

Inteligencia artificial generativa y pensamiento crítico en estudiantes de Educación Básica y Bachillerato: una revisión sistemática

Generative Artificial Intelligence and Critical Thinking Among Elementary and High School Students: A Systematic Review

Rosa Irene Molina Ríos¹

MINEDEC, Cuenca-Ecuador, (rirenemol@hotmail.com)

<https://orcid.org/0009-0004-4903-0283>

Fabián Eugenio Bravo Navas²

MINEDEC, Cuenca-Ecuador, (fabianbravoseito@hotmail.com)

<https://orcid.org/0009-0006-5352-1231>

María Augusta Cordero Pacheco³

MINEDEC, Cuenca-Ecuador, (maicordero2011@hotmail.com)

<https://orcid.org/0009-0000-2853-9970>

Ruth Maribel Gallegos Tandazo⁴

MINEDEC, Loja-Ecuador, (ruchis27071989@gmail.com)

<https://orcid.org/0009-0008-0087-3449>

Brayan Jhoel Campoverde Gallegos⁵

UNL, Loja-Ecuador, (campoverde.jbrayan@gmail.com)

<https://orcid.org/0009-0008-2499-8203>

Cómo citar

Molina Ríos, R. I., Bravo Navas, F. E., Cordero Pacheco, M. A., Gallegos Tandazo, R. M., & Campoverde Gallegos, B. J. (2026). Inteligencia artificial generativa y pensamiento crítico en estudiantes de Educación Básica y Bachillerato: una revisión sistemática. *Saber Estratégico Internacional*, 4(3), 1422–1432.

<https://doi.org/10.70577/s.e.v4i3.133>

RESUMEN

El crecimiento de la inteligencia artificial generativa (IAG) ha modificado las formas en que los estudiantes buscan información, producen textos, resuelven problemas y reciben retroalimentación. En Educación Básica y Bachillerato, esta expansión plantea una cuestión pedagógica central: determinar si las herramientas generativas pueden fortalecer el pensamiento crítico o si, por el contrario, favorecen la dependencia cognitiva y la aceptación acrítica de respuestas automatizadas. El objetivo de esta revisión sistemática fue sintetizar evidencia reciente sobre la relación entre el uso pedagógico de IAG y el desarrollo del análisis, la evaluación de información, la argumentación, el cuestionamiento y la resolución de problemas en estudiantes de niveles escolares. Se siguieron criterios inspirados en PRISMA 2020 y se priorizaron publicaciones revisadas por pares entre 2022 y 2026, con especial atención a estudios K-12, educación secundaria y revisiones recientes sobre IAG y pensamiento de orden superior. La síntesis muestra un efecto de doble naturaleza. Cuando la IAG se emplea como interlocutor, fuente que debe ser verificada, generador de perspectivas o apoyo para procesos de indagación, puede estimular la comparación de argumentos, el cuestionamiento, la reflexión y la resolución de problemas. En cambio, cuando se utiliza principalmente para obtener productos terminados o respuestas inmediatas, aumenta el riesgo de descarga cognitiva, dependencia, menor verificación y aprendizaje superficial. Se concluye que el impacto de la IAG no depende únicamente de la tecnología, sino del diseño de las tareas, la mediación docente, la alfabetización en IA y la exigencia de que el estudiante justifique, contraste y reconstruya las respuestas generadas.

Palabras clave: inteligencia artificial generativa; pensamiento crítico; ChatGPT; educación básica; bachillerato; alfabetización en IA; resolución de problemas.

ABSTRACT

The rapid growth of generative artificial intelligence (GenAI) has changed how students search for information, produce texts, solve problems, and receive feedback. In primary and secondary education, this expansion raises a central pedagogical question: whether generative tools can strengthen critical thinking or instead encourage cognitive dependence and uncritical acceptance of automated responses. This systematic review aimed to synthesize recent evidence on the relationship between pedagogical uses of GenAI and the development of analysis, information evaluation, argumentation, questioning, and problem-solving among school-age learners. Criteria inspired by PRISMA 2020 were followed, prioritizing peer-reviewed publications from 2022 to 2026, particularly K-12 and secondary education studies and recent reviews of GenAI and higher-order thinking. The synthesis reveals a dual effect. When GenAI is used as a dialogic partner, a source to be verified, a generator of alternative perspectives, or support for inquiry, it can stimulate comparison of arguments, questioning, reflection, and problem-solving. When it is used mainly to obtain finished products or immediate answers, risks of cognitive offloading, dependence, weak verification, and surface learning increase. The review concludes that the educational impact of GenAI depends less on the tool itself than on task design, teacher mediation, AI literacy, and requirements for students to justify, contrast, and reconstruct AI-generated responses.

Keywords: generative artificial intelligence; critical thinking; K-12 education; secondary education; AI literacy; problem-solving.

INTRODUCCIÓN

La incorporación de herramientas de inteligencia artificial generativa en educación se aceleró desde la aparición pública de sistemas conversacionales capaces de producir textos, explicaciones, código, imágenes y retroalimentación en lenguaje natural. Su presencia en el trabajo escolar ha dejado de ser marginal: estudiantes y docentes pueden utilizarlas para buscar ideas, explicar conceptos, revisar escritos, construir ejemplos y ensayar soluciones. Sin embargo, la facilidad con la que estos sistemas generan respuestas plausibles también obliga a reconsiderar qué significa aprender, argumentar y resolver problemas en un entorno donde una respuesta puede obtenerse en segundos.

En el ámbito K-12, la investigación reciente muestra que la educación relacionada con IA puede contribuir a la alfabetización tecnológica, la resolución de problemas y la reflexión ética cuando se apoya en experiencias prácticas y tareas cuidadosamente diseñadas (Lee & Kwon, 2024; Liu & Zhong, 2024). A la vez, las revisiones centradas en ChatGPT advierten problemas de exactitud, integridad académica, dependencia y necesidad de formación crítica para estudiantes y docentes (Zhang & Tur, 2024). La evidencia, por tanto, no permite asumir que la simple presencia de una herramienta generativa produzca aprendizaje de mayor nivel.

El pensamiento crítico es particularmente relevante en este escenario porque implica analizar información, valorar evidencia, identificar supuestos, contrastar perspectivas, justificar conclusiones y revisar el propio razonamiento. En ambientes mediados por IAG aparece además una dimensión específica: evaluar la credibilidad de una salida generada, comprobar fuentes, detectar errores o sesgos y decidir qué partes de una respuesta pueden utilizarse, modificarse o rechazarse. Una revisión reciente sobre pensamiento crítico en contextos de aprendizaje con IAG encontró precisamente una reorientación hacia estas competencias de verificación y evaluación de resultados automatizados (Ngo Cong-Lem & Nguyen Thi Thuy-Dung, 2026).

Los resultados disponibles son prometedores pero heterogéneos. El metaanálisis de Deng et al. (2025), basado en estudios experimentales, encontró efectos positivos globales de ChatGPT sobre el rendimiento y las disposiciones de pensamiento de orden superior, aunque también una reducción del esfuerzo mental y una elevada heterogeneidad entre estudios. En educación escolar, estudios cualitativos recientes muestran que la IAG puede favorecer razonamiento, argumentación y consideración de múltiples perspectivas cuando se configura como socio dialógico y no como autoridad que entrega la respuesta final (Tang & Putra, 2026). De forma complementaria, Tang et al. (2026) muestran que el cuestionamiento crítico depende de las condiciones pedagógicas, la intervención del docente y el diseño de la interacción.

En consecuencia, la pregunta relevante no es únicamente si la IAG es beneficiosa o perjudicial, sino bajo qué condiciones su uso moviliza procesos cognitivos de mayor complejidad. Esta revisión se concentra en Educación Básica y Bachillerato y examina la IAG como herramienta de aprendizaje, no solamente como contenido curricular. Se busca distinguir entre usos que amplían el razonamiento del estudiante y usos que sustituyen la actividad cognitiva que la tarea pretendía desarrollar.

1.1. Objetivo

Analizar sistemáticamente la evidencia reciente sobre la influencia del uso pedagógico de herramientas de inteligencia artificial generativa en el pensamiento crítico de estudiantes de Educación Básica y Bachillerato, con énfasis en análisis, argumentación, evaluación de información y resolución de problemas.

1.2. Preguntas de investigación

1. ¿Qué efectos o asociaciones reporta la literatura reciente entre el uso educativo de IAG y el pensamiento crítico en estudiantes de niveles escolares?
2. ¿Qué estrategias pedagógicas favorecen que la IAG estimule análisis, argumentación, cuestionamiento y resolución de problemas?
3. ¿Qué riesgos cognitivos aparecen cuando los estudiantes utilizan IAG para obtener respuestas o productos terminados?
4. ¿Qué condiciones de mediación docente, alfabetización en IA y diseño de tareas son necesarias para un uso formativo de estas herramientas?

MATERIAL Y MÉTODOS

Se realizó una revisión sistemática de orientación narrativa siguiendo los principios de transparencia, identificación, selección y síntesis establecidos por PRISMA 2020 (Page et al., 2021). Dado que la IAG educativa es un campo reciente y de rápida evolución, se combinó evidencia empírica escolar con revisiones sistemáticas y metaanálisis recientes que permitieran contextualizar los hallazgos K-12. La búsqueda fue actualizada en septiembre de 2026.

Estrategia de búsqueda

Se emplearon combinaciones en español e inglés de los términos: “generative artificial intelligence”, “generative AI”, “ChatGPT”, “large language models”, “critical thinking”, “critical questioning”, “argumentation”, “problem solving”, “AI literacy”, “K-12”, “secondary education”, “high school”, “middle school”, “educación básica” y “bachillerato”. Se priorizaron registros localizables en bases y portales académicos de amplio uso (Scopus/Web of Science cuando el registro bibliográfico estaba disponible, ScienceDirect, SpringerLink, Wiley Online Library, ERIC y repositorios institucionales de publicaciones revisadas por pares).

Criterios de inclusión y exclusión

Criterio	Inclusión	Exclusión
Periodo	2022-2026, priorizando publicaciones posteriores a la expansión de la IAG conversacional.	Estudios anteriores cuando no aportaban marco indispensable.
Población	Estudiantes K-12, Educación Básica, secundaria o Bachillerato; revisiones amplias con resultados relevantes para estos niveles.	Trabajos exclusivamente profesionales o empresariales sin relación educativa.
Fenómeno	Uso de IAG/LLM/ChatGPT relacionado con pensamiento crítico, cuestionamiento, argumentación, razonamiento o resolución de problemas.	IA no generativa sin relación con procesos cognitivos de orden superior.
Tipo de publicación	Artículos revisados por pares, revisiones sistemáticas, metaanálisis y estudios empíricos pertinentes.	Opiniones sin base metodológica, publicidad, entradas informales o fuentes sin trazabilidad.
Idioma	Español o inglés.	Otros idiomas sin información suficiente para extracción.

Selección y síntesis

La selección se realizó por pertinencia temática y calidad de la información metodológica disponible. La síntesis final se organizó alrededor de un núcleo de estudios y revisiones recientes directamente relacionados con IAG, educación

escolar y/o pensamiento crítico. No se efectuó un nuevo metaanálisis debido a la heterogeneidad de diseños, poblaciones, instrumentos y resultados. En su lugar, se aplicó una síntesis temática comparativa centrada en cinco categorías: (a) análisis y evaluación de información, (b) argumentación y multiplicidad de perspectivas, (c) resolución de problemas, (d) dependencia y descarga cognitiva, y (e) condiciones pedagógicas.

Esta revisión debe interpretarse como una síntesis sistemática actualizada y reproducible a partir de los criterios descritos, pero no como un metaanálisis exhaustivo de todas las bases de datos existentes. La rapidez de publicación del campo y la predominancia todavía limitada de estudios longitudinales K-12 constituyen restricciones relevantes.

RESULTADOS

Panorama de la evidencia reciente

Las revisiones publicadas entre 2024 y 2026 coinciden en que la investigación escolar sobre IAG crece con rapidez, pero sigue siendo menor que la investigación universitaria. Zhang y Tur (2024) sintetizaron 13 trabajos sobre ChatGPT en K-12 y destacaron oportunidades para personalización y apoyo al aprendizaje, junto con riesgos de calidad de las respuestas e integridad académica. Lee y Kwon (2024), al revisar 25 estudios de educación en IA en aulas K-12, identificaron mejoras en alfabetización en IA, resolución de problemas y reflexión ética, especialmente mediante actividades prácticas y aprendizaje basado en proyectos. Liu y Zhong (2024) analizaron 45 estudios empíricos de educación en IA y concluyeron que el diseño instruccional, el aprendizaje colaborativo, los contextos auténticos y el andamiaje docente son elementos esenciales.

En una revisión específicamente centrada en IAG K-12, Zhang, Lai y Yu (2026) analizaron 45 estudios y reportaron beneficios potenciales en rendimiento, capacidades cognitivas y motivación, pero también contenido engañoso, dependencia tecnológica, dificultades de comprensión y riesgos de privacidad. Los autores subrayan que gran parte de la evidencia disponible es de corto plazo. En el campo del aprendizaje de lenguas, Liu, Bin Sihes y Lu (2025) revisaron 15 estudios y encontraron una naturaleza dual: dos tercios de los estudios incluidos informaron efectos favorables sobre pensamiento crítico, mientras que un tercio reportó efectos negativos o riesgos, particularmente asociados con dependencia, fiabilidad y deficiencias en la integración pedagógica.

IAG como estímulo para analizar y evaluar información

Uno de los usos con mayor potencial consiste en convertir la respuesta de la IA en un objeto de análisis y no en una solución definitiva. En este enfoque, el estudiante debe identificar afirmaciones, comprobar datos, localizar inconsistencias, contrastar fuentes y explicar por qué acepta o rechaza una respuesta. Este cambio es relevante porque la alfabetización crítica en IA exige comprender que una respuesta lingüísticamente convincente no equivale necesariamente a una respuesta verdadera.

Javier y Moorhouse (2024), en una experiencia con estudiantes de secundaria que aprendían inglés, plantearon un proceso de uso productivo y crítico de ChatGPT que incluía conocimiento de las limitaciones de la herramienta y evaluación de sus salidas. Los autores señalan que los estudiantes necesitan aprender a formular indicaciones, interpretar respuestas y reconocer posibles sesgos, falsedades y riesgos de dependencia. De forma convergente, Tang

et al. (2026) observaron que el cuestionamiento crítico de estudiantes de décimo grado se fortalece cuando las actividades les exigen interrogar, profundizar y evaluar el contenido generado.

Esto sugiere que la IAG puede actuar como un “texto imperfecto” deliberadamente sometido a escrutinio. Desde el punto de vista didáctico, una respuesta incorrecta o incompleta puede ser más formativa que una respuesta perfecta si obliga al estudiante a justificar la corrección. El aprendizaje crítico se produce, por tanto, en la actividad de verificación y reconstrucción, no en la mera recepción del contenido.

Argumentación, perspectivas múltiples y diálogo

Los sistemas generativos pueden producir explicaciones alternativas, objeciones, contraargumentos y perspectivas distintas sobre un mismo problema. Esta capacidad resulta pedagógicamente valiosa cuando se utiliza para ampliar el espacio de deliberación. Tang y Putra (2026) estudiaron a 21 estudiantes de dos escuelas secundarias que interactuaron con un chatbot diseñado como socio dialógico en ciencias. El análisis de los registros de conversación identificó toma de perspectiva, razonamiento, argumentación y pensamiento creativo. El hallazgo central no fue que el chatbot “enseñara” automáticamente pensamiento crítico, sino que su diseño dialógico provocaba respuestas que requerían explicar, defender, revisar y producir alternativas.

De manera similar, Tang et al. (2026) muestran que el nivel de cuestionamiento depende de la estructura de la tarea y de la intervención docente. Cuando el estudiante formula preguntas de seguimiento, solicita evidencia, detecta contradicciones o desafía una afirmación, la interacción humano-IA se aproxima a un diálogo socrático. Cuando se limita a preguntar “¿cuál es la respuesta?”, la misma tecnología puede reducir la actividad intelectual.

Resolución de problemas y pensamiento de orden superior

Los hallazgos generales sobre resolución de problemas son favorables, aunque deben interpretarse con cautela. Lee y Kwon (2024) concluyeron que los programas de educación en IA K-12 pueden fortalecer habilidades de resolución de problemas, mientras que Deng et al. (2025) encontraron, en 69 estudios experimentales de distintos niveles educativos, un efecto global positivo de ChatGPT sobre las disposiciones de pensamiento de orden superior. El metaanálisis reportó un tamaño de efecto significativo para estas disposiciones ($g \approx 0,70$), pero también heterogeneidad elevada, lo que indica que los resultados dependen fuertemente del contexto, la intervención y la forma de evaluación.

En educación preuniversitaria, la investigación cualitativa de 2025 sobre IAG como mediador cognitivo en pensamiento computacional encontró que la herramienta podía favorecer creatividad y fluidez computacional, pero exigía vigilancia epistémica, reflexión crítica y valoración de la viabilidad de las soluciones. Este patrón coincide con la idea de que la IAG es más útil cuando ayuda a explorar rutas de solución, depurar procedimientos y comparar alternativas, sin sustituir la decisión final del estudiante.

Riesgo de dependencia, descarga cognitiva y aprendizaje superficial

El principal riesgo identificado en la literatura es que la eficiencia de la IAG elimine precisamente el esfuerzo cognitivo que una tarea pretendía provocar. Si el estudiante delega la formulación del problema, la búsqueda de evidencia, la construcción del argumento y la redacción final, puede entregar un producto de alta apariencia sin haber realizado los procesos intelectuales asociados.

Deng et al. (2025) encontraron que las intervenciones con ChatGPT redujeron significativamente el esfuerzo mental en el conjunto de estudios analizados. Este resultado no es necesariamente negativo —una herramienta puede disminuir carga irrelevante—, pero se vuelve problemático cuando reduce el esfuerzo germano necesario para comprender, analizar o transferir conocimiento. Zhang, Lai y Yu (2026) también identificaron dependencia tecnológica entre los desafíos recurrentes de la IAG en K-12, mientras que Liu et al. (2025) señalan que la dependencia del usuario y la fiabilidad de la IAG son limitaciones centrales para el desarrollo del pensamiento crítico.

Por ello, una tarea escolar no debería evaluarse únicamente por la calidad del producto final cuando existe acceso a IAG. Es necesario observar el proceso: fuentes consultadas, decisiones tomadas, versiones sucesivas, justificación de cambios, errores detectados en la IA y reflexión sobre qué aportó el estudiante. Esta transformación de la evaluación es particularmente importante en Bachillerato, donde la argumentación y la autonomía intelectual forman parte de las competencias que se espera consolidar.

Síntesis de hallazgos

Dimensión	Uso de IAG con mayor potencial	Riesgo si se usa pasivamente	Implicación didáctica
Análisis	Comparar respuestas, detectar errores, clasificar evidencia.	Aceptar respuestas plausibles sin comprobarlas.	Exigir verificación con fuentes y explicación de discrepancias.
Argumentación	Generar contraargumentos y perspectivas alternativas.	Copiar argumentos generados sin apropiación conceptual.	Solicitar tesis propia, evidencia, contraargumento y réplica.
Resolución de problemas	Explorar estrategias, recibir pistas, depurar procedimientos.	Delegar el procedimiento completo y copiar la solución.	Usar IA como andamiaje; pedir al estudiante explicar cada decisión.
Metacognición	Comparar respuesta inicial propia con la de la IA y revisar el razonamiento.	Sustituir la reflexión por corrección automática.	Incluir diarios de decisiones y reflexión posterior.
Alfabetización informacional	Auditar citas, fuentes, sesgos y afirmaciones.	Confiar en referencias inventadas o información desactualizada.	Aplicar protocolos de comprobación y triangulación.

DISCUSIÓN

La evidencia sintetizada permite rechazar dos posiciones simplificadoras: considerar que la IAG, por sí sola, mejora el pensamiento crítico, o asumir que inevitablemente lo deteriora. Los resultados apuntan a un efecto condicionado.

La misma herramienta puede funcionar como atajo cognitivo o como provocador intelectual dependiendo de cómo se diseñe la actividad.

El mecanismo pedagógico más consistente es la transformación de la IAG de “proveedor de respuestas” a “interlocutor cuestionable”. Cuando el estudiante debe evaluar la salida generada, pedir fundamentos, contrastarla con fuentes, identificar omisiones y producir una conclusión propia, la tecnología introduce material adicional para analizar. Esto coincide con los estudios escolares de Tang et al. (2026) y Tang y Putra (2026), en los que el valor de la interacción estuvo vinculado a cuestionamiento, diálogo y argumentación.

Un segundo elemento es la mediación docente. La alfabetización en IA no puede reducirse a enseñar a escribir mejores prompts. Incluye comprender límites, probabilidades de error, sesgos, privacidad, atribución, integridad académica y criterios de confiabilidad. Liu y Zhong (2024) y Zhang, Lai y Yu (2026) resaltan que las capacidades del profesorado y el diseño instruccional son condiciones para una integración responsable. Esto es especialmente importante en Educación Básica, donde los estudiantes pueden poseer menor experiencia para distinguir autoridad aparente de evidencia verificable.

Un tercer aspecto es la evaluación. Si la consigna puede resolverse completamente con una petición simple a un chatbot, la tarea mide cada vez menos el razonamiento del estudiante. La respuesta no debería ser únicamente prohibir la herramienta, sino rediseñar actividades para que la evidencia del aprendizaje esté en el proceso. Pueden incorporarse defensas orales, comparación de fuentes, anotación de prompts, identificación de errores de la IA, versiones comentadas, resolución de casos contextualizados y producción de conclusiones justificadas.

Finalmente, la evidencia disponible continúa presentando limitaciones. Gran parte de los estudios son de corta duración, utilizan muestras pequeñas o se concentran en educación superior. Incluso cuando se reportan mejoras en pensamiento de orden superior, existe heterogeneidad metodológica y diversidad de instrumentos. Por ello, se requieren estudios longitudinales en primaria, secundaria y Bachillerato que midan transferencia: no solo si el estudiante produce una mejor respuesta con IA, sino si luego puede analizar y resolver problemas de forma independiente.

Propuesta pedagógica derivada de la revisión

A partir de la evidencia se propone el ciclo VERIFICA para integrar IAG sin desplazar el pensamiento crítico. No constituye un instrumento validado, sino una síntesis didáctica derivada de los patrones encontrados en la literatura revisada.

V - Valorar el problema: El estudiante formula primero qué necesita resolver y qué sabe antes de consultar la IA.

E - Explorar: Utiliza la IAG para obtener explicaciones, alternativas, preguntas o posibles soluciones, evitando solicitar directamente un producto final.

R - Revisar: Identifica afirmaciones comprobables, posibles errores, supuestos y aspectos que requieren evidencia.

I - Investigar: Contrasta la respuesta con fuentes académicas, textos de clase, datos o documentos confiables.

F - Fundamentar: Construye una posición propia y explica por qué acepta, modifica o rechaza lo generado.

I - Iterar: Formula nuevas preguntas, contraejemplos o condiciones para mejorar el razonamiento.

C - Comunicar: Presenta una solución o argumento propio, señalando de manera transparente el papel de la IA.

A - Autoevaluar: Reflexiona sobre qué aprendió, qué hizo la herramienta y qué parte del razonamiento pudo realizar sin ella.

Implicaciones para Educación Básica y Bachillerato

- En Educación Básica, conviene priorizar tareas breves de verificación: comparar una respuesta de IA con el libro de texto, localizar dos errores, distinguir hechos de opiniones y explicar con palabras propias una corrección.
- En Educación General Básica Superior, pueden incorporarse debates, análisis de noticias, resolución de problemas con estrategias alternativas y ejercicios de detección de sesgos o referencias inexistentes.
- En Bachillerato, la IAG puede utilizarse para generar contraargumentos, simular posiciones, revisar modelos de solución, analizar casos y apoyar proyectos de investigación, siempre que el estudiante documente fuentes, decisiones y modificaciones.
- En todos los niveles, la evaluación debe valorar la justificación y el proceso. Una respuesta correcta sin evidencia del razonamiento ofrece información limitada sobre el aprendizaje cuando existen herramientas capaces de producirla automáticamente.

Limitaciones de la revisión

El campo cambia con gran rapidez y nuevos estudios pueden modificar las conclusiones en periodos breves. La evidencia específicamente centrada en estudiantes de Educación Básica y Bachillerato todavía es menor que la disponible en educación superior. Además, los estudios utilizan definiciones e instrumentos distintos para pensamiento crítico, razonamiento, cuestionamiento y habilidades de orden superior, lo que dificulta comparaciones directas. Esta revisión realizó una síntesis narrativa y no calculó un efecto combinado propio. Los resultados cuantitativos citados proceden de metaanálisis publicados y deben interpretarse en función de sus poblaciones y diseños.

CONCLUSIONES

La inteligencia artificial generativa puede contribuir al pensamiento crítico de estudiantes de Educación Básica y Bachillerato, pero el beneficio no surge de la generación automática de respuestas. Surge de las actividades cognitivas que el estudiante realiza alrededor de esas respuestas: preguntar, comprobar, comparar, argumentar, corregir y decidir.

Los resultados más consistentes favorecen un modelo de uso en el que la IAG funciona como andamiaje, interlocutor y fuente provisional que debe ser evaluada. La mediación docente y el diseño de tareas son determinantes. Pedir una respuesta terminada tiende a reducir oportunidades de razonamiento; pedir que el estudiante audite, contradiga, mejore o defienda una respuesta puede convertir la interacción en una experiencia de pensamiento de orden superior.

Por tanto, la pregunta educativa no debería ser simplemente si se permite o se prohíbe la IAG. El desafío es diseñar experiencias en las que el estudiante conserve la responsabilidad epistemológica: debe saber qué afirma, con qué

evidencia lo sostiene, qué errores detectó, qué decisiones tomó y por qué su conclusión es razonable. En ese escenario, la IAG puede ampliar el espacio para pensar sin reemplazar al estudiante como autor de su razonamiento.

REFERENCIAS BIBLIOGRÁFICAS

- Anwar, N. (2025). The use of generative artificial intelligence to develop student research, critical thinking, and problem-solving skills. *Trends in Higher Education*, 4(3), 34. <https://doi.org/10.3390/higheredu4030034>
- Belghith, Y., Mahdavi Goloujeh, A., Magerko, B., Long, D., McKlin, T., & Roberts, J. (2024). Testing, socializing, exploring: Characterizing middle schoolers' approaches to and conceptions of ChatGPT. *Proceedings of the CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3613904.3642332>
- Bueie, A. A., Skar, G. B., & Graham, S. (2026). High school students' use and beliefs about generative artificial intelligence and writing in school. *Reading and Writing*, 39, 597–627. <https://doi.org/10.1007/s11145-025-10702-3>
- 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, 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Javier, D. R. C., & Moorhouse, B. L. (2024). Developing secondary school English language learners' productive and critical use of ChatGPT. *TESOL Journal*, 15(2), e755. <https://doi.org/10.1002/tesj.755>
- Lee, S. J., & Kwon, K. (2024). A systematic review of AI education in K-12 classrooms from 2018 to 2023: Topics, strategies, and learning outcomes. *Computers and Education: Artificial Intelligence*, 6, 100211. <https://doi.org/10.1016/j.caeai.2024.100211>
- Liu, J., Bin Sihes, A. J., & Lu, Y. (2025). How do generative artificial intelligence (AI) tools and large language models (LLMs) influence language learners' critical thinking in EFL education? A systematic review. *Smart Learning Environments*, 12, 48. <https://doi.org/10.1186/s40561-025-00406-0>
- Liu, X., & Zhong, B. (2024). A systematic review on how educators teach AI in K-12 education. *Educational Research Review*, 45, 100642. <https://doi.org/10.1016/j.edurev.2024.100642>
- Ngo Cong-Lem, & Nguyen Thi Thuy-Dung. (2026). Towards a framework for understanding and assessing critical thinking in generative AI-enhanced learning environments: A scoping review. [Scoping review; publication record, Monash University].
- 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., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>

- Tang, K.-S., Cooper, G., Rappa, N., & Edwards, J. (2026). Critical questioning with generative AI: Developing AI literacy in secondary education. *Thinking Skills and Creativity*, 59, 102043. <https://doi.org/10.1016/j.tsc.2025.102043>
- Tang, K.-S., & Putra, G. B. S. (2026). Generative AI as a dialogic partner: Enhancing multiple perspectives, reasoning, and argumentation in science education with customized chatbots. *Journal of Science Education and Technology*, 35, 128–140. <https://doi.org/10.1007/s10956-025-10240-1>
- Yusuf, A., & Md Noor, N. (2024). Generative AI in education and research: A systematic mapping review. *Review of Education*, 12, e3489. <https://doi.org/10.1002/rev3.3489>
- Zhang, P., & Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. *European Journal of Education*, 59, e12599. <https://doi.org/10.1111/ejed.12599>
- Zhang, T., Lai, Y. C., & Yu, L. H. P. (2026). Generative artificial intelligence in K-12 education: A systematic review. *Research and Practice in Technology Enhanced Learning*, 21, Article 34. <https://doi.org/10.58459/rptel.2026.21034>