

La inteligencia artificial generativa como acompañante pedagógico

Generative Artificial Intelligence as a Pedagogical Companion

<https://doi.org/10.47606/ACVEN/PH0524>

Emma Lourdes Durand-Gómez^{*1}

<https://orcid.org/0000-0002-1896-9958>

edurand@ispa.edu.pe

Mario Alfredo Zevallos-Luna¹

<https://orcid.org/0000-0003-3477-8606>

mzevallos@ispa.edu.pe

Alfredo H. Laura-Anco¹

<https://orcid.org/0000-0001-9066-3649>

alaura@ispa.edu.pe

Isabel Olin-Zegarra¹

<https://orcid.org/0000-0002-3678-3997>

iolin@ispa.edu.pe

Magda Medina-Velásquez¹

<https://orcid.org/0000-0001-9184-4839>

mmedina@ispa.edu.pe

Recibido: 23/03/2026

Aceptado: 27/06/2026

RESUMEN

El presente artículo ofrece un estado del arte narrativo-sistematizador sobre la inteligencia artificial generativa (IAG) como acompañante pedagógico en la educación superior contemporánea. La pregunta orientadora fue: ¿cuáles son las funciones pedagógicas de la IAG, sus condiciones de eficacia y las tensiones éticas derivadas de su integración en contextos formales de educación superior, según la evidencia publicada entre 2022 y mayo de 2026? La búsqueda documental partió de consultas sistemáticas en Scopus, Web of Science, SciELO, Redalyc y Latindex 2.0. De 773 registros brutos, la eliminación de duplicados, el cribado por título y resumen, y la lectura analítica de texto completo con codificación independiente por dos autores dejaron un corpus de 29 artículos científicos publicados entre 2022 y mayo de 2026. Del análisis temático se derivaron cinco categorías que son la tutoría inteligente, el andamiaje adaptativo, la mediación dialógica, la evaluación formativa y las tensiones éticas y epistémicas emergentes. Los resultados muestran que la IAG fortalece la personalización del aprendizaje y la retroalimentación formativa en condiciones pedagógicas específicas. Las tensiones asociadas a la dependencia cognitiva, la autoría académica, los sesgos algorítmicos y la gobernanza educativa persisten sin que exista una resolución operacional. Se concluye que la IAG constituye un constructo sociotécnico cuya integración pedagógica depende de la calidad del diseño instruccional más que de la sofisticación técnica del modelo.

Palabras Clave: tutoría inteligente; andamiaje adaptativo; mediación dialógica; evaluación formativa; ética de la inteligencia artificial.

1. ¹Escuela de Educación Superior Pedagógica Pública Arequipa. Perú

* Autor de correspondencia: edurand@ispa.edu.pe

ABSTRACT

This article presents a narrative-systematizing state-of-the-art overview of generative artificial intelligence (GAI) as a pedagogical tool in contemporary higher education. The guiding question was: what are the pedagogical functions of GAI, its conditions for effectiveness, and the ethical tensions arising from its integration into formal higher education contexts, based on evidence published between 2022 and May 2026? The literature search involved systematic queries in Scopus, Web of Science, SciELO, Redalyc, and Latindex 2.0. From 773 raw records, duplicate removal, screening by title and abstract, and full-text analytical reading with independent coding by two authors yielded a corpus of 29 scientific articles published between 2022 and May 2026. Thematic analysis resulted in five categories: intelligent tutoring, adaptive scaffolding, dialogic mediation, formative assessment, and emerging ethical and epistemic tensions. The results show that GAI strengthens personalized learning and formative feedback under specific pedagogical conditions. However, tensions associated with cognitive dependency, academic authorship, algorithmic biases, and educational governance persist without an operational resolution. It is concluded that GAI is a sociotechnical construct whose pedagogical integration depends more on the quality of the instructional design than on the technical sophistication of the model.

Keywords: intelligent tutoring; adaptive scaffolding; dialogic mediation; formative assessment; artificial intelligence ethics.

INTRODUCCIÓN

La masificación de ChatGPT introdujo una discontinuidad en el debate pedagógico que la literatura académica aún procesa con marcos conceptuales en construcción. Los modelos de lenguaje (LLM) pasaron de ser objetos de especulación técnica a herramientas activas en las aulas universitarias con una velocidad que ninguna política educativa anticipó con suficiencia (Kasneci et al., 2023). La aceleración amplificó la dispersión metodológica en los estudios sobre la Inteligencia Artificial Generativa (IAG) educativa y la tendencia a tratar el acceso a la herramienta como un indicador de calidad pedagógica, sin que se llegue a examinar las condiciones bajo las cuales produce aprendizaje genuino.

Dos marcos teóricos del procesamiento cognitivo intentan atender el dilema de fondo. La teoría de la carga cognitiva de Sweller (1988) anticipaba que cualquier recurso capaz de reducir el esfuerzo de procesamiento puede liberar capacidad para el aprendizaje profundo o inhibirlo, debiéndose considerar la naturaleza de la tarea y la estructura de la ayuda ofrecida. La hipótesis de los niveles de procesamiento de Craik & Lockhart (1972) agregó que la huella mnémica durable dependía de la profundidad con que el sujeto elaboraba la información, y en tal sentido, una retroalimentación automática y fluida podría producir la ilusión de comprensión durante el proceso de estudio sin que se llegase a generar la elaboración cognitiva que el aprendizaje duradero requiere. Ambos marcos anticipaban lo que varios estudios recientes empiezan a documentar de forma convergente y es que la IAG no opera de manera neutral sobre el pensamiento crítico del estudiante.

La revisión de la literatura disponible revela un vacío analítico preciso. Los estudios sobre funciones pedagógicas específicas de la IAG centrado en la tutoría, el estudio del andamiaje, la mediación dialógica y la evaluación, se han producido de manera fragmentada sin que se aprecien marcos comparativos capaces de contrastar las

condiciones de eficacia entre las categorías ni de examinar las tensiones éticas de forma transversal. La presente revisión viene a contribuir para llenar ese espacio.

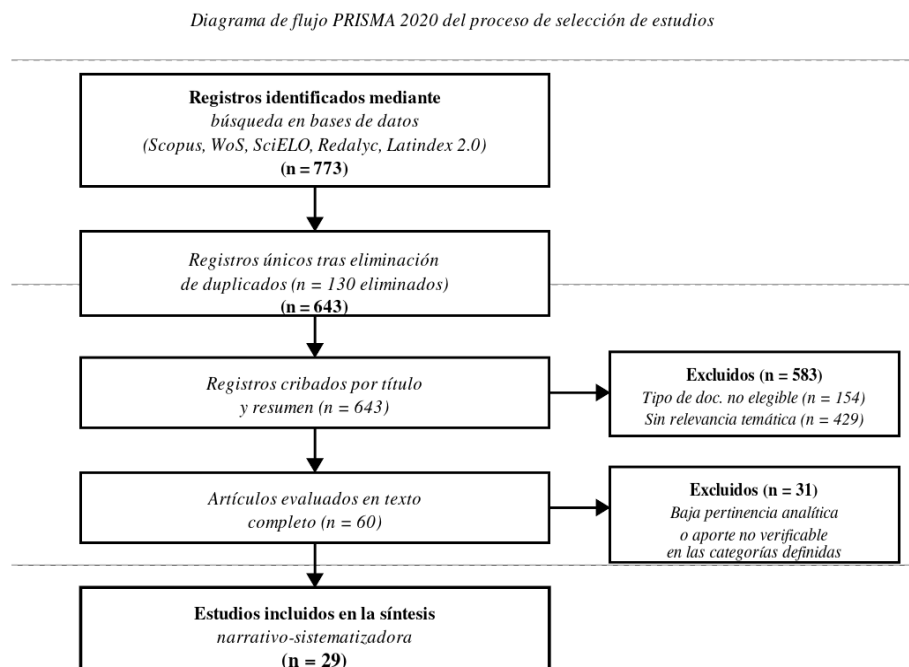
La pregunta orientadora en términos PICO [(P: estudiantes de educación superior; I: IAG como acompañante pedagógico; C: entornos con diseños instruccionales distintos o sin uso de IAG; O: calidad de la mediación pedagógica, autonomía del aprendiz y tensiones éticas derivadas)], fue: ¿cuáles son las funciones pedagógicas de la IAG, sus condiciones de eficacia y las tensiones éticas derivadas de su integración, según la evidencia publicada entre 2022 y mayo de 2026? El objetivo fue sistematizar críticamente esa evidencia e identificar tendencias, funciones emergentes, tensiones conceptuales y vacíos de investigación.

METODOLOGÍA

El diseño adoptado fue el de estado del arte narrativo-sistematizador con criterios de transparencia y trazabilidad del proceso de selección consonantes con el protocolo PRISMA 2020 (Page et al., 2021). La búsqueda fue desarrollada entre el 12 de abril y el 16 de mayo de 2026; partió del artículo de Kasneci et al. (2023) y se amplió con el mapeo bibliográfico en Connected Papers. Las bases consultadas fueron Scopus, Web of Science (WoS), SciELO, Redalyc y Latindex 2.0. Las ecuaciones combinaron operadores booleanos con los términos: “generative artificial intelligence”, “ChatGPT”, “large language model*”, “intelligent tutoring”, “adaptive scaffolding”, “dialogic mediation”, “formative assessment” y “education”.

Figura 1.

Diagrama de flujo PRISMA 2020 del proceso de selección de estudios



Nota. Construido conforme a Page et al. (2021). TI = Tutoría Inteligente; AA = Andamiaje Adaptativo; MD = Mediación Dialógica; EF = Evaluación Formativa; TE = Tensiones Éticas/Epistémicas.

La gestión de registros se realizó en Zotero. Los criterios de inclusión abarcaron artículos científicos publicados entre enero de 2022 y mayo de 2026, en inglés o español, relativos al análisis de funciones pedagógicas de la IAG, condiciones de uso o tensiones éticas en la educación superior o en la formación docente. Se excluyeron documentos sin arbitraje, tesis, capítulos de libro, informes institucionales y estudios que abordaran la IAG exclusivamente como herramienta de gestión administrativa.

De 773 registros brutos, la eliminación de 130 duplicados en Zotero dejó 643 únicos. El cribado por título y resumen excluyó 583 entradas (154 por tipo de documento no elegible y 429 por ausencia de relevancia temática). Los 60 registros restantes se sometieron a lectura de texto completo, y 31 fueron descartados por baja pertinencia analítica o por aportes no verificables en las categorías definidas. El corpus final quedó conformado por 29 artículos (véase Figura 1). La codificación temática siguió los principios del análisis cualitativo (Saldaña, 2021). Dos autores codificaron de manera independiente las categorías emergentes, al igual que los hallazgos centrales y las limitaciones de cada estudio. Las discrepancias se resolvieron mediante discusión razonada hasta alcanzar el consenso. Un tercer autor realizó la auditoría de consistencia interna de la distribución por categorías.

RESULTADOS

Los 29 estudios analizados evidencian un campo en expansión donde tres corresponden al bienio 2022–2023 (10.3 %), siete al año 2024 (24.1 %), trece al 2025 (44.8 %) y seis al año 2026 (20.7 %), observándose un punto de inflexión perceptible a partir de la masificación de ChatGPT. La procedencia geográfica abarca América Latina (n= 11; 37.9%), Europa (n= 7; 24.1%), Asia y Oceanía (n= 8; 27.6%) e investigaciones de alcance internacional o multicontinental (n= 3; 10.3%). La Tabla 1 resume los estudios. La Tabla 2 ofrece la comparación multidimensional de las cinco categorías emergentes.

Tabla 1.

Matriz de síntesis de los estudios seleccionados (E01–E29)

ID	Autoría / Año	País	Diseño metodológico	Muestra	Herramienta de IA	Cat.
E01	Kasneci et al. (2023)	Internacional	Revisión conceptual	N/A	LLM en general	TI
E02	Berrios Zepeda & Márquez Mora (2025)	Nicaragua	Cuasiexperimental longitudinal	111 est.	IAG	TI
E03	Latif et al. (2026)	Internacional	Revisión sistemática	86 estudios	ITS / LLM	TI
E04	Looi et al. (2026)	Singapur	Experimental	40 est. (primaria)	LLM multicapa	TI
E05	Scheepers & Stott (2026)	Sudáfrica	Estudio de diseño	Docentes en formación	LLM (Flipped Interaction)	TI
E06	Yang et al. (2024)	Taiwán	Cuasiexperimental (11 sem.)	N/E	PyTutor (ChatGPT)	TI
E07	Hilliger et al. (2026)	Chile	DBR	276 est., 10 cursos	Chatbot SRL / GAI	AA
E08	Huang et al. (2026)	Internacional	Metaanálisis de 3 niveles	85 estudios; 9 564 part.	IA de retroalimentación	AA
E09	Marimon et al. (2025)	Argentina	Cuantitativo	316 est.	ChatGPT	AA

E10	Wang et al. (2025)	Hong Kong / China	DBR	25 est.	SRLMentor (RAG + LLM)	AA
E11	Xu (2026)	China	Cuasiexperimental mixto	50 est.	LLM como par	AA
E12	Yang et al. (2026)	Taiwán / Japón	Cuantitativo	N/E	LLM + XAI	AA
E13	Alm & Watanabe (2023)	Internacional	Análisis teórico-crítico	N/A	ChatGPT (enfoque freireano)	MD
E14	Muñoz Martínez et al. (2025)	Iberoamérica	Exploratorio (panel de expertos)	N/E	IAG	MD
E15	Zúñiga Zárate (2025)	México	Mixto	20 est.	ChatGPT + blogs	MD
E16	Romeu Fontanillas et al. (2025)	Iberoamérica	Descriptivo (ed. en línea)	N/E	IAG	MD
E17	Wu et al. (2025)	China	Cuantitativo pre-post	51 est.	IA conversacional	MD
E18	Leiva-Guerrero et al. (2026)	Chile	Cualitativo-exploratorio	N/E	IAG	EF
E19	López Gil & Moreno Mosquera (2025)	Colombia	Mixto (escritura de tesis)	Posgrado	ChatGPT	EF
E20	Kashiha (2025)	Malasia	Mixto (escritura L2)	N/E	ChatGPT	EF
E21	Meyer et al. (2024)	Alemania	ECA	459 est. (EFL, secundaria)	LLM	EF
E22	Ossa & Willatt (2023)	Chile	Implementación institucional	N/E	ChatGPT-4	EF
E23	Hernández-Nodarse et al. (2025)	Ecuador	Cuantitativo	494 est.	ChatGPT	EF
E24	Gallent-Torres et al. (2023)	Iberoamérica	Revisión documental crítica	N/A	IAG	TE
E25	González-Fernández et al. (2025)	Internacional	Estado del arte	N/A	IAG	TE
E26	Jiménez León & Cisneros Chacón (2026)	México	Cualitativo multimétodo	N/E	IAG	TE
E27	Flores-Vivar & García-Peñalvo (2023)	Internacional	Revisión documental (ODS4)	N/A	IAG	TE
E28	Bozkurt (2024)	Internacional	Ensayo teórico-crítico	N/A	GenAI	TE
E29	Silvola et al. (2025)	Finlandia	Cualitativo	22 est.	IA en general	TE

Nota. Cat. = Categoría temática emergente; TI= Tutoría Inteligente; AA= Andamiaje Adaptativo; MD= Mediación Dialógica; EF= Evaluación Formativa; TE= Tensiones Éticas/Epistémicas. DBR= Design-Based Research; ECA= Ensayo Controlado Aleatorizado; EFL= English as a Foreign Language; est.= estudiantes; N/A= no aplica (estudios teóricos/conceptuales); N/E = no especificado en la fuente; part. = participantes; RAG= Retrieval-Augmented Generation; SRL= Self-Regulated Learning; XAI= Explainable AI.

Tabla 2.

Matriz comparativa multidimensional de las categorías emergentes

Categoría emergente	Función pedagógica de la IAG	Condición crítica de eficacia	Tensión teórica principal	Estado de la evidencia
Tutoría inteligente (E01-E06)	Modelado del razonamiento; retroalimentación inmediata en tareas estructuradas;	Diseño que preserve el esfuerzo cognitivo del estudiante; alternancia entre reglas explícitas y	Paradoja del tutor: mayor beneficio para estudiantes con menor base de conocimiento; mayor riesgo de	Empírica y en expansión; predominan diseños cuasiexperimentales de corto plazo; ausencia de

Andamiaje adaptativo (E07–E12)	detección de errores conceptuales Detección de metas de aprendizaje; apoyo a la autorregulación cognitiva y motivacional; feedback adaptado al ritmo del aprendiz	apertura conversacional Guía docente explícita que active la intencionalidad del estudiante; diseño instruccional orientado a la reflexión, no a la inercia	dependencia en ese mismo grupo Dependencia del andamio artificial; uso por inercia en ausencia de orientación; asimetría entre percepción de utilidad y rendimiento real	seguimientos longitudinales Más consolidada; el metaanálisis de Huang et al. (2026) provee la única estimación cuantitativa del campo (efecto moderado sobre SRL)
Mediación dialógica (E13–E17)	Estructuración conversacional que exige preguntas del estudiante; detonación de procesos de contraste y justificación	Formación explícita en uso crítico; diseño que obligue a preguntar, contrastar y justificar; integración con escritura reflexiva	Indistinguibilidad entre conversación que construye conocimiento y conversación plausible sin elaboración conceptual	Predominantemente teórica y exploratoria; pocos diseños controlados; sin criterios empíricos de operacionalización del diálogo epistémico
Evaluación formativa (E18–E23)	Retroalimentación masiva, inmediata y ajustada al texto del estudiante; modelo híbrido de tres movimientos (IA + docente + estudiante)	Supervisión docente; criterios de evaluación explícitos; retroalimentación de la IA como punto de partida, no como punto de llegada	Brecha entre calidad perceptible de la retroalimentación automática y profundidad del aprendizaje real; riesgos de integridad académica	Empírica; incluye un ECA con 459 estudiantes (Meyer et al., 2024); sin estudios de escalabilidad institucional ni de efectos a largo plazo
Tensiones éticas / epistémicas (E24–E29)	Marco normativo y de gobernanza para la integración reflexiva; categorías teóricas emergentes (cocreación, soberanía tecnológica)	Políticas institucionales que distingan usos legítimos de los fraudulentos; formación docente tecnológica, didáctica y ética articulada	Desfase estructural entre producción normativa y operacionalización empírica; estudiantes negociando solos en condiciones de ambigüedad	Predominantemente teórica y documental; la operacionalización empírica de literacidad crítica en IA y justicia epistémica está en fase inicial

Nota. Las síntesis por celda derivan del análisis cualitativo del corpus (E01–E29). Los efectos descritos son cualitativos (patrones recurrentes en la mayoría de los estudios de cada categoría) y no implican magnitudes de efecto agregadas.

Tutoría inteligente (E01–E06)

Los sistemas de tutoría inteligente (ITS) llevan décadas intentando replicar al tutor humano pero lo que cambia con los LLM es el método ya que la conversación abierta reemplaza los árboles de decisión y las respuestas predefinidas. Dos orientaciones metodológicas conviven en el campo. La primera apuesta por adoptar la apertura del LLM mediante reglas explícitas que impidan las desviaciones pedagógicas (Looi et al., 2026; Scheepers & Stott, 2026). La otra prefiere la flexibilidad conversacional y asume el riesgo de obtener una menor consistencia tutorial a cambio de mayor naturalidad dialógica (Berrios Zepeda & Márquez Mora, 2025). Latif et al. (2026), documentaron esa bifurcación donde el campo se desplaza hacia arquitecturas híbridas que no abandonan ninguno de los dos extremos.

Yang et al. (2024) nombran “paradoja del tutor” al hallazgo más inquietante de esta categoría: los estudiantes con menor base de conocimiento son quienes más mejoran con el apoyo del sistema, pero también quienes más se exponen a depender de él. La paradoja expresa, en términos empíricos, la tensión entre andamiaje y autonomía que Vygotski planteó en términos teóricos.

Andamiaje adaptativo (E07–E12)

El metaanálisis de Huang et al. (2026) provee la estimación cuantitativa más firme del campo reportando un efecto moderado sobre la autorregulación cognitiva y motivacional en 85 estudios con casi diez mil participantes. La heterogeneidad entre estudios obliga a tomar ese dato con cautela metodológica a pesar de que la dirección del efecto es consistente.

Hilliger et al. (2026) y Wang et al. (2025), trabajando con diseños de investigación basada en el diseño (DBR), demostraron que los chatbots pueden detectar metas de aprendizaje con una precisión cercana a la de evaluadores humanos. El dato central es que en ausencia de una adecuada guía docente explícita los estudiantes usan la herramienta por inercia. Por tanto, el andamio funciona cuando el diseño instruccional lo activa de forma deliberada.

Xu (2026) aporta un dato en otro plano señalando que cuando el LLM actúa como par y no como tutor, la ansiedad argumentativa disminuye. La razón es social asociada a que el temor a equivocarse se atenúa cuando el interlocutor no evalúa. Marimon et al. (2025) asomaron el mismo principio desde el ángulo de la confianza y de la percepción de utilidad de ChatGPT que pesa más que el rendimiento objetivo en la disposición de los estudiantes a usar la herramienta. La asimetría entre percepción y desempeño real constituye por sí misma un problema pedagógico de primer orden.

Mediación dialógica (E13–E17)

Con base en este componente, Mercer (2000) describió cómo el lenguaje construye conocimiento compartido entre personas. La pregunta que inaugura esta categoría es si un sistema que predice tokens estadísticamente puede producir algo análogo. El corpus no la responde con certeza. Delimita si cuándo la IAG al menos no obstaculiza la construcción de conocimiento y cuando obliga al estudiante a preguntar, contrastar y justificar, en lugar de simplemente recibir respuestas (Muñoz Martínez et al., 2025; Wu et al., 2025).

Alm & Watanabe (2023) desde la pedagogía de Freire plantean que la herramienta no garantiza nada por sí misma y que una IAG mal diseñada reproduce la pedagogía bancaria con la misma eficiencia que cualquier otro recurso pasivo.

También se ha documentado que la formación explícita en el uso crítico de la IAG marca una diferencia apreciable en los niveles de reflexión metacognitiva de los estudiantes (Romeu Fontanillas et al., 2025). Zúñiga Zárate (2025) demostró que la integración de ChatGPT con una escritura en blogs redujo la dependencia y potenció la metacognición. El problema de fondo permanece abierto evidenciando esto que no existe aún un criterio empírico claro para distinguir desde la observación externa una conversación con IA que construye conocimiento de una que solo produce respuestas plausibles.

Evaluación formativa (E18–E23)

Un patrón recurrente resalta en varios estudios de esta categoría. ChatGPT produce retroalimentación de calidad comparable a la de revisores pares y docentes en formación (López Gil & Moreno Mosquera, 2025), pero se queda por debajo de la retroalimentación de un docente experto en aquello que más importa pedagógicamente y que es el diálogo contextualizado y la adaptación a la historia de aprendizaje del estudiante (Kashiha, 2025). Meyer et al. (2024) añaden evidencia experimental en el campo del inglés como lengua extranjera (EFL), donde la retroalimentación generada por LLM produjo mejoras en la revisión de textos y la motivación; y se reportaron las emociones positivas asociadas al aprendizaje con efectos de magnitud pequeña a moderada ($d = 0.19$ para la calidad de la revisión).

Ossa & Willatt (2023) en Chile dicen que ChatGPT-4 generó retroalimentación masiva bajo supervisión docente con resultados funcionalmente aceptables. Para la retroalimentación por competencias en 494 estudiantes en Ecuador, Hernández-Nodarse et al. (2025) reportaron valoraciones positivas de ChatGPT aunque señalaron riesgos de integridad académica cuando la supervisión docente es insuficiente.

Leiva-Guerrero et al. (2026), desde un diseño cualitativo-exploratorio en Chile, mencionaron que la IAG fortaleció la retroalimentación cuando se combinó con criterios explícitos de evaluación y mediación docente activa.

El modelo que se configuró a partir de estos trabajos no postula el reemplazo del docente ni la inutilidad evaluativa de la IA. Propone si tres movimientos articulados en que la IAG produce un primer nivel de comentario útil, y luego el docente lo contextualiza; para que al final el estudiante decida qué hacer con él.

Tensiones éticas y epistémicas (E24–E29)

Cuatro problemas recorren esta categoría y ninguno cuenta con solución operacional clara. El primero es el más cuestionado y se centra en la autoría, cuando el texto lo produce en parte una máquina y los marcos legales e institucionales vigentes no saben cómo asignarla (Bozkurt, 2024).

El siguiente es la gobernanza resaltando que el debate institucional se ha concentrado en el fraude académico cuando el problema de fondo es quién controla los datos y las decisiones algorítmicas dentro de la institución (Jiménez León & Cisneros Chacón, 2026).

Se continúa con la equidad, porque las oportunidades que ofrece la IAG se distribuyen de manera asimétrica y la regulación llega tarde a los contextos de mayor vulnerabilidad educativa (Flores-Vivar & García-Peñalvo, 2023). El cuarto es la dependencia, revelando que los estudiantes construyen relaciones complejas con la tecnología; negocian su autonomía y siguen dependiendo de ella en condiciones que ningún reglamento norma todavía (Silvola et al., 2025).

La operacionalización empírica de nociones como justicia epistémica y literacidad crítica en IA o lo concerniente a la soberanía tecnológica está apenas en fase inicial. La mayor parte de la producción en esta categoría es teórica o de tipo documental (Flores-Vivar & García-Peñalvo, 2023; Gallent-Torres et al., 2023).

Esto abre un desfase entre la producción normativa y la investigación aplicada. Mientras González-Fernández et al. (2025) y Jiménez León & Cisneros Chacón (2026) proponen marcos racionales de gobernanza institucional, los estudiantes negocian solos en condiciones de una ambigüedad que ningún protocolo regula todavía (Silvola et al., 2025).

DISCUSIÓN

El conjunto del corpus reveló un campo en proceso de consolidación donde la variable determinante es la actividad cognitiva y dialógica que el diseño instruccional exige al estudiante. La presente revisión vino a rellenar el vacío analítico identificado marcado por la ausencia de marcos comparativos multidimensionales capaces de contrastar condiciones de eficacia entre categorías pedagógicas de la IAG en la educación superior entre 2022 y 2026. Los trabajos previos habían abordado la tutoría inteligente, el andamiaje adaptativo o la evaluación formativa de manera separada; pero la presente síntesis los articuló y evidencia que sus convergencias se traducen en preservar el razonamiento del aprendiz, provocar reflexión genuina, evitar la delegación del esfuerzo cognitivo; y que estas son más profundas que sus diferencias superficiales.

La brecha metodológica es común a las cinco categorías y se expresa en la concentración en diseños de corto plazo sin seguimiento posterior. Los trabajos sobre tutoría (Looi et al., 2026; Yang et al., 2024), andamiaje autorregulado (Hilliger et al., 2026; Wang et al., 2025) y evaluación formativa (Kashiha, 2025; Meyer et al., 2024) son intervenciones realmente breves. No se sabe si las mejoras observadas se consolidan en el tiempo o si emerge dependencia del andamio artificial en su ausencia. El contraste con métodos tradicionales es también escaso. La mayoría de los estudios carece de un grupo de comparación explícito lo que limita el análisis de causalidad y la validez externa de las conclusiones.

Con respecto a la distribución geográfica de la producción América Latina aporta más de un tercio de las obras (37.9%). Se refleja así un crecimiento perceptible de la investigación regional sobre IAG educativa. Sin embargo, la integración de esa producción en los debates teóricos globales es aún fragmentaria. La educación básica y media aparece de manera marginal en el corpus explorado pese a que los desafíos de equidad, brecha digital y sobre todo de diversidad lingüística son más acuciantes en esos niveles (UNESCO, 2023). El riesgo es que la evidencia generada en contextos de alta conectividad termine orientando políticas educativas globales sin arraigo en las realidades del Sur Global.

Existe también una debilidad teórica compartida. Las nociones como literacidad crítica en IA, la agencia epistémica y el análisis o aspiraciones de cocreación humano-máquina (Bozkurt, 2024; Jiménez León & Cisneros Chacón, 2026) se enuncian como aspiraciones, pero todavía carecen de los criterios de operacionalización empírica. La agenda de trabajo requerirá precisar qué funciones cognitivas no deberían delegarse en la máquina y bajo qué condiciones el acompañamiento de la IAG potencia o erosiona la autonomía intelectual del estudiante.

La principal limitación del presente estudio radica en la restricción aplicada a bases de datos indexadas en inglés y español. Lo señalado puede haber dejado fuera experiencias valiosas de otras regiones o en idiomas distintos y además la decisión de partir de un único artículo seminal pudo haber condicionado el mapeo bibliográfico inicial.

CONCLUSIONES

La evidencia sistematizada establece que el impacto de la IAG sobre el aprendizaje en la educación superior va más allá de ser intrínseco a la herramienta y depende de la calidad del diseño instruccional y de la deliberación pedagógica del docente; y también se desprende de los marcos institucionales que regulan su uso. Las cinco categorías emergentes convergen en señalar que la IAG amplifica o inhibe el pensamiento de orden superior según el tipo de procesamiento cognitivo que la situación de aprendizaje demanda, estando esto en plena consonancia con lo anticipado por la teoría de la carga cognitiva (Sweller, 1988) y la hipótesis de los niveles de procesamiento (Craik & Lockhart, 1972). Los diseños longitudinales con un grupo de comparación explícito son actualmente escasos en todas las categorías y constituyen la deuda metodológica más apremiante del campo.

Las implicaciones pedagógicas apuntalan a la formación docente como la prioridad no negociable. Los profesores que incorporan la IAG sin una preparación específica tienden a replicar rutinas habituales con soporte tecnológico distinto, neutralizando los beneficios potenciales de la herramienta. Y los marcos normativos deben distinguir con precisión los usos académicamente legítimos de la IA de los fraudulentos y actualizarse con la periodicidad que la aceleración tecnológica exige. La pregunta abierta para futuros trabajos es saber cuáles son las condiciones mínimas de mediación pedagógica que preservan el

pensamiento crítico cuando la IAG es omnipresente en la formación universitaria. La interrogante define la agenda de investigación del próximo ciclo.

REFERENCIAS

- Alm, A., & Watanabe, Y. (2023). Integrating ChatGPT in language education: A Freirean perspective. *Iranian Journal of Language Teaching Research*, 11(3), 19–30. <https://doi.org/10.30466/ijltr.2023.121404>
- Berrios Zepeda, R., & Márquez Mora, L. (2025). Agente de inteligencia artificial generativa en investigación científica: Un análisis explicativo del aprendizaje en el aula. *RIED. Revista Iberoamericana de Educación a Distancia*, 28(2), 39–55. <https://doi.org/10.5944/ried.28.2.43545>
- Bozkurt, A. (2024). GenAI et al.: Co-creation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis*, 16(1), 1–10. <https://doi.org/10.55982/openpraxis.16.1.654>
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684. [https://doi.org/10.1016/S0022-5371\(72\)80001-X](https://doi.org/10.1016/S0022-5371(72)80001-X)
- Flores-Vivar, J. M., & García-Peñalvo, F. J. (2023). Reflexiones sobre la ética, potencialidades y retos de la inteligencia artificial en el marco de la educación de calidad (ODS4). *Comunicar*, 31(74), 37–47. <https://doi.org/10.3916/C74-2023-03>
- Gallent-Torres, C., Zapata-González, A., & Ortego-Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: Una mirada desde la ética y la integridad académica. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 29(2), art. M5. <https://doi.org/10.30827/relieve.v29i2.29134>
- González-Fernández, M. O., Romero-López, M. A., Sgreccia, N. F., & Latorre Medina, M. J. (2025). Marcos normativos para una IA ética y confiable en la educación superior: Estado de la cuestión. *RIED. Revista Iberoamericana de Educación a Distancia*, 28(2), 181–208. <https://doi.org/10.5944/ried.28.2.43511>
- Hernandez-Nodarse, M., Lopez Ramos, A. R., Iniguez Apolo, L. M., Garcia Morales, J. A., & Paula Chica, M. G. (2025). Teaching experience in competency-based assessment supported by ChatGPT. *Universidad Ciencia Y Tecnología*, 29(127), 31–40. <https://doi.org/10.47460/uct.v29i127.959>
- Hilliger, I., Pérez-Sanagustín, M., González, C., Villalobos, E., & Celis, S. (2026). A GAI-based chatbot for scaffolding self-regulated learning: Insights from a design-based research approach. *Journal of Learning Analytics*, 13(1), 74–88. <https://doi.org/10.18608/jla.2026.9143>
- Huang, C.-Q., Lu, L.-N., Huang, Q., Zhang, Y.-R., Luo, M.-S., Su, Y., & Zheng, J.-X. (2026). Effects of artificial intelligence feedback on students' self-regulated learning in higher education: A three-level meta-analysis. *Educational Psychology Review*, 38, 64. <https://doi.org/10.1007/s10648-026-10166-z>
- Jiménez León, R., & Cisneros Chacón, E. J. (2026). Inteligencia artificial generativa e integridad académica en educación superior: Hacia un modelo de gobernanza institucional para la innovación educativa. *Reencuentro. Análisis de Problemas Universitarios*, 37(89), 249–298. <https://doi.org/10.24275/uamxoc-dcsh/reencuentro/202589-12>
- Kashiha, H. (2025). From algorithms to annotations: Rethinking feedback practices in academic writing through AI-human comparison. *Journal of Second Language Writing*, 70, 101254. <https://doi.org/10.1016/j.jslw.2025.101254>

- 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>
- Latif, E., Liu, V., & Zhai, X. (2026). A systematic review of intelligent and robot tutoring systems: Evolution, pedagogical design, and AI-driven classification. *Smart Learning Environments*, 13(1), 1-22. <https://doi.org/10.1186/s40561-025-00427-9>
- Leiva-Guerrero, M. V., Araya Zamorano, I., Escobar Collins, R., & Silva Castro, F. (2026). Feedback on learning with generative artificial intelligence in university students. *RIED. Revista Iberoamericana de Educación a Distancia*, 29(1), 241–267. <https://doi.org/10.5944/ried.45547>
- Looi, C. K., Liu, Z., & Sun, D. (2026). Taming the black box: Design principles for rule-integrated LLM tutoring systems in primary school mathematical problem solving. *Computers and Education: Artificial Intelligence*, 10, 100586. <https://doi.org/10.1016/j.caeai.2026.100586>
- López Gil, K., & Moreno Mosquera, E. (2025). Retroalimentación formativa en la escritura de tesis en posgrado: Comparación entre ChatGPT y revisores pares en un círculo de escritura. *Revista Virtual Universidad Católica del Norte*, 74, 123–160. <https://doi.org/10.35575/rvucn.n74a6>
- Marimon, F., Arias Valle, M. B., Coria Augusto, C. J., & Larrea Arnau, C. M. (2025). Del optimismo a la confianza: El impacto de ChatGPT en la confianza de los estudiantes en el aprendizaje asistido por IA. *RIED. Revista Iberoamericana de Educación a Distancia*, 28(2), 131–153. <https://doi.org/10.5944/ried.28.2.43238>
- Mercer, N. (2000). *Words and minds: How we use language to think together*. Routledge. <https://doi.org/10.4324/9780203464984>
- Meyer, J., Jansen, T., Schiller, R., Liebenow, L. W., Steinbach, M., Horbach, A., & Fleckenstein, J. (2024). Using LLMs to bring evidence-based feedback into the classroom: AI-generated feedback increases secondary students' text revision, motivation, and positive emotions. *Computers and Education: Artificial Intelligence*, 6, 100199. <https://doi.org/10.1016/j.caeai.2023.100199>
- Muñoz Martínez, C., Roger-Monzó, V., & Castelló-Sirvent, F. (2025). IA generativa y pensamiento crítico en la educación universitaria a distancia: Desafíos y oportunidades. *RIED. Revista Iberoamericana de Educación a Distancia*, 28(2), 233–273. <https://doi.org/10.5944/ried.28.2.43556>
- Ossa, C., & Willatt, C. (2023). Uso de inteligencia artificial generativa para retroalimentar escritura académica en procesos de formación inicial docente. *European Journal of Education and Psychology*, 16(2), 1–16. <https://doi.org/10.32457/ejep.v16i2.2412>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Gherzi, D., Guyatt, G. H., & Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Romeu Fontanillas, T., Romero Carbonell, M., Guitert Catasús, M., & Baztán Quemada, P. (2025). Desafíos de la inteligencia artificial generativa en educación superior: Fomentando su uso crítico en el estudiantado. *RIED. Revista Iberoamericana de Educación a Distancia*, 28(2), 209–231. <https://doi.org/10.5944/ried.28.2.43535>
- Saldaña, J. (2021). *The coding manual for qualitative researchers* (4.ª ed.). Sage.

- Scheepers, S. J., & Stott, A. E. (2026). Designing flipped-interaction prompts: A framework for generative AI as an intelligent tutor in higher education. *Education Sciences*, 16(4), 573. <https://doi.org/10.3390/educsci16040573>
- Silvola, A., Kajamaa, A., Merikko, J., & Muukkonen, H. (2025). AI-mediated sensemaking in higher education students' learning processes: Tensions, sensemaking practices, and AI-assigned purposes. *British Journal of Educational Technology*, 56(5), 2001–2018. <https://doi.org/10.1111/bjet.13606>
- Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- Wang, S., Luo, X., & Hew, K. F. (2025). Enhancing self-regulated learning using generative AI: Development and evaluation of SRLMentor. *IEEE Transactions on Learning Technologies*, 18, 1048–1061. <https://doi.org/10.1109/TLT.2025.3634216>
- Wu, B., He, Y. N., Song, Y., & Li, H. H. (2026). Fostering critical thinking in higher education: an intelligent dialogue-based approach empowered by conversational AI. *Interactive Learning Environments*, 34(4), 2061–2078. <https://doi.org/10.1080/10494820.2025.2538750>
- Xu, B. (2026). Mitigating interpersonal uncertainty in collaborative argumentation: Using LLMs as scaffolds for group cohesion, dialogic behavior and learning performance in CSCL. *Learning and Motivation*, 94, Article 102249. <https://doi.org/10.1016/j.lmot.2026.102249>
- Yang, A. C. M., Lin, J.-Y., Lin, C.-Y., & Ogata, H. (2024). Enhancing Python learning with PyTutor: Efficacy of a ChatGPT-based intelligent tutoring system in programming education. *Computers and Education: Artificial Intelligence*, 7, 100309. <https://doi.org/10.1016/j.caeai.2024.100309>
- Yang, C. C. Y., Li, M., & Huang, A. Y. Q. (2026). Leveraging large language models to enhance self-regulated learning in programming education with explainable AI. *Journal of Computer Assisted Learning*, 42(2), e70206. <https://doi.org/10.1002/jcal.70206>
- Zúñiga Zárate, A. G. (2025). Análisis de la integración de ChatGPT en el aprendizaje universitario con apoyo de blogs. *Comunicar*, 33(80), 60–74. <https://doi.org/10.58262/C80-2025-06>