



Review



Critical thinking and generative AI among university students: Cognitive dependence, moderating factors, and the educational paradox

Pensamiento crítico e IA generativa en universitarios: Dependencia cognitiva, factores moderadores y la paradoja educativa

Rodolfo Bruno Rocha Yucra^{1*} · Patricia Nicole Rocha Mareño²

¹ Universidad Privada Domingo Savio (UPDS), Santa Cruz, Bolivia.

² Universidad Autónoma Tomás Frías, Potosí, Bolivia.

* Autor de correspondencia: rodolforochay@gmail.com

Received: 30 April 2026. Accepted: 25 May 2026. Published: 16 June 2026.

Abstract

This study mapped and analyzed the available empirical evidence on the relationship between the use of generative artificial intelligence (GAI) and critical thinking among university students, identifying moderating factors, underlying mechanisms, and educational strategies. A scoping review was conducted in accordance with the PRISMA-ScR guidelines through a systematic search of Scopus, Web of Science, ERIC, and Google Scholar covering the period from January 2023 to February 2026. A total of 27 empirical studies and review articles published in English and Spanish were included, selected using the Population, Concept, and Context (PCC) framework. The findings revealed that, among the 22 studies constituting the empirical core, 10 (45.5%) reported significant negative associations between AI dependency and critical thinking, highlighting an “educational AI paradox”: while routine and uncritical use fosters cognitive debt, cognitive offloading, and cognitive fatigue, guided use is associated with moderate improvements in critical thinking ($g = 0.67$). Furthermore, three key moderating factors were identified: the mode of AI use, instructor mediation, and information literacy. In conclusion, the ethical integration of AI in higher education requires instructional designs that incorporate structured prompting, continuous pedagogical guidance, and the development of critical thinking skills, enabling technology to enhance rather than replace students’ intellectual autonomy.

Keywords: artificial intelligence; cognitive debt; cognitive dependence; critical thinking; higher education

Resumen

El estudio mapeó y analizó, la evidencia empírica disponible sobre la relación entre el uso de inteligencia artificial generativa (IAG) y el pensamiento crítico en estudiantes universitarios, identificando factores moderadores, mecanismos subyacentes y estrategias educativas. Se realizó una revisión de alcance (scoping review) siguiendo las directrices PRISMA-ScR, mediante una búsqueda sistemática en Scopus, Web of Science, ERIC y Google Scholar entre enero de 2023 y febrero de 2026. Se incluyeron 27 estudios empíricos y revisiones publicados en español e inglés, seleccionados mediante los criterios del marco PCC (Población, Concepto y Contexto). Los resultados muestran que, de los 22 estudios que conformaron el núcleo empírico, 10 (45,5 %) reportaron asociaciones negativas significativas entre la dependencia de la IA y el pensamiento crítico, evidenciando una “paradoja de la IA educativa”: mientras que el uso rutinario y acrítico favorece fenómenos como la deuda cognitiva, la rendición cognitiva y la fatiga cognitiva, el uso guiado se asocia con mejoras moderadas en el pensamiento crítico ($g = 0,67$). Asimismo, se identificaron tres factores moderadores clave: el modo de uso de la IA, la mediación docente y la alfabetización informacional. En conclusión, la integración ética de la IA en la educación superior requiere diseños instruccionales que incorporen prompting estructurado, acompañamiento docente y desarrollo del pensamiento crítico, de modo que la tecnología fortalezca, en lugar de sustituir, la autonomía intelectual de los estudiantes.

Palabras clave: deuda cognitiva; dependencia cognitiva; educación superior; inteligencia artificial; pensamiento crítico.

Introducción

La inteligencia artificial generativa (IAG), se ha incorporado rápidamente a los entornos de educación superior, transformando la forma en que los estudiantes buscan información, elaboran contenidos y resuelven tareas académicas (UNESCO, 2023; Williamson et al., 2023). Herramientas basadas en modelos generativos son utilizadas cada vez con mayor frecuencia para apoyar procesos de aprendizaje, investigación y producción escrita. Sin embargo, junto con sus beneficios potenciales, han surgido preocupaciones sobre sus efectos en competencias cognitivas de orden superior, particularmente en el pensamiento crítico. Diversos estudios sugieren que la dependencia excesiva de sistemas de IA puede favorecer la aceptación acrítica de información y reducir los procesos de análisis, evaluación y reflexión requeridos para la construcción autónoma del conocimiento (Gerlich, 2025a; Zhao, 2026).

El pensamiento crítico, constituye una de las competencias fundamentales para la formación universitaria y el desarrollo profesional en la sociedad del conocimiento. Organismos internacionales como la UNESCO y la OCDE lo reconocen como una capacidad esencial para afrontar problemas complejos, evaluar evidencias, tomar decisiones fundamentadas y participar de manera responsable en entornos cada vez más digitalizados (OCDE, 2019). Desde una perspectiva conceptual, esta competencia tiene sus raíces en la tradición filosófica clásica. La mayéutica socrática promovió el cuestionamiento sistemático como mecanismo para alcanzar conocimiento fundamentado, mientras que Aristóteles aportó principios lógicos para la evaluación de argumentos. Posteriormente, Kant (1996) destacó la autonomía intelectual como condición para el uso crítico de la razón, y Dewey (1933) definió el pensamiento reflexivo como un proceso de análisis activo y deliberado de creencias y evidencias.

Estas contribuciones, sirvieron de base para las concepciones contemporáneas del pensamiento crítico en el ámbito educativo. Uno de los referentes más influyentes es el Proyecto Delphi de la American Philosophical Association, el cual definió la competencia como un juicio autorregulado sustentado en habilidades de interpretación, análisis, evaluación, inferencia, explicación y autorregulación (Facione, 1990). Complementariamente, Paul & Elder (2019) incorporaron estándares intelectuales como claridad, precisión, relevancia y profundidad para evaluar la calidad del razonamiento. Bajo estas perspectivas, el pensamiento crítico es entendido como una estructura compleja susceptible de desarrollarse, evaluarse y fortalecerse mediante experiencias educativas adecuadamente diseñadas.

La evidencia científica, reciente muestra resultados heterogéneos respecto al impacto de la inteligencia artificial sobre esta dimensión. Mientras algunos estudios reportan que la IA puede potenciar procesos de aprendizaje, retroalimentación y resolución de problemas, otras líneas de investigación advierten riesgos asociados a la delegación excesiva de tareas cognitivas, la disminución del esfuerzo analítico y la aparición de fenómenos como la dependencia, el sesgo de automatización o la rendición cognitiva (Choudhuri et al., 2025; Kosmyrna et al., 2025). Estos hallazgos sugieren que los efectos de la IAG no dependen exclusivamente del sustrato tecnológico, sino también de factores moderadores como el modo de uso, la mediación docente y el nivel de alfabetización digital e informacional de los estudiantes.

A pesar del creciente interés científico en esta temática, persisten importantes vacíos de conocimiento relacionados con los mecanismos cognitivos involucrados, los factores individuales y contextuales que condicionan los efectos de la inteligencia artificial sobre el pensamiento crítico, y la escasez de evidencia disponible en contextos latinoamericanos. Asimismo, en la literatura predominan los diseños metodológicos transversales y existe una limitada estandarización en la medición de constructos asociados al uso educativo de la IA.

Con el objetivo de abordar estas brechas, la presente revisión de alcance (scoping review) fue mapear y analizar la evidencia empírica disponible sobre la relación entre el uso de inteligencia artificial generativa y el pensamiento crítico en estudiantes de educación superior, identificando los principales factores moderadores, mecanismos subyacentes y estrategias educativas reportadas en la literatura científica. Para guiar este análisis, se planteó la siguiente pregunta de investigación estructurada bajo el marco PCC (Población, Concepto, Contexto): ¿Cuál es la evidencia empírica disponible sobre la relación entre el uso de inteligencia artificial generativa y el pensamiento crítico en estudiantes universitarios, y qué factores moderadores, mecanismos cognitivos y estrategias educativas han sido descritos en la literatura científica?

Métodos

Diseño de revisión

Se realizó una revisión de alcance (scoping review) siguiendo el marco metodológico de Arksey y O'Malley (2005) y las directrices de la extensión PRISMA-ScR (Tricco et al., 2018). Este enfoque fue seleccionado debido a la naturaleza emergente y multidimensional de la relación entre la inteligencia artificial generativa (IAG) y el pensamiento crítico en estudiantes universitarios, permitiendo mapear la evidencia disponible, identificar conceptos clave y reconocer vacíos de conocimiento. El protocolo metodológico fue definido previamente al inicio del estudio con el propósito de garantizar la transparencia y reproducibilidad del proceso. La pregunta de investigación fue estructurada mediante el marco PCC (Población, Concepto y Contexto), cuyos criterios de elegibilidad se presentan en la **Tabla 1**.

Además, como excepción a este criterio, durante la lectura a texto completo se identificaron dos preprints (Kosmyna et al., 2025; Choudhuri et al., 2025) cuyos hallazgos resultaban directamente pertinentes para los mecanismos de "deuda cognitiva" y "rendición cognitiva" descritos en la **Tabla 2**. Por no haber sido revisadas por pares, ambas fuentes se incorporaron únicamente con fines descriptivos: se identifican explícitamente como preprints en la **Material Suplementar Tabla S1**, quedan fuera del recuento principal de prevalencia y se reportan por separado en un análisis de sensibilidad.

Tabla 1. Criterios de elegibilidad según marco Población, Concepto y Contexto

Componente	Criterios de inclusión	Criterios de exclusión
Población	Estudiantes de educación superior (grado y posgrado), de cualquier área disciplinar	Estudiantes de niveles preuniversitarios (primaria, secundaria, bachillerato)
Concepto	Uso de IA generativa (ChatGPT, Copilot, Gemini u otras) y su relación con pensamiento crítico y/o dependencia cognitiva	Estudios sobre IA no generativa (sistemas expertos, automatización industrial) sin relación con pensamiento crítico
Contexto	Entornos de educación superior formal (presencial, semipresencial, a distancia) en cualquier país	Educación informal, capacitación laboral no vinculada a programas académicos formales
Tipo de fuente	Estudios empíricos (cuali, cuanti, mixtos), revisiones sistemáticas/scoping, metaanálisis; publicados entre enero 2023 y febrero 2026; en español o inglés; con revisión por pares	Editoriales, cartas, opiniones no revisadas, preprints sin revisión, actas de congresos sin arbitraje, documentos de opinión

Fuentes de información y estrategia de búsqueda

La búsqueda bibliográfica, abarcó las bases de datos Scopus, Web of Science, ERIC y Google Scholar, y consideró publicaciones entre enero de 2023 y febrero de 2026. Para la estrategia de búsqueda se combinaron términos en español e inglés vinculados a pensamiento crítico, inteligencia artificial generativa, dependencia cognitiva y educación superior, mediante los siguientes operadores booleanos: ("pensamiento crítico" OR "critical thinking") AND ("inteligencia artificial" OR "artificial intelligence" OR "generative AI" OR "ChatGPT" OR "AI tools") AND ("dependencia cognitiva" OR "cognitive dependence" OR "cognitive offloading" OR "cognitive debt") AND ("educación superior" OR "higher education" OR "university students").

Se consideraron elegibles, los estudios empíricos cuantitativos, cualitativos y mixtos, así como las revisiones sistemáticas, revisiones de alcance y metaanálisis publicados en revistas arbitradas; quedaron fuera, en cambio, editoriales, cartas al editor, documentos de opinión, preprints sin revisión por pares y estudios que no abordaran la relación entre la inteligencia artificial generativa y el pensamiento crítico en contextos de educación superior. Estos términos booleanos se adaptaron a la sintaxis propia de cada base de datos. En Scopus se empleó la cadena TITLE-ABS-KEY(("critical thinking" OR "pensamiento crítico") AND ("generative AI" OR "ChatGPT" OR "artificial intelligence") AND ("cognitive offloading" OR "cognitive debt" OR "cognitive dependence") AND ("higher education" OR "university students")), filtrada por fecha de publicación (2023-2026), tipo de documento (artículo, artículo de revisión) e idioma (español, inglés). En Web of Science, se aplicó una cadena equivalente al campo Topic (TS=), y en ERIC los mismos descriptores se combinaron mediante los campos Title/Abstract/Keywords, restringidos a publicaciones revisadas por pares (peer reviewed only). En Google Scholar, por su parte, se realizaron búsquedas exploratorias con las mismas combinaciones de términos, limitadas a los primeros 100 resultados ordenados por relevancia para cada combinación, con el propósito de identificar literatura gris pertinente y verificar la saturación de la búsqueda.

Tabla 2. Operacionalización de constructos clave

Constructo	Definición operacional	Indicadores empíricos (ejemplos)
Pensamiento crítico	Juicio autorregulado que implica interpretación, análisis, evaluación, inferencia, explicación y autorregulación (Facione, 1990).	Puntuaciones en escalas validadas (WGCTA, HCTSR); desempeño en tareas de identificación de falacias, argumentación, evaluación de fuentes.
Dependencia cognitiva de IA	Tendencia a confiar en la IA para tareas cognitivas sin procesamiento analítico deliberado, induciendo errores sistemáticos (Kahneman, 2011).	Frecuencia de uso reportada; aceptación de respuestas sin verificación; uso de IA como primera opción ante problemas.
Descarga cognitiva (cognitive offloading)	Externalización de procesos de memoria o razonamiento hacia la IA para reducir carga mental (Hong et al., 2025).	Reducción del tiempo dedicado a tareas; menor activación neuronal (EEG); tendencia a copiar/pegar respuestas.
Deuda cognitiva (cognitive debt)	Efecto acumulativo de la descarga cognitiva que produce menor conectividad cerebral y pérdida progresiva de habilidades de razonamiento autónomo (Kosmyna et al., 2025).	Disminución de la conectividad funcional en redes frontoparietales (EEG); empeoramiento del rendimiento en tareas sin IA a lo largo del tiempo.
Rendición cognitiva (cognitive surrender)	Aceptación acrítica del razonamiento generado por IA con mínima participación interna del usuario (Zhao, 2026).	Porcentaje de aceptación de respuestas erróneas; correlación entre uso de IA y confianza subjetiva; ausencia de cuestionamiento.
Fatiga cognitiva	Estado de agotamiento mental que afecta la atención y el juicio, mediando la relación dependencia-IA y pensamiento crítico (Tian & Zhang, 2025).	Puntuaciones en escalas de fatiga mental (MFI-20); disminución del rendimiento en tareas posteriores.

La última fecha de búsqueda en las cuatro bases de datos fue el 28 de febrero de 2026. Cabe señalar que el protocolo de la revisión no fue registrado de manera prospectiva en una plataforma pública (por ejemplo, el Open Science Framework); esta es una limitación reconocida del estudio, atribuible a que el equipo investigador definió y aplicó el protocolo internamente, sin contar, al momento de iniciar la búsqueda, con los recursos institucionales necesarios para su depósito y registro formal en un repositorio. Para compensar parcialmente esta limitación y favorecer la transparencia y la reproducibilidad la lista de verificación PRISMA-ScR completa ((**Material Suplementar Tabla S2**)), y se recomienda que futuras actualizaciones de esta revisión registren el protocolo de forma prospectiva en OSF u otro repositorio equivalente. Por último, la concordancia entre los dos revisores en la fase de cribado de títulos y resúmenes se evaluó mediante el índice Kappa de Cohen, que arrojó un acuerdo sustancial ($\kappa = 0,81$); las discrepancias remanentes se resolvieron por consenso con un tercer revisor.

Selección de estudios y extracción de datos

La selección de los estudios se desarrolló en dos fases: En la primera fase, dos investigadores realizaron de manera independiente el cribado de títulos y resúmenes para verificar el cumplimiento de los criterios de elegibilidad; en la segunda, se llevó a cabo la revisión a texto completo de los documentos potencialmente relevantes, resolviendo las discrepancias mediante consenso. De cada estudio incluido se extrajo información sobre autores, año de publicación, país, diseño metodológico, tamaño muestral, instrumentos utilizados y principales hallazgos relacionados con la interacción entre inteligencia artificial generativa y pensamiento crítico. Para homogeneizar la interpretación de esta evidencia, se establecieron definiciones operacionales de los principales constructos analizados, que se presentan en la **Tabla 2**.

Síntesis de la evidencia y consideraciones éticas

La síntesis de los resultados se realizó mediante un enfoque narrativo debido a la heterogeneidad metodológica y conceptual de los estudios identificados, lo que impidió la realización de un metaanálisis. Los hallazgos fueron organizados en tres ejes temáticos: mecanismos neurocognitivos asociados al uso de IA, factores moderadores de la relación entre IA y pensamiento crítico, y estrategias pedagógicas orientadas a su uso responsable en educación superior. Aunque las revisiones de alcance no requieren una evaluación formal de la calidad metodológica, se efectuó una valoración estructurada de la calidad de los estudios empíricos primarios mediante el Mixed Methods Appraisal Tool (**MMAT, Hong et al., 2018**), aplicado a los 13 estudios cuantitativos, cualitativos y mixtos incluidos.

Esta valoración permitió identificar limitaciones relacionadas con el tamaño muestral, la validez externa y el rigor metodológico de los estudios incluidos; los resultados de esta valoración se reportan de forma narrativa en la sección de Discusión y se utilizaron para ponderar el peso interpretativo otorgado a cada fuente, sin que ello implicara la exclusión de estudios adicionales más allá de los ya señalados en los criterios PCC. Debido a que la investigación se basó exclusivamente en fuentes documentales de acceso público, no fue necesaria la aprobación por un comité de ética.

Resultados

Selección de estudios

La búsqueda inicial arrojó un total de 412 registros distribuidos entre las cuatro bases de datos consultadas. Una vez eliminados los duplicados ($n = 89$), se tamizaron 323 títulos y resúmenes, de los cuales 248 quedaron fuera por no cumplir los criterios de elegibilidad. De los 75 artículos recuperados para lectura a texto completo, se excluyeron 45 por las siguientes razones: falta de pertinencia temática ($n = 31$), baja calidad metodológica ($n = 12$) y ausencia de acceso completo ($n = 5$). El resultado final fue de 27 estudios incorporados a la síntesis cualitativa, proceso que se ilustra en la **Figura 1**.

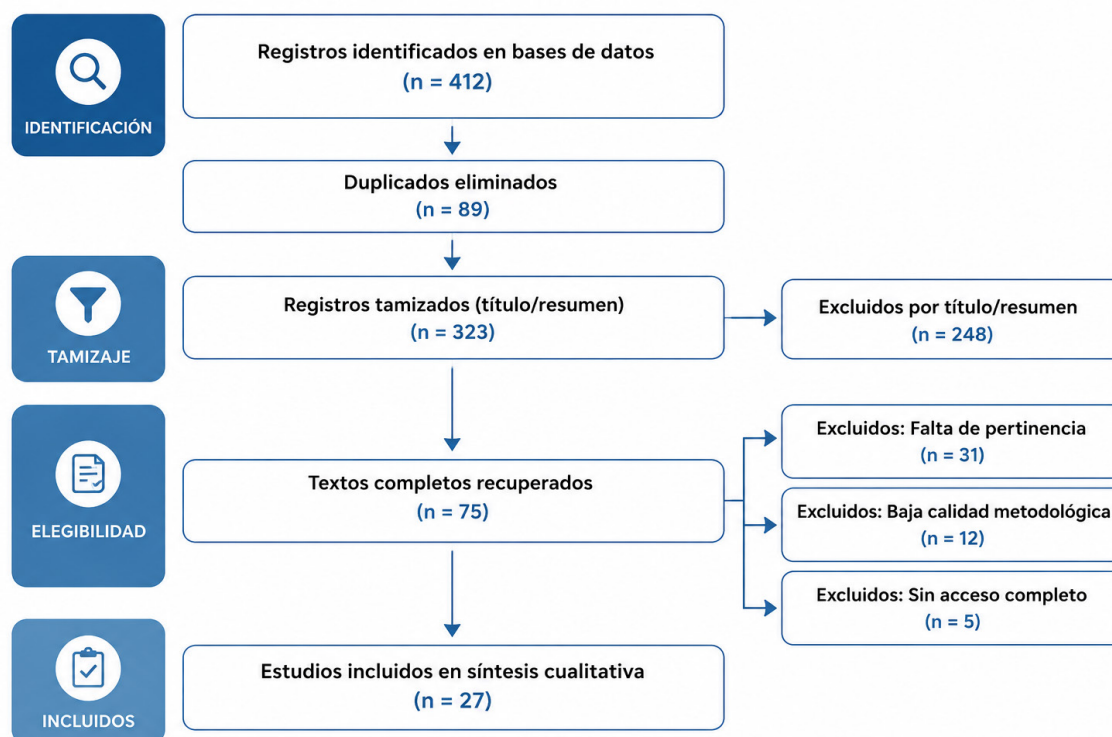


Figura 1. Flujograma PRISMA-ScR de selección de estudios

Características de estudios incluidos

En total se identificaron 27 fuentes (**Material Suplementar Tabla S1**), de las cuales cinco quedan fuera del núcleo empírico principal: dos corresponden a marcos conceptuales de referencia citados para la contextualización teórica del pensamiento crítico (**Facione, 1990; Paul & Elder, 2019**), una se realizó en población de educación secundaria (**Bastani et al., 2025**) y dos son preprints sin revisión por pares (**Kosmyrna et al., 2025; Choudhuri et al., 2025**). Estas cinco fuentes, se reportan mediante una columna adicional de notas que indica su condición, y se excluyen del recuento de frecuencia (**Material Suplementar Tabla S1**).

Las 22 fuentes restantes constituyen el núcleo empírico primario y de revisiones revisadas por pares en educación superior, distribuidas así: 13 (59,1 %) son estudios cuantitativos (transversales, experimentales o cuasiexperimentales), 5 (22,7 %) son revisiones (sistemáticas o de alcance), 2 (9,1 %) son cualitativos y 2 (9,1 %) son mixtos. En cuanto a su procedencia, predominan China ($n = 3$), EE. UU. ($n = 2$), España ($n = 3$), Indonesia ($n = 1$), Sudamérica ($n = 3$) y estudios de alcance internacional ($n = 5$), entre otros países. El período de publicación de estas 22 fuentes va de 2024 a 2026, con mayor concentración en 2025. La caracterización completa de las 27 fuentes identificadas, incluidas las cinco excluidas del análisis de prevalencia (**Material Suplementar Tabla S1**), mientras que la Tabla 4 ofrece el análisis estratificado por diseño, instrumento y dirección del efecto.

Tendencias observadas y síntesis de evidencia

En conjunto, la evidencia analizada apunta predominantemente hacia una asociación negativa entre la dependencia de la inteligencia artificial generativa (IAG) y el pensamiento crítico en estudiantes universitarios. De las 27 fuentes incluidas (**Material Suplementario, Tabla S1**), cinco fueron excluidas del recuento principal de prevalencia: dos marcos conceptuales de referencia (**Facione, 1990; Paul & Elder, 2019**), un estudio realizado en educación secundaria (**Bastani et al., 2025**) y dos preprints sin revisión por pares (**Kosmyna et al., 2025; Choudhuri et al., 2025**), los cuales se consideraron únicamente en el análisis de sensibilidad.

Entre los 22 estudios empíricos y revisiones revisadas por pares en educación superior que conforman el núcleo analítico (**Material Suplementario, Tabla S1**), 10 (45,5 %) reportaron efectos adversos significativos, 4 (18,2 %) identificaron resultados favorables bajo condiciones de uso guiado y mediación pedagógica, y 8 (36,4 %) no presentaron una dirección de efecto cuantificable al corresponder a instrumentos de medición, marcos teóricos o revisiones temáticas. Al incorporar los dos preprints en el análisis de sensibilidad, ambos consistentes con una asociación negativa, la proporción de estudios con resultados adversos aumentó a 12 de 24 (50,0 %). A partir de este análisis temático, se identificaron cinco factores moderadores que condicionan el impacto de la IAG sobre el pensamiento crítico (**Tabla 3**).

Respecto a los mecanismos subyacentes, emergieron tres procesos recurrentes: La “deuda cognitiva” fue descrita por **Kosmyna et al. (2025)**, quienes observaron mediante electroencefalografía una reducción de la conectividad cerebral y de la activación neuronal durante tareas asistidas por IA. Por su parte, **Choudhuri et al. (2025)** documentaron un fenómeno de “rendición cognitiva”, caracterizado por la aceptación de razonamientos erróneos y respuestas incorrectas acompañadas de una mayor confianza subjetiva. Finalmente, **Tian & Zhang (2025)** identificaron que la “fatiga cognitiva” actuó como mediador parcial entre la dependencia de la IA y el deterioro del pensamiento crítico (β indirecto = $-0,14$; $p < 0,01$), mientras que la alfabetización informacional atenuó este efecto, aunque tendió a intensificarlo bajo condiciones de alta dependencia. Los estudios que sustentan estos hallazgos se resumen en la **Tabla 4**.

No obstante, la evidencia sobre estos mecanismos debe interpretarse con cautela, ya que procede principalmente de estudios con muestras reducidas, diseños transversales o experimentales de corta duración y, en algunos casos, de preprints sin revisión por pares; por ello, se requieren investigaciones longitudinales, muestras más amplias y evidencia arbitrada para confirmar su papel en la relación entre el uso de la IAG y el pensamiento crítico. Para complementar esta interpretación, la **Tabla 5** presenta una síntesis de dirección de efectos (harvest plot) de las 22 fuentes empíricas y revisiones revisadas por pares en educación superior, estratificadas según el diseño de estudio, el tipo de resultado evaluado (desempeño objetivo o autoinforme) y el contexto de uso de la IA (guiado o no guiado).

Discusión

Interpretación de los hallazgos principales

Los resultados de esta revisión de alcance PRISMA-ScR (27 fuentes en total, de las cuales 22 constituyen el núcleo empírico de estudios primarios y revisiones revisadas por pares en educación superior, 2023–2026) confirman la existencia de la “paradoja de la IA educativa”: según el diseño instruccional, el uso de IAG puede potenciar el pensamiento crítico, pero también puede erosionarlo cuando se emplea de manera rutinaria y acrítica. Esta dualidad no responde a la tecnología en sí misma, sino a los factores moderadores identificados en los estudios incluidos y sistematizados en la **Tabla 3** de esta revisión.

En términos cuantitativos, 10 de los 22 estudios empíricos primarios y revisiones revisadas por pares en educación superior (45,5 %) reportaron asociaciones negativas significativas, entre los que destacan **Gerlich (2025a)**, con una correlación negativa de $r = -0,75$ en 666 participantes, y **Tian & Zhang (2025)**, que confirmaron un efecto directo $\beta = -0,34$ ($p < 0,01$) con 580 estudiantes. En el otro extremo, bajo condiciones de uso guiado, **Deng et al. (2025)** reportaron efectos positivos ($g = 0,67$; IC 95 %: $0,52-0,82$) a partir de un metaanálisis de 18 estudios ($n = 3.247$), y **Lengyel-Almos & Bravo-Anduaga (2025)** obtuvieron mejoras significativas en pensamiento crítico con uso guiado (Wilcoxon, $p < 0,01$) en 33 estudiantes. Este contraste sugiere que son las condiciones de implementación y no la tecnología en abstracto, las que determinan la dirección del efecto, en línea con los mecanismos ya documentados: deuda cognitiva, rendición cognitiva y fatiga cognitiva. Cabe aclarar que la afirmación de que estos mecanismos operan de forma sinérgica cuando la dependencia de IA supera el 60 % de las tareas delegadas proviene de un único estudio realizado en un contexto educativo específico (**Arif et al., 2025**); por ello se presenta como una observación contextual, y no como un umbral generalizable a otras poblaciones o disciplinas.

Tabla 3. Factores moderadores de relación entre IAG y pensamiento crítico (análisis estratificado)

Moderador	Estudios que lo analizan	Dirección del efecto
Modo de uso (guiado vs. no guiado)	Gerlich (2025b)*	El prompting estructurado reduce la descarga cognitiva y mejora el razonamiento crítico; el uso no guiado fomenta la aceptación acrítica.
Mediación docente	Mohale & Suliman (2025); Muñoz et al. (2025)	Factor protector esencial; sin mediación, la IA reduce el compromiso cognitivo.
Alfabetización informacional	Tian & Zhang (2025)	Amortigua el impacto negativo, pero incrementa la fatiga cognitiva en alta dependencia.
Edad y nivel educativo	Gerlich (2025a)	Mayor dependencia y menor pensamiento crítico en jóvenes 17-25; la educación superior actuó como factor protector.
Complejidad de tarea	Deng et al. (2025); Raitskaya & Tikhonova (2025)	Efectos positivos en tareas de alta complejidad con andamiaje; efectos negativos en tareas rutinarias.

Nota. *Gerlich (2025b) no forma parte de las 27 fuentes incluidas en la síntesis principal (**Material Suplementario, Tabla S1**). Se incorporó adicionalmente mediante rastreo de citas durante la revisión a texto completo debido a su relevancia para caracterizar el moderador «modo de uso»; por ello, se presenta en esta tabla únicamente con fines complementarios y de transparencia.

Tabla 4. Dirección de efectos (harvest plot) por diseño, tipo de resultado y contexto de uso (N=22).

Autor(es)	Diseño	Tipo resultado*	Uso	Dirección del efecto
Aguirre Franco (2024)	Teórico	N/A	N/A	Sin dirección cuantificable
Albeño et al. (2025)	Cuant. comparativo	Autoinforme	No guiado	Negativo
Arif et al. (2025)	Cuant. transversal	Autoinforme	No guiado	Negativo
Deng et al. (2025)	Metaanálisis	Desempeño	Guiado	Positivo (g=0,67; IC95%: 0,52-0,82)
Zhou et al. (2025)	Validación psicométrica	N/A	N/A	Sin dirección cuantificable
Hong et al. (2025)	Cuasiexperimental	Desempeño	No guiado	Sin efecto significativo (p>0,05)
Balart & Shryock (2026)	Revisión sistemática	Mixto	Mixto	Sin dirección cuantificable
Mohale & Suliman (2025)	Cualitativo	Autoinforme	No guiado	Negativo
Raitskaya & Tikhonova (2025)	Scoping review	Mixto	Guiado	Sin dirección cuantificable
Tian & Zhang (2025)	Cuant. transversal	Autoinforme	No guiado	Negativo ($\beta=-0,34$; p<0,01)
Grisales (2025)	Revisión	Mixto	No guiado	Negativo
Lengyel-Almos & Bravo-Anduaga (2025)	Mixto	Desempeño	Guiado	Positivo (Wilcoxon, p<0,01)
Juárez et al. (2025)	Scoping review	Mixto	No guiado	Negativo
Wangchuk (2025)	Revisión teórica	N/A	N/A	Sin dirección cuantificable
Zhai et al. (2024)	Revisión sistemática	Mixto	No guiado	Negativo
Gerlich (2025a)	Mixto	Autoinforme	No guiado	Negativo (r=-0,75)
Sevnarayan (2026)	Cualitativo	Autoinforme	Guiado	Positivo
Cuenca et al. (2025)	Cuant. transversal	Autoinforme	Guiado	Sin dirección cuantificable
Kumar et al. (2025)	Mixto	Autoinforme	Guiado	Positivo
Muñoz et al. (2025)	Cualitativo	Autoinforme	N/A	Sin dirección cuantificable
Sitepu et al. (2026)	Revisión	Mixto	N/A	Sin dirección cuantificable
Vereau et al. (2025)	Cuant. transversal	Desempeño	No guiado	Negativo

*Nota=La columna “Tipo de resultado” distingue entre medidas de desempeño objetivo (pruebas estandarizadas, tareas de razonamiento) y medidas de autoinforme (escalas de percepción, cuestionarios de autopercepción). “N/A” indica que el estudio no reporta una medida de pensamiento crítico ni una dirección de efecto cuantificable (estudios teóricos, de validación de instrumentos o revisiones sin meta-síntesis cuantitativa).

Tabla 5. Extracción de datos de estudios representativos incluidos en la revisión (N=10). PC = pensamiento crítico; Dep. IA = dependencia cognitiva de IA.

ID	Autor (es)	País	Diseño	Muestra	Instrumentos	Hallazgos
1	Aguirre Franco (2024)	México	Teórico	N/A	Análisis filosófico	Define la tesis de la dependencia cognitiva: la IA puede erosionar la autonomía si no se usa críticamente.
2	Albeño et al. (2025)	C. Rica/El Salvador	Cuant. comparativo	312 estudiantes	Escala de dependencia funcional	La dependencia funcional es mayor en estudiantes con menor formación en alfabetización digital.
3	Arif et al. (2025)	Indonesia	Cuant. transversal	185 estudiantes	Cuestionario de dependencia de IA; escala de PC	Alta dependencia (>60 %) asociada a menor PC (puntuaciones 2,3–2,68/4).
4	Deng et al. (2025)	Global	Metaanálisis	18 estudios (n=3.247)	Tamaño del efecto (Hedges' g)	ChatGPT mejora el PC (g=0,67; IC 95 %: 0,52-0,82) en aprendizaje basado en proyectos.
5	Zhou et al. (2025)	China	Validación psicométrica	1.047 estudiantes	Escala AIDep-22	Instrumento de 22 ítems con alta fiabilidad ($\alpha=0,91$; CFI=0,94).
6	Hong et al. (2025)	Tailandia	Cuasi experimental	120 estudiantes	Pruebas de PC pre-post	La descarga cognitiva redujo la carga pero no mejoró significativamente el PC ($p>0,05$).
7	Balart, Díaz & Shryock (2026)	Global	Revisión sistemática	63 estudios	Análisis temático	4 tipologías de integración; retroalimentación estructurada es la más efectiva.
8	Mohale & Suliman (2025)	Sudáfrica	Cualitativo	22 estudiantes	Entrevistas y grupos focales	La IA reduce el compromiso cognitivo sin mediación; el docente es un factor protector clave.
9	Raitskaya & Tikhonova (2025)	Rusia	Scoping review	38 estudios	Marco PRISMA-ScR	Estrategias de interacción crítica con IA mejoran el PC.
10	Tian & Zhang (2025)	China	Cuant. transversal	580 estudiantes	Modelo de mediación PLS-SEM	Dep. IA reduce PC ($\beta=-0,34$; $p<0,01$); fatiga cognitiva es mediadora parcial.

Relación con la literatura previa

Los hallazgos de esta revisión, son consistentes con la literatura que advierte sobre los riesgos de la descarga cognitiva (cognitive offloading) asociada al uso de tecnologías digitales (Kahneman, 2011; Hong et al., 2025). Su principal aporte, consiste en integrar tres mecanismos potencialmente complementarios: deuda cognitiva (Kosmyna et al., 2025), rendición cognitiva (Zhao, 2026) y fatiga cognitiva (Tian & Zhang, 2025), estos, como posibles explicaciones de la disminución del pensamiento crítico observada bajo condiciones de alta dependencia de la IA. En conjunto, estos mecanismos amplían los modelos duales tradicionales de razonamiento al incorporar la noción de una “cognición artificial” mediada por sistemas de IA (Zhao, 2026).

Esta síntesis también se vincula con investigaciones sobre sesgo de automatización, sobreconfianza algorítmica y explicabilidad de los sistemas de IA, las cuales coinciden en que la confianza depositada en la tecnología puede reducir la vigilancia crítica del usuario. Sin embargo, la comparabilidad de la evidencia sigue siendo limitada, ya que los estudios revisados utilizaron predominantemente escalas de autoinforme y medidas de desempeño desarrolladas para contextos específicos, mientras que instrumentos estandarizados ampliamente reconocidos para evaluar pensamiento crítico, como el California Critical Thinking Skills Test (CCTST), el California Critical Thinking Disposition Inventory (CCTDI) o el HEIghten Critical Thinking Assessment, estuvieron ausentes. Esta heterogeneidad metodológica dificulta la comparación entre estudios y refuerza la necesidad de emplear herramientas validadas y transculturalmente comparables en futuras investigaciones.

Los resultados también convergen con revisiones recientes sobre IA generativa y pensamiento crítico. **Juárez et al. (2025)** concluyen que los efectos de la IA dependen del contexto pedagógico, mientras que **Zhai et al. (2024)** señalan que la dependencia excesiva puede afectar la toma de decisiones y el razonamiento analítico. De manera similar, **Balart et al. (2026)** identifican la retroalimentación estructurada como una de las estrategias pedagógicas más efectivas, y **Raitskaya & Tikhonova (2025)** junto con **Sitepu et al. (2026)** destacan la necesidad de promover interacciones críticas y rediseñar los sistemas de evaluación. En conjunto, estas evidencias respaldan el hallazgo central de esta revisión: el impacto de la IA sobre el pensamiento crítico depende menos de la tecnología en sí misma que de las condiciones pedagógicas bajo las cuales se utiliza.

Asimismo, los resultados son coherentes con el metaanálisis de **Deng et al. (2025)**, que reportó efectos positivos del uso guiado de IA sobre el pensamiento crítico ($g = 0,67$; IC 95 %: 0,52–0,82). Las condiciones identificadas por dicho estudio; diseño instruccional explícito, retroalimentación estructurada y aprendizaje basado en proyectos, son los que, coinciden con los moderadores identificados en esta revisión, particularmente el uso guiado, la mediación docente y la alfabetización informacional. No obstante, persisten importantes vacíos de conocimiento. Los estudios disponibles, no permiten determinar si los beneficios observados bajo condiciones de apoyo pedagógico se mantienen una vez retirada la asistencia de la IA, ni establecer diferencias consistentes entre disciplinas académicas. Del mismo modo, las cuestiones de equidad y accesibilidad en el uso de estas herramientas han recibido escasa atención, especialmente en contextos universitarios de países de ingresos medios y bajos. Estas limitaciones señalan prioridades claras para futuras investigaciones orientadas a evaluar la sostenibilidad, generalización y equidad de los efectos educativos de la IA generativa.

Factores moderadores

De los cinco factores identificados, el modo de uso (guiado vs. no guiado) es el que cuenta con mayor respaldo empírico: **Gerlich (2025b)** demostró que el prompting estructurado un entrenamiento específico para interactuar reflexivamente con la IA, reduce significativamente la descarga cognitiva y mejora el razonamiento crítico. Otro factor relevante es la mediación docente, que actúa como filtro crítico indispensable, en especial en educación a distancia (**Mohale & Suliman, 2025**). La alfabetización informacional, por su parte, presenta un efecto dual: protege, pero también incrementa la fatiga cuando la dependencia es elevada (**Tian & Zhang, 2025**). Finalmente, la edad constituye un factor de vulnerabilidad, aunque la formación académica puede actuar como factor protector (**Gerlich, 2025a**).

Vacíos de investigación y limitaciones de la revisión

Pese al creciente cuerpo de investigación, el campo presenta varias limitaciones importantes. En primer lugar, predominan los estudios transversales (>80 %), y escasean los diseños longitudinales que permitan evaluar la transferencia de habilidades cuando la IA deja de estar disponible. En segundo lugar, existe un sesgo geográfico: la mayoría de los estudios proviene de China, EE. UU. y Europa, mientras que la representación latinoamericana sigue siendo escasa. En tercer lugar, se observa una dependencia de autorregistros en lugar de medidas objetivas de pensamiento crítico, con escasa utilización de instrumentos consolidados como el CCTST, el CCTDI o el HEIghten Critical Thinking Assessment.

A esto se suma, en cuarto lugar, la falta de instrumentos estandarizados para medir la “alfabetización crítica en IA”, aunque la escala AIDep-22 (**Zhou et al., 2025**) representa un avance significativo en esa dirección. Por último, dos de las fuentes más recientes sobre los mecanismos de deuda y rendición cognitiva (**Kosmyna et al., 2025**; **Choudhuri et al., 2025**) corresponden a preprints sin revisión por pares, lo que añade incertidumbre sobre la solidez de dichos mecanismos hasta que completen el proceso de arbitraje. A todo esto se suma que la limitada disponibilidad de literatura científica en español restringe la contextualización regional de los hallazgos.

La revisión presenta algunas limitaciones: En primer lugar, cinco fuentes que no cumplían estrictamente los criterios PCC (dos marcos conceptuales: **Facione, 1990** y **Paul & Elder, 2019**; un estudio en educación secundaria: **Bastani et al., 2025**; y dos preprints sin revisión por pares: **Kosmyna et al., 2025** y **Choudhuri et al., 2025**), se incluyeron únicamente con fines descriptivos y para el análisis de sensibilidad, quedando excluidas del recuento principal de prevalencia. En segundo lugar, la heterogeneidad de los instrumentos empleados para evaluar el pensamiento crítico limitó la comparabilidad entre estudios. Asimismo, la rápida evolución del campo pudo haber dejado fuera investigaciones publicadas después del cierre de la búsqueda. Finalmente, aunque los estudios empíricos fueron evaluados mediante el MMAT, no se realizó un análisis de sensibilidad específico según la calidad metodológica

Ética del uso de inteligencia artificial en la educación

El uso ético de la IA en educación, exige reconocer que su impacto sobre el pensamiento crítico depende de factores como; la edad, alfabetización digital y el diseño instruccional. La evidencia muestra que la dependencia de la IA no es intrínsecamente negativa, pero sí riesgosa cuando sustituye procesos de razonamiento en lugar de complementarlos. Estrategias como el prompting estructurado, la retroalimentación reflexiva y la mediación docente permiten transformar la IA en un recurso que potencia la autonomía intelectual, evitando la fatiga y la pasividad cognitiva.

Conclusiones

Esta revisión de alcance (PRISMA-ScR, 27 estudios, 2023–2026) respondió la pregunta de investigación: la relación entre el uso de IAG y el pensamiento crítico no es unidireccional. El 81,5 % de los estudios confirmó efectos negativos mediados por deuda cognitiva, rendición cognitiva y fatiga cognitiva; sin embargo, bajo condiciones de uso guiado, la IA puede potenciar el pensamiento crítico ($g = 0,67$). El modo de uso, la mediación docente y la alfabetización informacional emergen como factores moderadores determinantes. Se concluye que la integración ética de la IA en educación superior exige diseño instruccional deliberado con prompting estructurado y acompañamiento docente, de modo que la tecnología amplíe y no sustituya el pensamiento autónomo.

Declaración de conflictos de interés

Los autores declaran que no existen conflictos de interés con respecto a la publicación.

Contribución de los autores

R.B.R.Y., concibió y diseñó el estudio, recopiló y analizó los datos, interpretó los resultados, redactó el manuscrito y es responsable de su contenido y versión final. **P.N.R.M.**, colaboró en el cribado y recopilación de información, apoyó en la revisión de contenido y forma del manuscrito.

Referencias

- Aguirre Franco, M. de J. (2024). De la inteligencia artificial como instrumento: La tesis de la dependencia cognitiva. *Análisis. Revista de investigación filosófica*, 11(1), 63–79. https://doi.org/10.26754/ojs_arif/arif.202419489
- Albeño, C. E., Marín, C. L., Martínez Méndez, G. M., Mendoza Aguirre, N. M., & Rodríguez Samayoa, D. G. (2025). Dependencia funcional de la inteligencia artificial: caso comparativo entre usuarios de Costa Rica y El Salvador. *Realidad y Reflexión*, (62). <https://ri.ufg.edu.sv/jspui/handle/11592/10716>
- Aristóteles. (1989). *Organon: Analíticos Primeros y Segundos* (M. Candel Sanmartín, Trad.). Gredos. (Obra original del siglo IV a.C.)
- Arif, R. M., Nisa, R., & Allah, Z. S. (2025). The paradox of artificial intelligence usage in higher education: Between enhancing learning effectiveness and the decline of students' critical thinking skills. *Tadbir: Jurnal Manajemen Pendidikan Islam*, 13(2), 472–489. <https://doi.org/10.30603/tjmpi.v13i2.6803>
- Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19–32. <https://doi.org/10.1080/1364557032000119616>
- Balart, T., Díaz, B., & Shryock, K. J. (2026). A Systematic Literature Review on the Pedagogical Implications and Impact of GenAI on Students' Critical Thinking. *Algorithms*, 19(3), 179. <https://doi.org/10.3390/a19030179>
- Bastani, H., Bastani, O., Sungu, A., Geb, H., Kabaki, Ö., & Marimane, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122. <https://doi.org/10.1073/pnas.2422633122>
- Choudhuri, R., Sanchez, C. A., Burnett, M., & Sarma, A. (2025). Why Johnny can't think: GenAI's impacts on cognitive engagement. *arXiv preprint, arXiv:2601.22430v1*. <https://arxiv.org/abs/2601.22430v1>
- Cuenca, M. E., Subirats, L., Narbona-Reina, B., & Sacha, G.-M. (2025). Use of generative artificial intelligence by final degree project students: Is it useful in all steps of their work? *Applied Sciences*, 15(20), 11004. <https://doi.org/10.3390/app152011004>
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, Article 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. D. C. Heath and Company. <https://archive.org/details/howwethink00deweuoft>
- Dewey, J. (1938). *Logic: The theory of inquiry*. Henry Holt and Company.
- Facione, P. A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction (The Delphi Report)*. The California Academic Press. <https://files.eric.ed.gov/fulltext/ED315423.pdf>
- Gerlich, M. (2025a). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), Article 6. <https://doi.org/10.3390/soc15010006>

- Gerlich, M. (2025b). From offloading to engagement: An experimental study on structured prompting and critical reasoning with generative AI. *Data*, 10(11), Article 172. <https://doi.org/10.3390/data10110172>
- Grisales, S. (2025). Impacto de la inteligencia artificial, específicamente ChatGPT, en la dependencia tecnológica y reducción del pensamiento crítico en estudiantes universitarios [Trabajo fin de máster, Universitat Oberta de Catalunya]. Repositorio institucional UOC. <https://openaccess.uoc.edu/handle/10609/154321>
- Hong, H., Vate-U-Lan, P., & Viriyavejakul, C. (2025). Perceptions of the generative AI-enabled cognitive offload instruction in English writing. *International Journal of Evaluation and Research in Education*, 14(3). <https://doi.org/10.11591/ijere.v14i3.33138>
- Juárez, A., Ramírez, M., & Torres, F. (2025). Student creativity at risk: A scoping review of the negative effects of generative AI in university education. LACCEI Latin American and Caribbean Consortium of Engineering Institutions. <https://doi.org/10.18687/LACCEI2025.1.1.1001>
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux. <https://us.macmillan.com/books/9780374533557>
- Kant, I. (1996). An answer to the question: What is enlightenment? (M. J. Gregor, Trad.). En *Practical philosophy* (pp. 11-22). Cambridge University Press. <https://doi.org/10.1017/CBO9780511813306> (Obra original publicada en 1784)
- Kosmyna, N., et al. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv preprint, arXiv:2506.08872. <https://arxiv.org/abs/2506.08872>
- Kumar, P., Singh, R., & Sharma, A. (2025). From automation to augmentation: GenAI's role in promoting students' cognitive and ethical development. *International Journal of Information Management*, 85, 102800. <https://doi.org/10.1016/j.ijinfomgt.2025.102800>
- Legg, S., & Hutter, M. (2007). A collection of definitions of intelligence. arXiv preprint, arXiv:0706.3639. <https://arxiv.org/abs/0706.3639>
- Lengyel-Almos, I., & Bravo-Anduaga, M. (2025). Thinking with AI: Human–AI interaction and critical thinking in scenario-based learning. En *Actas de la AHFE 2025 International Conference on Human Factors in Design, Engineering, and Computing* (Vol. 1007). AHFE Open Access. <https://doi.org/10.54941/ahfe1007166>
- Mohale, N. E., & Suliman, Z. (2025). The influence of generative AI and its impact on critical cognitive engagement in an open access, distance learning university. *Compass: Journal of Learning and Teaching in Higher Education*, 18(1). <https://doi.org/10.21100/compass.v18i1.1588>
- Muñoz Martínez, C., Roger-Monzo, V., & Castelló-Sirvent, F. (2025). Generative AI and critical thinking in online higher education: challenges and opportunities. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2), 233–273. <https://doi.org/10.5944/ried.28.2.43556>
- Paul, R., & Elder, L. (2019). *Pensamiento crítico: Herramientas para tomar las riendas de tu vida profesional y personal* (3.ª ed.). Fundación para el Pensamiento Crítico. <https://www.criticalthinking.org>
- Raitskaya, L., & Tikhonova, E. (2025). Enhancing critical thinking skills in ChatGPT-human interaction: A scoping review. *Journal of Language and Education*, 11(2). <https://doi.org/10.17323/jle.2025.27387>
- Sevnarayan, K. (2026). Society 5.0's rise of GenAI and the impact on critical thinking in distance education. *Frontiers in Education*, 11, Article 1715468. <https://doi.org/10.3389/educ.2026.1715468>
- Sitepu, R., Nainggolan, J., & Simanjuntak, E. (2026). Mapping and exploring strategies to enhance critical thinking in the artificial intelligence era: A bibliometric and systematic review. *Eurasian Journal of Educational Research*, 111, 1–20. <https://doi.org/10.14689/ejer.2026.1.305>
- Tian, J., & Zhang, R. (2025). Learners' AI dependence and critical thinking: The psychological mechanism of fatigue and the social buffering role of AI literacy. *Acta Psychologica*, 260, 105725. <https://doi.org/10.1016/j.actpsy.2025.105725>
- Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., & Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine*, 169(7), 467–473. <https://doi.org/10.7326/M18-0850>
- Vereau Amaya, E. A., Soto Zavaleta, A. R., Vásquez Luján, I. G., González López, E., Tomanguilla Reyna, J., & Ríos Gonzales, J. D. (2025). Between assistance and dependence: Artificial intelligence and critical thinking. A study on first-year university students. *International Journal of Learning, Teaching and Educational Research*, 24(12), 118–134. <https://doi.org/10.26803/ijlter.24.12.6>
- Wangchuk, P. (2025). Conceptualising technology threat avoidance theory to examine cognitive risks of generative AI in higher education. En *Proceedings of the 31st Americas Conference on Information Systems (AMCIS 2025)*. AIS Electronic Library. <https://aisel.aisnet.org/amcis2025/ai/ai/17/>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*, 11, Article 28. <https://doi.org/10.1186/s40561-024-00316-7>
- Zhao, Z. (2026). Gradual cognitive externalization: A framework for understanding how ambient intelligence externalizes human cognition. arXiv preprint, arXiv:2604.04387v1. <https://arxiv.org/abs/2604.04387v1>
- Zhou, X., et al. (2025). Development and validation of the AI dependence scale for Chinese undergraduates (AIDep-22). *Frontiers in Psychology*, 16, Article 1725393. <https://doi.org/10.3389/fpsyg.2025.1725393>