi.org/10.1016/j.caeai.2023.100134>
- Zhang, P., & Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. *European Journal of Education*, 59, e12599. <https://doi.org/10.1111/ejed.12599>
- Zhang, S., Shan, C., Lee, J. S. Y., Che, S., & Kim, J. H. (2023). Effect of chatbot-assisted language learning: A meta-analysis. *Education and Information Technologies*, 28, 15223–15243. <https://doi.org/10.1007/s10639-023-11805-6>

**Conflicto de intereses:**

Los autores declaran que no existe conflicto de interés