REVISTA INVECOM
ISSN en línea: 2739-0063

Inteligencia artificial generativa y comprensión lectora en estudiantes universitarios

Generative Artificial Intelligence and Reading Comprehension Among College Students

Recibido: 24/06/2026 - **Aceptado:** 12/09/2026

Jhonny Santiago Gutiérrez Ortega
https://orcid.org/0009-0004-9550-4486
jgutierrezo@unitru.edu.pe
Universidad Nacional de Trujillo. Trujillo, Perú

William Marcelino Paliza Franco
https://orcid.org/0000-0003-2289-2938
wpaliza7@ucvvirtual.edu.pe
Universidad César Vallejo. Lima, Perú

José Emilio Álvarez Trujillo
https://orcid.org/0000-0002-5148-9438
jalvarezt@unitru.edu.pe
Universidad Nacional de Trujillo. Trujillo, Perú

Janeth Tomanguilla Reyna
https://orcid.org/0000-0002-7460-7214
jtomanguillar@ucvvirtual.edu.pe
Universidad César Vallejo. Trujillo, Perú

Resumen

El objetivo de este trabajo fue determinar la relación entre la interacción académica con inteligencia artificial generativa (IAG) y la comprensión lectora en estudiantes universitarios. Se realizó un estudio cuantitativo, observacional, transversal y correlacional con 400 estudiantes mayores de 18 años de una universidad pública peruana, seleccionados mediante muestreo intencional por accesibilidad. Se aplicaron una escala de 20 ítems sobre interacción académica con IAG y una prueba de comprensión lectora con 10 indicadores dicotómicos. La estructura interna de ambas puntuaciones globales fue examinada mediante análisis factorial exploratorio y análisis paralelo. La consistencia interna fue adecuada para IAG (α = .858) y lectura (KR-20 = .787). Se identificó una asociación positiva y fuerte entre las puntuaciones globales, r_s = .782, IC 95 % [.737, .819], p < .001. La relación fue mayor con la comprensión inferencial (r_s = .720) que con la literal (r_s = .609) y la crítica (r_s = .571). Los seis dominios de IAG también presentaron asociaciones positivas con la lectura total. Se concluyó que una interacción académica más informada, frecuente y reflexiva con la IAG se relacionó con un mejor desempeño lector. El diseño transversal, el muestreo no probabilístico y las limitaciones de medición impiden establecer causalidad y restringen la generalización.

Palabras clave: inteligencia artificial, comprensión de lectura, estudiante universitario.

Abstract

The objective of this study was to determine the relationship between academic interaction with generative artificial intelligence (GAI) and reading comprehension among college students. A quantitative, observational, cross-sectional, and correlational study was conducted with 400 students over the age of 18 from a Peruvian public university, selected through purposive sampling based on accessibility. A 20-item scale measuring academic interaction with GAI and a reading comprehension test with 10 dichotomous indicators were administered. The internal structure of both overall scores was examined using exploratory factor analysis and parallel analysis. Internal consistency was adequate for the IAG (α = .858) and the reading test (KR-20 = .787). A strong positive association was identified between the total scores, r_s = .782, 95% CI [.737, .819], p < .001. The relationship was stronger with inferential comprehension (r_s = .720) than with literal comprehension (r_s = .609) and critical comprehension (r_s = .571). The six domains of IAG also showed positive associations with total reading scores. It

SciELO Venezuela latindex LatinREV CC BY

was concluded that a more informed, frequent, and reflective academic engagement with IAG was associated with better reading performance. The cross-sectional design, non-probability sampling, and measurement limitations prevent the establishment of causality and restrict generalizability.

Keywords: artificial intelligence, reading comprehension, college student.

Introducción

La inteligencia artificial generativa (IAG) se incorporó rápidamente a las prácticas universitarias mediante herramientas capaces de producir explicaciones, resúmenes, borradores y retroalimentación. Su valor educativo no depende solo del acceso, sino de la finalidad de uso, la calidad de las instrucciones y la revisión humana de sus respuestas (Chen & Gong, 2025; Cordero et al., 2025). La evidencia sobre educación superior describe aplicaciones heterogéneas y advierte que la adopción tecnológica debe examinarse junto con los procesos cognitivos que sostiene o desplaza (Crompton & Burke, 2023).

Las percepciones estudiantiles combinan utilidad y cautela. Chan y Hu (2023) encontraron una valoración favorable del apoyo personalizado y la eficiencia, junto con preocupaciones sobre exactitud, privacidad y ética. Resultados semejantes se observaron en estudios sobre familiaridad, beneficios y riesgos (Almassaad et al., 2024; Gasaymeh et al., 2024). La aceptación de ChatGPT se relaciona con expectativas de desempeño y hábito (Strzelecki, 2024), pero su utilización puede producir consecuencias favorables o problemáticas según los motivos y patrones de uso (Abbas et al., 2024). En escritura académica, las ventajas para organizar y expresar ideas coexisten con preguntas sobre autoría e integridad (Alkamel & Alwagieh, 2024; Barrett & Pack, 2023; Cotton et al., 2024).

En este estudio, la interacción académica con IAG comprende seis dominios de contenido: conocimiento, frecuencia de uso, beneficios percibidos, conciencia de riesgos, integración pedagógica y responsabilidad ética. Esta delimitación recoge componentes examinados en escalas y encuestas recientes (Almassaad et al., 2024; Nemt-Allah et al., 2024). No obstante, los dominios no se asumieron como subescalas independientes: su interpretación principal recayó en una puntuación general, cuya estructura se evaluó empíricamente en la muestra.

La comprensión lectora implica recuperar información explícita, construir inferencias y valorar críticamente las afirmaciones. Gordillo Alfonso y Flórez (2009) distinguen los niveles literal, inferencial y crítico; en el ámbito universitario, estos procesos permiten interpretar discursos especializados, comparar evidencias y adoptar posiciones razonadas. La relación con la IAG es ambivalente: la simplificación asistida puede apoyar la comprensión y la inferencia (Celik et al., 2024), pero la disponibilidad de respuestas no elimina la necesidad de interpretar y verificar el contenido (Juanda & Afandi, 2024).

Los modelos generativos producen textos plausibles que pueden contener errores o sesgos (Alvero et al., 2024). También persisten dificultades para reconocer la autoría sintética (Fleckenstein et al., 2024) y preocupaciones sobre exactitud, privacidad e integridad (Tlili et al., 2023). Por ello, la alfabetización crítica en IAG demanda evaluar la pertinencia y confiabilidad del producto, decidir qué aceptar o rechazar y conservar la responsabilidad epistémica durante la colaboración humano-IA (Ou et al., 2024a, 2024b; Nguyen et al., 2024).

La asistencia generativa puede apoyar el aprendizaje lingüístico y la retroalimentación (Karataş et al., 2024; Mohammed & Khalid, 2025; Muslimin et al., 2024); sin embargo, el uso sustitutivo puede reducir la planificación, el monitoreo y la elaboración profunda. Fan et al. (2025) denominaron pereza metacognitiva a este desplazamiento, mientras que Stojanov et al. (2024) identificaron perfiles diferenciados de dependencia. El uso productivo exige formular instrucciones y evaluar críticamente los resultados (Kim et al., 2025).

La investigación disponible se concentra en percepciones, escritura y aceptación, y ofrece menos evidencia sobre el desempeño lector objetivo. En América Latina, los antecedentes siguen siendo escasos y se orientan principalmente a la evaluación y utilidad didáctica (Charpentier-Jiménez, 2024; Ortega-Rodríguez & Pericacho-Gómez, 2025). Una síntesis reciente confirma oportunidades de apoyo a la lectura y la escritura, pero también dependencia, pérdida de reflexión y heterogeneidad metodológica (Sanz-Tejeda et al., 2026). En consecuencia, permanece abierto el problema de si la interacción académica con IAG se relaciona con la comprensión lectora en universitarios peruanos.

El objetivo general fue determinar la relación entre la interacción académica con IAG y la comprensión lectora en estudiantes universitarios. Los objetivos específicos fueron describir los dominios de IAG, caracterizar los niveles literal, inferencial y crítico, y estimar las relaciones entre los dominios de IAG y los componentes lectores. Se planteó que ambas puntuaciones globales presentarían una relación estadísticamente significativa.

Metodología

Enfoque y diseño

Se realizó un estudio cuantitativo, observacional, transversal y correlacional. Las variables fueron medidas una sola vez, sin manipulación ni asignación de condiciones. El diseño permitió estimar asociaciones, pero no establecer temporalidad ni causalidad.

Participantes y muestreo

Participaron 400 estudiantes de 18 años o más de la Universidad Nacional de Trujillo. Se utilizó un muestreo no probabilístico intencional por accesibilidad. Fueron invitados estudiantes disponibles de Educación y Ciencias de la Comunicación, Ingeniería Agroindustrial, Ingeniería Agrícola, Ingeniería de Sistemas, Ingeniería Química, Ingeniería Informática, Administración y Economía. Se incluyeron estudiantes matriculados en 2026 que aceptaron participar y completaron ambos instrumentos; se excluyeron registros incompletos, duplicados o técnicamente inválidos. Una sensibilidad aproximada indicó que n = 400 permitía detectar correlaciones cercanas a .14 con α = .05, potencia de .80 y contraste bilateral.

Instrumentos

Interacción académica con IAG. Se aplicó una escala de 20 ítems con respuestas de 1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo). Los reactivos cubrieron conocimiento (1-3), frecuencia (4-7), beneficios (8-10), riesgos y preocupaciones (11-13), integración pedagógica (14-17) y ética (18-20). El total osciló teóricamente entre 20 y 100; una puntuación mayor representó una interacción académica más amplia, incluyendo mayor conciencia de riesgos y responsabilidad ética. Los dominios se consideraron indicadores de contenido y sus análisis se trataron como exploratorios. Para la descripción se utilizaron intervalos de igual amplitud: 20-46, 47-73 y 74-100 para el total; 3-7, 8-11 y 12-15 en dominios de tres ítems; y 4-9, 10-15 y 16-20 en dominios de cuatro ítems.

Comprensión lectora. Se utilizó la prueba diagnóstica de Gordillo Alfonso y Flórez (2009), basada en el texto argumentativo «Es imposible ganar la guerra contra las drogas». La matriz consolidada registró 10 indicadores dicotómicos: tres literales, cinco inferenciales y dos críticos. Cada indicador recibió 1 cuando la respuesta cumplió el criterio de logro y 0 cuando fue incorrecta o insuficiente. El total varió entre 0 y 10. Los niveles se definieron así: literal, bajo 0-1, medio 2 y alto 3; inferencial, bajo 0-1, medio 2-3 y alto 4-5; crítico, bajo 0, medio 1 y alto 2; total, bajo 0-3, medio 4-6 y alto 7-10.

Procedimiento

La aplicación se realizó presencialmente en junio de 2026. Se explicó el propósito, la voluntariedad, la confidencialidad y las instrucciones. Los registros se identificaron con códigos EST-001 a EST-400. La depuración incluyó la verificación de rangos, valores ausentes, duplicados y sumas: los 400 códigos fueron únicos, no hubo datos perdidos ni perfiles completos duplicados y los totales coincidieron con sus ítems.

Análisis de datos

Los análisis se reprodujeron en Python 3.12 con SciPy 1.17 y scikit-learn 1.8. Se calcularon frecuencias, porcentajes, medias y desviaciones estándar. La consistencia interna se estimó mediante alfa de Cronbach y KR-20. Como evidencia exploratoria de estructura interna, se calcularon KMO, la prueba de esfericidad de Bartlett, análisis factorial de máxima verosimilitud y análisis paralelo con 1000 matrices aleatorias. Las asociaciones se estimaron mediante rho de Spearman; para el resultado principal se obtuvo un IC de 95 % con remuestreo bootstrap. Las 28 asociaciones exploratorias de la Tabla 5 se sometieron a ajuste de Holm y conservaron $p < .001$. Como análisis complementario, se calcularon chi-cuadrado y V de Cramer. Se adoptó α = .05.

Consideraciones éticas

El estudio involucró a adultos y procedimientos educativos de riesgo mínimo. La participación fue voluntaria y sin consecuencias académicas; se informó el propósito y el uso científico de los datos, y la matriz analítica no incluyó identificadores directos. El protocolo no fue evaluado previamente por un comité de ética, situación que se declara sin atribuir una aprobación inexistente. Para la postulación se recomienda adjuntar, si la institución competente lo emite, un pronunciamiento formal sobre la exención aplicable; esta gestión no puede sustituirse mediante aprobación retrospectiva.

Resultados y discusión

Calidad de los datos y estructura de las puntuaciones

La puntuación global de IAG mostró consistencia adecuada (α = .858). El KMO fue .920 y la prueba de esfericidad de Bartlett resultó significativa, $\chi^2(190) = 1494.69$, $p < .001$. El análisis paralelo retuvo un factor general; sus cargas oscilaron entre .403 y .564 y explicó el 27.2 % de la varianza. En lectura, el KR-20 fue .787, KMO = .882 y la prueba de esfericidad de Bartlett fue significativa, $\chi^2(45) = 674.35$, $p < .001$; el único factor retenido explicó el 34.3 % de la varianza, con cargas entre .482 y .610. Estos resultados respaldaron el uso de las puntuaciones globales, pero no justificaron tratar los seis dominios de IAG como subescalas autónomas.

Tabla 1 *Estadísticos descriptivos y evidencia psicométrica de las puntuaciones globales*

Instrumento	Ítems	Rango observado	M	DE	Fiabilidad	KMO
Interacción con IAG	20	24-95	59.77	14.73	α = .858	.920
Comprensión lectora	10	0-10	4.61	2.92	KR-20 = .787	.882

Nota. n = 400. En ambos instrumentos, Bartlett fue significativo (p < .001) y el análisis paralelo retuvo un factor general.

Tabla 2 *Niveles de interacción académica con IAG y dominios de contenido*

Variable/dominio	Bajo f (%)	Medio f (%)	Alto f (%)	Total
IAG total	78 (19.5)	245 (61.3)	77 (19.2)	400
Conocimiento	128 (32.0)	182 (45.5)	90 (22.5)	400
Frecuencia	106 (26.5)	227 (56.8)	67 (16.8)	400
Beneficios	140 (35.0)	176 (44.0)	84 (21.0)	400
Riesgos	133 (33.2)	173 (43.2)	94 (23.5)	400
Integración	115 (28.7)	214 (53.5)	71 (17.8)	400
Ética	128 (32.0)	196 (49.0)	76 (19.0)	400

Nota. Los niveles se construyeron mediante intervalos de igual amplitud y se utilizaron con fines descriptivos.

Predominó el nivel medio de interacción con IAG. El conocimiento presentó la mayor proporción en el nivel alto, mientras la frecuencia y la integración se concentraron en el nivel medio. El patrón sugiere una adopción extendida, pero todavía desigual. Esta combinación de familiaridad, beneficios y reservas coincide con Chan y Hu (2023), Almassaad et al. (2024) y Gasaymeh et al. (2024).

Tabla 3 *Niveles de comprensión lectora*

Componente	Bajo f (%)	Medio f (%)	Alto f (%)	Total
Literal	224 (56.0)	95 (23.8)	81 (20.2)	400
Inferencial	149 (37.2)	136 (34.0)	115 (28.7)	400
Crítica	143 (35.8)	151 (37.8)	106 (26.5)	400
Total	202 (50.5)	116 (29.0)	82 (20.5)	400

Nota. Los puntos de corte se especifican en la metodología.

La mitad de los participantes se ubicó en el nivel bajo de comprensión total y el nivel literal concentró la mayor proporción en el nivel bajo. El resultado indica dificultades para recuperar y articular información antes de realizar operaciones inferenciales y críticas. Esta secuencia es coherente con la interdependencia de los niveles descrita por Gordillo Alfonso y Flórez (2009).

Tabla 4 *Correlación entre interacción con IAG y comprensión lectora*

Variables	n	r_s	IC 95 %	p	Magnitud
IAG total - lectura total	400	.782	[.737, .819]	< .001	Fuerte

Nota. IC calculado mediante bootstrap. La clasificación de magnitud se utiliza de forma descriptiva.

En relación con el objetivo general, se buscó determinar la relación entre la interacción académica con inteligencia artificial generativa y la comprensión lectora en estudiantes universitarios. Este planteamiento se sustentó en los aportes de Ou et al. (2024a, 2024b), quienes conciben la alfabetización crítica en IAG como la capacidad para comprender el funcionamiento y las limitaciones de estas herramientas, valorar la confiabilidad de sus respuestas y decidir de manera responsable qué información aceptar, cuestionar o rechazar. La importancia de este enfoque radica en que la interacción académica con IAG no se reduce a la frecuencia de uso, sino que involucra conocimientos, percepción de beneficios, conciencia de riesgos, integración pedagógica y responsabilidad ética.

En relación con ello, Nguyen et al. (2024) explicaron la interacción con sistemas generativos desde una perspectiva de colaboración humano-IA. En esta relación, la herramienta puede apoyar la búsqueda, organización y reformulación de información, pero el estudiante conserva la responsabilidad de interpretar, contrastar y validar el contenido producido. Desde este enfoque, la comprensión lectora interviene antes, durante y después del uso de la IAG: permite formular solicitudes pertinentes, comprender las respuestas generadas, compararlas con el texto original y detectar omisiones, inconsistencias o afirmaciones carentes de sustento. Por tanto, la interacción productiva con IAG no sustituye la lectura, sino que demanda una participación cognitiva activa del estudiante.

Los resultados presentados en la Tabla 4 evidenciaron una relación positiva, fuerte y estadísticamente significativa entre la interacción académica con IAG y la comprensión lectora (r_s = .782; IC 95 % [.737, .819]; p < .001). En consecuencia, se rechazó la hipótesis nula. El coeficiente indicó que los estudiantes con puntuaciones más altas en la interacción académica con IAG tendieron a obtener mejores resultados en comprensión lectora. Asimismo, el intervalo de confianza fue relativamente estrecho y mantuvo sus límites dentro de una magnitud elevada, lo que mostró precisión en la estimación de la asociación dentro de la muestra analizada.

El resultado coincidió parcialmente con Celik et al. (2024), quienes encontraron que la simplificación de textos mediante inteligencia artificial podía favorecer la comprensión y la elaboración de inferencias. La coincidencia se explica porque, en ambos estudios, la IAG facilitó operaciones relacionadas con la lectura, como aclarar expresiones complejas, reorganizar información y hacer más accesible el contenido. Estas funciones pueden reducir dificultades lingüísticas iniciales y permitir que el estudiante concentre sus recursos cognitivos en establecer relaciones, completar significados y comprender la estructura global del texto.

Sin embargo, los estudios no fueron directamente equivalentes. Celik et al. (2024) evaluaron una intervención específica de simplificación textual, mientras que la presente investigación analizó una puntuación global de interacción académica con IAG mediante un diseño transversal y correlacional. Por ello, sus resultados permiten respaldar la posibilidad de que determinadas funciones de la IAG apoyen la lectura, pero no demuestran que la asociación identificada en los estudiantes universitarios peruanos se haya originado exclusivamente en la simplificación de los textos.

El hallazgo también guardó relación con la síntesis de Sanz-Tejeda et al. (2026), quienes identificaron oportunidades de la IAG para apoyar la lectura y la escritura académicas. La coincidencia radicó en que las herramientas generativas pueden facilitar la explicación de conceptos, la producción de resúmenes, la reorganización de información y la retroalimentación inmediata. Estas acciones pueden funcionar como formas de andamiaje cuando ayudan al estudiante a regresar al texto, contrastar interpretaciones y elaborar una comprensión propia.

No obstante, Sanz-Tejeda et al. (2026) también advirtieron sobre la dependencia, la pérdida de reflexión y la heterogeneidad de los resultados disponibles. Esta advertencia permite precisar que la asociación positiva encontrada no debe atribuirse a cualquier forma de uso. La interacción evaluada comprendió conocimiento, beneficios, riesgos, integración pedagógica y responsabilidad ética; por tanto, una puntuación elevada representó una relación más amplia e informada con la IAG y no únicamente una mayor frecuencia de acceso.

Los resultados contrastaron con Fan et al. (2025), quienes advirtieron que la sustitución de procesos cognitivos por respuestas automatizadas puede favorecer la denominada pereza metacognitiva. La diferencia puede explicarse por la calidad del uso. Cuando la IAG se utiliza para obtener una respuesta terminada y evitar la lectura, puede reducir la planificación, el monitoreo y la revisión. En cambio, cuando se emplea para formular

preguntas, comparar explicaciones y verificar la coherencia de una respuesta, puede estimular las mismas operaciones cognitivas que intervienen en la comprensión lectora.

Esta interpretación encuentra apoyo en Abbas et al. (2024) y Stojanov et al. (2024), quienes mostraron que el uso académico de ChatGPT no constituye una práctica homogénea. Los estudiantes pueden adoptar patrones de apoyo, complementariedad o dependencia. La coincidencia con el presente estudio se encuentra en que los efectos educativos no dependen exclusivamente de utilizar la herramienta, sino de la finalidad, la intensidad y el grado de control cognitivo conservado por el estudiante. Por ello, una interacción reflexiva puede relacionarse con mejores resultados, mientras que un uso sustitutivo puede debilitar la elaboración personal.

Una primera explicación del coeficiente obtenido es que la IAG haya funcionado como andamiaje cognitivo. Los estudiantes pudieron emplearla para aclarar conceptos, solicitar ejemplos, reformular fragmentos complejos, identificar ideas centrales y contrastar interpretaciones. Estas acciones podrían favorecer la integración de información y la construcción de inferencias, siempre que las respuestas generadas fueran sometidas a evaluación crítica y no aceptadas automáticamente.

Sin embargo, también debe considerarse una explicación inversa. Los estudiantes con mejores capacidades lectoras pudieron estar en condiciones de interactuar de manera más efectiva con la IAG. Una comprensión más desarrollada facilita formular instrucciones precisas, reconocer si una respuesta conserva el sentido del texto, identificar errores y solicitar correcciones. Desde esta perspectiva, el desempeño lector no sería únicamente un posible resultado de la interacción con IAG, sino también una capacidad previa que mejora la calidad de dicha interacción.

Asimismo, la asociación pudo estar influida por variables no controladas. El rendimiento académico general, la experiencia tecnológica, el área profesional, el ciclo de estudios, la frecuencia de lectura, la autorregulación y el acceso a recursos digitales podrían relacionarse simultáneamente con el uso de IAG y la comprensión lectora. La ausencia de control sobre estos factores impide establecer cuánto de la correlación correspondió específicamente a la interacción con las herramientas generativas.

También debe considerarse la diferencia entre las mediciones. La interacción académica con IAG se evaluó mediante autoinforme, mientras que la comprensión lectora se examinó mediante indicadores de desempeño. Aunque esta combinación disminuye el riesgo de que toda la asociación proceda de un mismo método de medición, las respuestas sobre IAG pudieron estar influidas por la deseabilidad social o por la percepción que cada estudiante tenía de sus propias competencias. Por ello, futuros estudios deberían complementar los autoinformes con registros de uso, análisis de las instrucciones formuladas y evaluación de las respuestas generadas.

En consecuencia, la relación observada no demostró que la IAG hubiera producido una mejora en la comprensión lectora. El diseño transversal permitió establecer covariación, pero no temporalidad ni causalidad. Tampoco fue posible determinar si la interacción con IAG fortaleció la lectura, si los mejores lectores utilizaron estas herramientas con mayor eficacia o si ambas variables dependieron de capacidades académicas previas.

Tabla 5 *Correlaciones exploratorias de los dominios de IAG con los componentes lectores*

Dominio IAG	Literal r_s	Inferencial r_s	Crítica r_s	Total r_s	IC 95 % total
Conocimiento	.413	.544	.447	.578	[.509, .640]
Frecuencia	.468	.588	.447	.621	[.557, .678]
Beneficios	.455	.480	.432	.555	[.484, .620]
Riesgos	.486	.536	.403	.586	[.518, .647]
Integración	.503	.552	.419	.607	[.541, .666]
Ética	.443	.566	.427	.594	[.527, .654]
IAG total	.609	.720	.571	.782	[.737, .819]

Nota. Todas las correlaciones fueron significativas después del ajuste de Holm (p ajustada < .001). Los dominios son indicadores exploratorios de contenido, no subescalas independientes.

La comprensión inferencial presentó la asociación más elevada con la puntuación global de IAG. Este patrón es coherente con tareas de uso académico que requieren completar significados, relacionar información y valorar la coherencia. Celik et al. (2024) observaron beneficios en la inferencia bajo una intervención de simplificación, mientras Juanda y Afandi (2024) mostraron que la respuesta automatizada no sustituye la interpretación humana.

La frecuencia y la integración mostraron las asociaciones más altas con la lectura total entre los dominios, pero sus diferencias fueron pequeñas y no deben interpretarse como una jerarquía causal. La correlación positiva

del dominio de riesgos refleja mayor conciencia de problemas de fiabilidad, dependencia y privacidad, no mayor exposición al daño. Del mismo modo, la ética representa verificación y responsabilidad. Estas relaciones son coherentes con la necesidad de evaluación crítica señalada por Tlili et al. (2023), Alvero et al. (2024) y Fleckenstein et al. (2024).

Tabla 6 *Distribución conjunta de los niveles globales de IAG y comprensión lectora*

Nivel de interacción con IAG	Comprensión baja f (%)	Comprensión media f (%)	Comprensión alta f (%)	Total
Bajo	70 (89,7)	8 (10,3)	0 (0,0)	78
Medio	131 (53,5)	91 (37,1)	23 (9,4)	245
Alto	1 (1,3)	17 (22,1)	59 (76,6)	77
Total	202 (50,5)	116 (29,0)	82 (20,5)	400

Nota. Datos tomados de la muestra.

Para esta distribución se obtiene:

$\chi^2(4) = 226{,}51$

$p < ,001$

V de Cramer = ,532

Asociación estadísticamente significativa y de magnitud alta.

La distribución conjunta mostró un patrón progresivo entre ambas variables. Entre los estudiantes con baja interacción académica con IAG predominó ampliamente la comprensión lectora baja, sin registrarse casos en el nivel alto. En el grupo con interacción media, la comprensión baja continuó siendo la categoría más frecuente, aunque aumentó la presencia del nivel medio. En cambio, entre quienes presentaron alta interacción con IAG predominó la comprensión lectora alta. La prueba de chi-cuadrado confirmó que la distribución de los niveles lectores varió significativamente según el nivel de interacción con IAG, mientras que la V de Cramer evidenció una asociación de magnitud alta. Este resultado complementa la correlación de Spearman obtenida con las puntuaciones originales y muestra que los niveles elevados de interacción académica con IAG tendieron a presentarse conjuntamente con niveles superiores de comprensión lectora. Sin embargo, debido al diseño transversal y correlacional, esta asociación no permite establecer una relación causal.

Limitaciones e implicaciones

El diseño transversal no permite establecer temporalidad ni causalidad. El muestreo por accesibilidad y la participación de una sola universidad restringen la generalización. La escala de IAG reúne dominios conceptualmente diversos; aunque la evidencia exploratoria respaldó un factor general, la varianza explicada fue limitada y los dominios no deben tratarse como subescalas consolidadas. El uso de IAG fue autorreportado y susceptible a deseabilidad social. La prueba lectora redujo tareas complejas a indicadores dicotómicos y no se dispuso de una segunda calificación independiente para estimar el acuerdo interevaluador. Tampoco se midieron rendimiento previo, hábitos lectores, motivación ni competencia digital. Finalmente, la ausencia de revisión ética previa constituye una limitación procedimental y editorial.

Futuros estudios deberían validar los instrumentos mediante jueces, entrevistas cognitivas, análisis factorial confirmatorio y confiabilidad interevaluador; utilizar muestras probabilísticas o multicéntricas; registrar interacciones objetivas con IAG; y emplear diseños longitudinales o experimentales. En la práctica, los resultados apoyan actividades que exijan contrastar la respuesta generada con fuentes, justificar inferencias, identificar sesgos y documentar la participación de la IAG. La evidencia no respalda aumentar indiscriminadamente la frecuencia de uso.

Conclusiones

La interacción académica con IAG se relacionó de manera positiva y fuerte con la comprensión lectora en la muestra analizada, con una asociación particularmente elevada en el componente inferencial. Los seis dominios de contenido mostraron relaciones positivas, pero deben interpretarse de forma exploratoria y no como factores independientes ni como predictores causales.

Los hallazgos respaldan una integración pedagógica basada en la alfabetización crítica, la verificación de información y la responsabilidad ética. No demuestran que la IAG mejore por sí misma la lectura: también es

plausible que estudiantes con mejores competencias lectoras utilicen la tecnología de manera más estratégica. La principal contribución es documentar esta relación en universitarios de una institución pública peruana y delimitar condiciones para investigarla con mayor rigor.

Referencias

Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education, 21*, 10. https://doi.org/10.1186/s41239-024-00444-7

Alkamel, M. A. A., & Alwagieh, N. A. S. (2024). Utilizing an adaptable artificial intelligence writing tool (ChatGPT) to enhance academic writing skills among Yemeni university EFL students. *Social Sciences & Humanities Open, 10*, 101095. https://doi.org/10.1016/j.ssaho.2024.101095

Almassaad, A., Alajlan, H., & Alebaikan, R. (2024). Student perceptions of generative artificial intelligence: Investigating utilization, benefits, and challenges in higher education. *Systems, 12*(10), 385. https://doi.org/10.3390/systems12100385

Alvero, A. J., Lee, J., Regla-Vargas, A., Kizilcec, R. F., Joachims, T., & Antonio, A. L. (2024). Large language models, social demography, and hegemony: Comparing authorship in human and synthetic text. *Journal of Big Data, 11*, 138. https://doi.org/10.1186/s40537-024-00986-7

Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education, 20*, 59. https://doi.org/10.1186/s41239-023-00427-0

Celik, F., Yangın Ersanlı, C., & Arslanbay, G. (2024). Does AI simplification of authentic blog texts improve reading comprehension, inferencing, and anxiety? A one-shot intervention in Turkish EFL context. *International Review of Research in Open and Distributed Learning, 25*(3), 287–303. https://doi.org/10.19173/irrodl.v25i3.7779

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*, 43. https://doi.org/10.1186/s41239-023-00411-8

Charpentier-Jiménez, W. (2024). Assessing artificial intelligence and professors' calibration in English as a foreign language writing courses at a Costa Rican public university. *Actualidades Investigativas en Educación, 24*(1), 533–557. https://doi.org/10.15517/aie.v24i1.55612

Chen, C., & Gong, Y. (2025). The role of AI-assisted learning in academic writing: A mixed-methods study on Chinese as a second language students. *Education Sciences, 15*(2), 141. https://doi.org/10.3390/educsci15020141

Cordero, J., Torres-Zambrano, J., & Cordero-Castillo, A. (2025). Integration of generative artificial intelligence in higher education: Best practices. *Education Sciences, 15*(1), 32. https://doi.org/10.3390/educsci15010032

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*, 22. https://doi.org/10.1186/s41239-023-00392-8

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology, 56*(2), 489–530. https://doi.org/10.1111/bjet.13544

Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence, 6*, 100209. https://doi.org/10.1016/j.caeai.2024.100209

Gasaymeh, A.-M. M., Beirat, M. A., & Abu Qbeita, A. A. (2024). University students' insights of generative artificial intelligence (AI) writing tools. *Education Sciences, 14*(10), 1062. https://doi.org/10.3390/educsci14101062

Gordillo Alfonso, A., & Flórez, M. del P. (2009). Los niveles de comprensión lectora: Hacia una enunciación investigativa y reflexiva para mejorar la comprensión lectora en estudiantes universitarios. *Actualidades Pedagógicas*, (53), 95–107. https://doi.org/10.19052/ap.1048

Juanda, J., & Afandi, I. (2024). Assessing text comprehension proficiency: Indonesian higher education students

vs. ChatGPT. *XLinguae, 17*(1), 49–68. https://doi.org/10.18355/XL.2024.17.01.04
Karataş, F., Yaşar Abedi, F., Ozek Gunyel, F., Karadeniz, D., & Kuzgun, Y. (2024). Incorporating AI in foreign language education: An investigation into ChatGPT's effect on foreign language learners. *Education and Information Technologies, 29*, 19343–19366. https://doi.org/10.1007/s10639-024-12574-6
Kim, J., Yu, S., Detrick, R., & Li, N. (2025). Exploring students' perspectives on generative AI-assisted academic writing. *Education and Information Technologies, 30*, 1265–1300. https://doi.org/10.1007/s10639-024-12878-7
Mohammed, S. J., & Khalid, M. W. (2025). Under the world of AI-generated feedback on writing: Mirroring motivation, foreign language peace of mind, trait emotional intelligence, and writing development. *Language Testing in Asia, 15*, 7. https://doi.org/10.1186/s40468-025-00343-2
Muslimin, A. I., Nur, M., & Francisca Maria, I. (2024). Evaluating Cami AI across SAMR stages: Students' achievement and perceptions in EFL writing instruction. *Online Learning, 28*(2). https://doi.org/10.24059/olj.v28i2.4246
Nemt-Allah, M., Khalifa, W., Badawy, M., Elbably, Y., & Ibrahim, A. (2024). Validating the ChatGPT usage scale: Psychometric properties and factor structures among postgraduate students. *BMC Psychology, 12*, 497. https://doi.org/10.1186/s40359-024-01983-4
Nguyen, A., Hong, Y., Dang, B., & Huang, X. (2024). Human-AI collaboration patterns in AI-assisted academic writing. *Studies in Higher Education, 49*(5), 847–864. https://doi.org/10.1080/03075079.2024.2323593
Ortega-Rodríguez, P. J., & Pericacho-Gómez, F. J. (2025). La utilidad didáctica percibida del ChatGPT por parte del alumnado universitario. *Pixel-Bit. Revista de Medios y Educación, 72*, 159–178. https://doi.org/10.12795/pixelbit.109778
Ou, A. W., Stöhr, C., & Malmström, H. (2024a). Academic communication with AI-powered language tools in higher education: From a post-humanist perspective. *System, 121*, 103225. https://doi.org/10.1016/j.system.2024.103225
Ou, A. W., Khuder, B., Franzetti, S., & Negretti, R. (2024b). Conceptualising and cultivating critical GAI literacy in doctoral academic writing. *Journal of Second Language Writing, 66*, 101156. https://doi.org/10.1016/j.jslw.2024.101156
Sanz-Tejeda, A., Domínguez-Oller, J. C., Baldaquí-Escandell, J. M., Gómez-Díaz, R., & García-Rodríguez, A. (2026). The impact of generative AI on academic reading and writing: A synthesis of recent evidence (2023–2025). *Frontiers in Education, 10*, 1711718. https://doi.org/10.3389/feduc.2025.1711718
Stojanov, A., Liu, Q., & Koh, J. H. L. (2024). University students' self-reported reliance on ChatGPT for learning: A latent profile analysis. *Computers and Education: Artificial Intelligence, 6*, 100243. https://doi.org/10.1016/j.caeai.2024.100243
Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education, 49*, 223–245. https://doi.org/10.1007/s10755-023-09686-1
Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, 15. https://doi.org/10.1186/s40561-023-00237-x

CONTRIBUCIÓN DE LA AUTORÍA

1. Conceptualización: Jhonny Santiago Gutiérrez Ortega, William Marcelino Paliza Franco, José Emilio Álvarez Trujillo y Janeth Tomanguilla Reyna.
2. Curación de datos: Jhonny Santiago Gutiérrez Ortega y Janeth Tomanguilla Reyna.
3. Análisis formal: William Marcelino Paliza Franco y José Emilio Álvarez Trujillo.
4. Adquisición de fondos: No se recibió financiación externa.
5. Investigación: Todos los autores.
6. Metodología: Todos los autores.
7. Dirección del proyecto: Jhonny Santiago Gutiérrez Ortega.
8. Recursos: Todos los autores.
9. Software: William Marcelino Paliza Franco y José Emilio Álvarez Trujillo.
10. Supervisión: Todos los autores.
11. Validación: Todos los autores.
12. Visualización: William Marcelino Paliza Franco.
13. Redacción – borrador original: Jhonny Santiago Gutiérrez Ortega y Janeth Tomanguilla Reyna.
14. Redacción – corrección de pruebas y edición: Todos los autores.