

DE LA RESOLUCIÓN AL RAZONAMIENTO: REVISIÓN SISTEMÁTICA DE LAS BRECHAS COGNITIVAS GENERADAS POR EL USO DE IA EN EL APRENDIZAJE DE LAS MATEMÁTICAS

FROM PROBLEM-SOLVING TO REASONING: A SYSTEMATIC REVIEW OF COGNITIVE GAPS RESULTING FROM THE USE OF AI IN MATHEMATICS LEARNING

Ana Elizabeth González González¹

Víctor Raúl Camargo Colmenares²

John Jairo González González³

Cómo citar: González González, A. E., Camargo Colmenares, V. R., & González González, J. J. (2026). De la resolución al razonamiento: revisión sistemática de las brechas cognitivas generadas por el uso de la IA en el aprendizaje de las matemáticas. *Conocimiento Global, 11*(S1 Especia), 343-364. <https://doi.org/10.70165/cglobal.v11iS1Especia.699>

Resumen

La inteligencia artificial ha hecho que la resolución de operaciones y problemas matemáticos por parte de los estudiantes sea eficaz, pero si ese uso no tiene una intención pedagógica dirigida a desarrollar el pensamiento matemático, pueden aparecer brechas. Este trabajo estudió las brechas cognitivas encontradas en la literatura mediante el análisis de fuentes bibliográficas por medio de una revisión sistemática a través del protocolo PRISMA, cubriendo estudios publicados entre 2022 y 2026, resultando de la búsqueda tras aplicar los cuatro criterios de inclusión/exclusión un total de 60 estudios. Los hallazgos indican que los estudiantes rinden cuando les asiste la IA, pero ese rendimiento no es mantenido cuando trabajan sin ella. Pierden capacidad para el desarrollo del razonamiento matemático y la argumentación, se apoyan en la herramienta antes que en su propio esfuerzo personal y tienen menor control metacognitivo y menor autorregulación de su propio aprendizaje. Estas brechas se acentúan en donde no hay mediación pedagógica. En consecuencia, se concluye que el problema no se encuentra en la IA sino en la forma en que se usa sin intencionalidad formativa. Es por eso que el docente debe convertir esa herramienta en andamiaje del razonamiento matemático.

Palabras clave: inteligencia artificial, educación matemática, pensamiento matemático, brechas cognitivas, revisión sistemática.

Abstract

Artificial intelligence has made it easier for students to solve mathematical problems and perform calculations, but if its use lacks a pedagogical purpose aimed at developing mathematical thinking, gaps may emerge. This study examined the cognitive gaps identified in the literature by analyzing bibliographic sources through a systematic review following the PRISMA protocol, covering studies published between 2022 and 2026. After applying the four inclusion and exclusion criteria,

Recepción: 02 de agosto de 2026 / Evaluación: 29 de agosto de 2026 / Aprobado: 18 de septiembre de 2026

¹ Doctora en Educación Matemática, Universidad Antonio Nariño – Bogotá, Colombia. E-Mail: anaelizabet.gonzalez@uptc.edu.co, ORCID: <https://orcid.org/0000-0002-0164-1059>

² Doctorando en Estadística, Universidad de los Andes (ULA) - Merida Venezuela E-Mail: camargomathvictor@outlook.com, ORCID: <https://orcid.org/0009-0004-0941-8769>

³ Doctorando en Educación, Universidad Pedagógica Experimental Libertador. E-Mail: jhonjg0150@gmail.com, ORCID: <https://orcid.org/0009-0002-3190-9586>



the search yielded a total of 60 studies. The findings indicate that students perform well when assisted by AI, but that performance is not sustained when they work without it. They lose the ability to develop mathematical reasoning and argumentation, rely on the tool rather than their own personal effort, and exhibit less metacognitive control and self-regulation of their own learning. These gaps are accentuated in the absence of pedagogical mediation. Consequently, it is concluded that the problem lies not with AI itself but with the way it is used without a formative intent. That is why teachers must transform this tool into a scaffold for mathematical reasoning.

Keywords: artificial intelligence, mathematics education, mathematical thinking, cognitive gaps, systematic review

Introducción

El proceso de expansión de la inteligencia artificial (IA) generativa tiene efectos en las condiciones de acceso que tienen los estudiantes al conocimiento matemático. La existencia de herramientas que efectúan operaciones, que desarrollan procedimientos para resolver un ejercicio matemático y que explican soluciones matemáticas paso a paso son recursos escolares que se usan como una consulta cotidiana, con grados de acierto que superan las correspondientes a tareas típicas de la educación básica o media (Kasneci, 2023). Esta disponibilidad transforma la actividad cognitiva del aprendiz, puesto que traslada hacia el sistema artificial procesos que, de manera tradicional, requieren esfuerzo mental.

Sin embargo, el aprender matemáticas no se agota en saber dar las respuestas correctas. La competencia matemática conjuga la fluidez en procedimientos con la comprensión conceptual, el razonamiento, la argumentación y la capacidad de justificar la veracidad de una solución (OECD, 2023). Cuando en la resolución de un problema se recurre a un agente externo, la respuesta puede ser correcta sin que el alumno haya realizado esos procesos de pensamiento que forman parte del aprendizaje auténtico; esa disociación entre el resolver y el razonar es lo que constituye el centro del problema de este estudio.

La literatura reciente ha ido dando cuenta tanto de las oportunidades como de las amenazas que puede ofrecer tal mediación. Varios estudios muestran cómo la IA puede ser una oportunidad al personalizar la retroalimentación, al adecuar los problemas para el nivel que tiene la persona aprendiente y al soportar sistemas de tutoría inteligente (Crompton y Burke, 2023; Mohebi, 2024). Otros estudios subrayan que, la delegación sistemática del esfuerzo cognitivo en recursos externos (externalización cognitiva), puede ir en detrimento de las propias funciones que la escuela busca desarrollar (Tlili, 2023; Wu et al., 2025). Esta tensión está dispersa en estudios aislados, de contextos y metodologías dispares, sin que exista una síntesis sistemática que dé cuenta del carácter de las brechas cognitivas que conlleva el uso de la IA en el aprendizaje matemático; vacío que sustenta esta revisión.

La cuestión se pone de manifiesto en un fenómeno que los profesores reportan cada vez con más frecuencia, en el sentido de que los estudiantes resuelven los problemas sin dificultad cuando tienen al alcance la ayuda que proporciona la IA, mas no en contextos de desempeño autónomo, como el aula o la evaluación presencial, siendo incapaces, entonces, de argumentar sus procedimientos o transferir lo que aparentemente se ha aprendido a nuevas situaciones (UNESCO, 2023). De este modo, se formula la pregunta que orienta esta investigación, es decir, ¿cuáles serían las brechas cognitivas que la evidencia científica documenta en torno del uso de la IA en el aprendizaje de las matemáticas y en qué condiciones pedagógicas emergerían o disminuirían?

La respuesta a esta pregunta aporta el conocimiento desde dos planos. Un plano teórico, porque su revisión sitúa un cuerpo de evidencia no estructurado y pendiente de gestionar con constructos como el pensamiento matemático, la transferencia del aprendizaje y la autorregulación, dando la posibilidad de entender el fenómeno en cuestión más allá de la anécdota escolar. Y un plano práctico, dado que se proporciona a los docentes, a las instituciones y a los responsables de la política educativa criterios organizados y fundamentados que permiten orientar el uso de la IA con fines formativos, de manera que la inteligencia artificial actúe como andamiaje del razonamiento y no como sustituto del estudiante.

En función de lo anterior, la finalidad del estudio es la de analizar la evidencia científica existente sobre las brechas cognitivas ocasionadas por el uso de la IA en la enseñanza de las matemáticas en el marco de una revisión sistemática de literatura bajo el protocolo PRISMA que considera estudios publicados entre el año 2022 y 2026; para ello se persiguen los propósitos de documentar los efectos del razonamiento, de la argumentación y de la metacognición matemática, caracterizar los patrones de dependencia y de delegación del esfuerzo cognitivo reportados en la literatura y, por último, establecer las condiciones pedagógicas que la evidencia relaciona con el incremento o la atenuación de dichas brechas.

Referente Teórico

El marco teórico establece cinco ejes que aportan los criterios interpretativos para las brechas cognitivas analizadas. En primer lugar, se delimita la resolución de problemas como actividad cognitiva y se describe el pensamiento matemático en sus elementos de razonar, argumentar y la metacognición. En segundo lugar, se analiza el proceso de exteriorización de los procesos cognitivos en la enseñanza mediada por la IA, el estado del campo de la inteligencia artificial y la enseñanza de matemáticas y, finalmente, la descripción de la transferencia de aprendizaje y de la auto-regulación, criterios que permiten distinguir la actuación en situaciones de aprendizaje asistidas de la actuación autónoma.

De la ejecución procedimental a la comprensión conceptual: la resolución de problemas como actividad cognitiva

La resolución de los problemas matemáticos se configura como una actividad cognitiva de alta complejidad, que se puede considerar como ir más allá de la recuperación mecánica de algoritmos o métodos previamente aprendidos. Resolver un problema es comprender la información que se presenta, seleccionar y coordinar las estrategias, supervisar cómo va la solución y comprobar la validez de los resultados que van surgiendo. Desde esta visión, el pensamiento matemático se establece a través de procesos de planificación, regulación y toma de decisiones que exigen una activa participación cognitiva del estudiante. Así, en la resolución de un problema no se logra únicamente alcanzando la respuesta final, sino también con los procesos intelectuales que permitan llegar a la misma (Schoenfeld, 1985; OECD, 2023).

La literatura reciente sobre competencia matemática sostiene que el aprendizaje significativo requiere la interrelación de distintas dimensiones del conocimiento; entre ellas, la comprensión conceptual y la fluidez procedimental son de especial relevancia y, a la vez, complementarias entre sí. La primera permite establecer los principios y las relaciones que hacen fundamentadas a las ideas matemáticas; mientras que la segunda se puede expresar a partir de la ejecución flexible de los procedimientos. Pese a su interrelación, poseer habilidades procedimentales no garantiza por sí mismo la comprensión de los conceptos que se hallan bajo el

mismo ni el desarrollo de un razonamiento matemático de mayor profundidad (Kilpatrick et al., 2001; OECD, 2023).

La construcción del saber matemático también requiere la posibilidad de movilizar distintos sistemas de representación. Los objetos matemáticos se explicitan mediante registros semióticos, como expresiones simbólicas, representaciones gráficas, lenguaje natural, esquemas geométricos, entre otros, cuya coordinación es necesario establecer para comprender. El aprendizaje se potencia cuando el estudiante es capaz de transitar entre estas formas de representación, poder realizar un análisis de las relaciones que establecen y poder utilizarlas para poder producir sentido; sin embargo, la ejecución de procedimientos aislados de la posibilidad de la coordinación representacional, en la mayor parte de las ocasiones, conducen a aprendizajes fragmentados y con baja transferibilidad (Duval, 2006; NCTM, 2023).

Desde esta óptica, el hecho de obtener la solución correcta de un problema no se considera una forma de razonar matemáticamente. Seguramente, llegar a una respuesta matemática exacta puede, entre otras, proceder de la simple reproducción de procedimientos memorizados o de la validación de recursos externos que pueden hacer que parte del procesamiento cognitivo no sea llevado a cabo por las personas. Por el contrario, el razonamiento matemático lleva implícito el hecho de justificar procedimientos, de establecer conexiones conceptuales, de argumentar la pertinencia de las soluciones e incluso la necesidad de reflexionar sobre las estrategias empleadas. De modo que la diferenciación entre ejecución procedimental y comprensión conceptual es un criterio analítico fundamental en el margen de las capacidades cognitivas de las herramientas de IA en el aprendizaje de las matemáticas.

El pensamiento matemático: razonamiento, argumentación y procesos metacognitivos

El pensamiento matemático se configura como un entramado de procesos cognitivos en que el estudiante comprende situaciones, realiza conexiones y genera significados a partir de la actividad matemática. Desde este punto de vista, Ulusoy y Girit-Yildiz (2024), afirman que el razonamiento matemático supone la trascendencia de la simple ejecución de procedimientos y para ello requiere generar inferencias justificadas que permitan comprender y sustentar las relaciones que existen en una situación problemática. Dichas inferencias pueden ser inducciones, cuando se encuentran regularidades a partir de casos concretos, deducciones cuando se generan conclusiones a partir de principios generales o enunciados generales, y abducciones cuando se ofrecen explicaciones plausibles, frente a situaciones inicialmente inciertas. Estas modalidades de razonamientos convierten la actividad matemática en un proceso de construcción intelectual, en vez de una simple reproducción de algoritmos ya aprendidos.

La construcción del conocimiento matemático también requiere que las ideas generadas por los individuos sean sometidas a procesos de validación. En este sentido, Cervantes-Barraza et al. (2020) afirman que la argumentación y la justificación son prácticas insustituibles del pensamiento matemático, y lo son porque permiten: explicar las razones que sostienen un procedimiento; dar lógica a las relaciones que enlazan las diferentes formas en que se plantea un problema, e, incluso, fundamentar la validez de la respuesta a la que se ha llegado. Desde esta visión, una respuesta considerada correcta no tiene suficiente valor formativo si no la acompaña un argumento que haga explícitas las relaciones y los principios que otorgan sentido a dicha respuesta. Esto implica que la actividad matemática no consiste únicamente en producir respuestas, sino que también supone construir y comunicar razones que las validen.

A su vez, Ellis et al., (2022) afirman que el pensamiento matemático tiene niveles de complejidad más altos cuando el estudiante supera los casos concretos y establece relaciones más

generales y abstractas. La generalización es el propio proceso de encontrar regularidades, reconocer estructuras invariantes y formar principios que se pueden aplicar a situaciones diferentes de las que se construyeron. Tal competencia de abstracción hace que se favorezca la transferencia del aprendizaje, y que se puedan poner en juego los contenidos matemáticos de forma flexible en situaciones nuevas y ante situaciones problemáticas que presentan relaciones subyacentes.

La metacognición es otra de las dimensiones que se añaden al razonamiento, la argumentación y la generalización para conformar otro componente necesario del pensamiento matemático. A tal efecto, Turgut y Uğurlu (2024) consideran que la resolución de problemas tal y como se espera de un estudiante que sigue un razonamiento en buena lógica no es más que una actividad para la que se requiere que el estudiante ejecute un control consciente sobre la actividad cognitiva, los procesos de planificación, supervisión y evaluación, es decir, el control sobre la actividad de resolución de los mismos problemas. De esta forma, la planificación puede orientarle hacia una selección adecuada de estrategias, la supervisión le permite por su parte hacer un control en sí mismo en el desarrollo del proceso para examinar el avance o detectar errores, y la evaluación da lugar a una reflexión crítica sobre las decisiones adoptadas y a una regulación de las acciones a seguir. Por tanto, la debilidad o complejidad de cualquiera de estos procesos puede ser entendida también como brecha cognitiva, pues determinan la incapacidad de construir un pensamiento matemático profundizado, transferible y autónomo.

Metodología

La presente investigación ha resultado en una estrategia documental de carácter cualitativo en sí misma consistente en una revisión sistemática de la literatura para identificar, analizar e integrar los hallazgos científicos en torno a las brechas cognitivas del uso de la inteligencia artificial en la enseñanza-aprendizaje de las matemáticas; revisión realizada siguiendo los lineamientos PRISMA (Page et al., 2021) para asegurar la transparencia, trazabilidad y reproducibilidad del estudio. Este diseño resulta ser particularmente significativo por la dispersión y heterogeneidad de la evidencia extraída de un corpus tan diverso proveniente de contextos, niveles educativos o enfoques metodológicos distintos y que se ha integrado mediante procedimientos sistemáticos de búsqueda, selección, valoración y síntesis.

Se llevó a cabo la recuperación documental sobre las fuentes académicas publicadas desde enero de 2022 hasta mayo de 2026, periodo que coincide con la expansión masiva de la inteligencia artificial generativa en los contextos escolares a partir de la liberación pública de los grandes modelos de lenguaje, circunstancia que agrupa, en dicho periodo, la producción científica sobre el fenómeno que se estudia. La búsqueda fue realizada en bases de datos académicas de ámbito internacional, como Scopus, Web of Science, ERIC y ScienceDirect, elegidas en función de su cobertura en el ámbito de la educación matemática, la tecnología educativa y la psicología del aprendizaje, integrando también Redalyc y SciELO con la finalidad de recuperar la producción científica iberoamericana y Google Scholar como recurso complementario de apoyo, para identificar los estudios que no se había recuperado a partir de las bases de datos principales.

Las técnicas empleadas para la generación de la estrategia de búsqueda reconocieron el uso de operadores booleanos (y, o) juntamente con las palabras clave en inglés como: ("artificial intelligence", "generative AI", "ChatGPT", "intelligent tutoring system", "mathematics education", "mathematics learning", "mathematical reasoning", "problem solving", "cognitive skill", "metacognition", "cognitive offloading", "reliance" y "dependence"); en español: ("inteligencia artificial", "IA generativa", "educación matemática", "aprendizaje de las matemáticas", "razonamiento matemático", "barreras cognitivas", "metacognición").

Se consideraron cuatro criterios de inclusión para asegurar la pertenencia del corpus documental: a) artículos científicos revisados por pares, publicados entre 2022 y 2026, b) documentos en texto completo, escrita en español o inglés, c) investigaciones en las que se tratase del uso de la inteligencia artificial en la enseñanza/aprendizaje de las matemáticas vinculadas a procesos cognitivos, como el razonamiento, la argumentación, metacognición, autorregulación, transferencia o dependencia tecnológica, d) investigaciones centradas en la evidencia empírica o revisiones de la literatura relacionada con el objeto de estudio. Igualmente se consideraron prioritarios aquellos trabajos de acceso abierto, por el hecho que facilita la comprobación de los resultados detallados y la transparencia de los resultados.

De manera complementaria, se definieron criterios de exclusión que permitieron depurar los registros recuperados. Fueron descartados los documentos correspondientes a literatura gris, tesis, capítulos de libro, editoriales, cartas al editor y resúmenes de congresos, dado que estos formatos no garantizan un proceso de arbitraje equivalente al de los artículos científicos. Igualmente, se eliminaron los registros duplicados entre bases de datos, los documentos sin acceso al texto completo y aquellas investigaciones que, si bien abordaban la inteligencia artificial en relación con las matemáticas, lo hacían desde perspectivas exclusivamente técnicas o computacionales, centradas en el rendimiento de los modelos para resolver tareas matemáticas, sin vincularse con los procesos de aprendizaje ni con los efectos cognitivos en los estudiantes.

La selección tuvo lugar a través de diversas fases en función de las etapas del protocolo PRISMA. En la fase de identificación se recuperaron 178 registros de las bases de datos que se consultaron. A partir de la eliminación de 34 estudios quedaron 144 documentos a los cuales se les aplicó el proceso de cribado mediante el análisis de títulos y resúmenes, se descartaron 61 por no cumplir los criterios de inclusión preestablecidos. Quedaron 83 artículos que se evaluaron a texto completo y se valoró su elegibilidad, eliminando 23 artículos que no se enmarcaban dentro de los requisitos predefinidos. Finalmente, el corpus documental quedó integrado por 60 estudios que constituyen la base de la revisión efectuada.

Con el fin de organizar y estructurar la información recuperada, se creó un instrumento de extracción de datos que facilitó la toma de información relacionada con los diferentes artículos incluidos en el meta-análisis, permitiendo registrar para cada una de estas fuentes los siguientes aspectos: autoría, año de publicación del artículo, país o región de procedencia, nivel educativo del colectivo estudiado, tipo de herramienta de inteligencia artificial utilizada, objetivo, enfoque del estudio, procesos cognitivos evaluados, condiciones pedagógicas de uso de la IA y resultados principales. Este proceso permitió una lectura elaborada de cada estudio y produjo homogeneidad en la información recogida, lo que representa un requisito necesario para poder establecer comparaciones pertinentes entre las fuentes bibliográficas revisadas.

La información extraída fue ordenada mediante una matriz de análisis de fuentes bibliográficas que favoreció la comparación de los objetivos, metodologías, resultados y tendencias de cada una de las investigaciones. A continuación, se llevó a cabo una codificación temática que permitió identificar categorías recurrentes, convergencias y divergencias entre estudios, tomando como base los elementos del pensamiento matemático señalados en el marco teórico. Este proceso posibilitó la agrupación de la evidencia en torno a cinco categorías de análisis: disociación entre desempeño mediado y desempeño autónomo, debilitamiento del razonamiento, argumentación y justificación, dependencia de la tecnología y delegación del esfuerzo cognitivo, afectación de la metacognición y la auto-regulación, uso sin mediación pedagógica como factor diferencial de las brechas. Las categorías mencionadas dirigieron la presentación y discusión de los resultados.

Para el desarrollo del estudio se articulan tres fases de trabajo entre sí. La primera fase se refiere a la recuperación documental, en la que se llevaron a cabo las ecuaciones de búsqueda, se hicieron las aplicaciones de los criterios de inclusión y exclusión y se constituyó el corpus final. La segunda fase abarca la etapa de análisis-sistematización de la información extraída y su organización en la matriz documental. Por último, la tercera fase se orientó a la síntesis e interpretación de los resultados, en la cual se constituyeron las categorías analíticas y se estructuraron los resultados que dan cuenta de los objetivos planteados.

Respecto a las consideraciones éticas, dado que la investigación usó únicamente fuentes secundarias de acceso científico, no implicó intervención con personas ni fue necesaria la obtención de consentimiento informado. No obstante, se respetaron de forma rigurosa los principios que hacen referencia al uso ético y responsable de la información científica, garantizando el reconocimiento del trabajo de otros a través de la debida citación de todas las fuentes consultadas, siguiendo las normas APA séptima edición. Se trató de representar los hallazgos descritos en cada investigación de forma fidedigna, evitando en todo momento cualquier distorsión o contextualización errónea de los resultados originales.

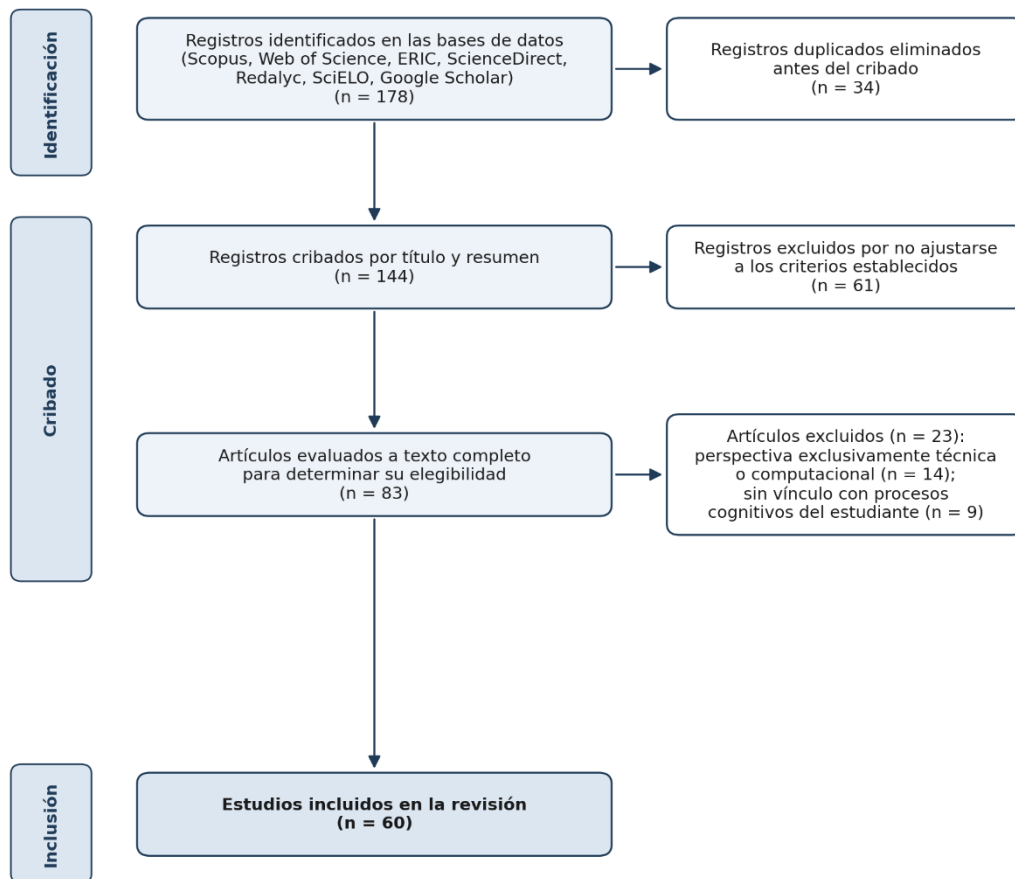
Resultados

El presente capítulo expone los hallazgos de la revisión sistemática en correspondencia con los propósitos formulados para el estudio y con las tres fases declaradas en la metodología. En consecuencia, la presentación se organiza en tres momentos: el primero da cuenta de la recuperación de las fuentes documentales y del proceso de selección que condujo a la conformación del corpus definitivo; el segundo presenta la sistematización y caracterización de los estudios incluidos a partir de la matriz de análisis documental; y el tercero desarrolla la síntesis interpretativa de la evidencia mediante las cinco categorías analíticas construidas, con las cuales se documentan los efectos del uso de la inteligencia artificial sobre el razonamiento, la argumentación y la metacognición matemática, se caracterizan los patrones de dependencia y de delegación del esfuerzo cognitivo y se establecen las condiciones pedagógicas asociadas al incremento o la atenuación de las brechas identificadas. Esta organización permite al lector transitar desde la evidencia del proceso de selección hacia la interpretación integrada de los resultados que responden a la pregunta de investigación.

Resultados etapa uno. Recuperación de las fuentes documentales

La primera etapa del proceso investigativo correspondió a la ejecución de las ecuaciones de búsqueda en las bases de datos seleccionadas, procedimiento que arrojó un total de 178 registros potencialmente pertinentes al objeto de estudio. Sobre este conjunto inicial se aplicaron de manera sucesiva los criterios de inclusión y exclusión definidos en la metodología, mediante un proceso de depuración desarrollado en las fases de identificación, cribado y elegibilidad que garantizó la calidad y la pertinencia del material seleccionado. Este recorrido condujo a la conformación de un corpus definitivo de 60 estudios, tal como muestra el diagrama PRISMA que se presenta en la figura 1.

Figura 1
Diagrama de flujo PRISMA del proceso de selección de los estudios



Nota. La figura ilustra las fases de identificación, cribado e inclusión que condujeron a la conformación del corpus documental de la revisión. Elaboración propia a partir del protocolo PRISMA (Page et al., 2021).

La depuración del corpus documental es mostrada a lo largo de este proceso, que se ve reflejado en el diagrama PRISMA, y sigue el rigor de este protocolo. A partir de los 178 registros, la eliminación de 34 duplicados dejó 144 documentos que fueron objeto de cribado desde la revisión de títulos y resúmenes, excluyéndose 61 registros que no contemplaban la relación entre inteligencia artificial, educación matemática y procesos cognitivos. Sobre los 83 artículos restantes, se evaluaron a texto completo, excluyéndose 23 estudios que seguían un enfoque técnico o computacional, centrados en el rendimiento de modelos matemáticos o sin considerar efectos cognitivos del aprendizaje; lo que ha permitido garantizar un corpus final configurado en relación con la amplitud del proceso de búsqueda y en relación con la precisión de la evidencia recuperada.

Los 60 estudios, incluidos en esta revisión, constituyen un corpus de gran diversidad contextual y metodológica. La producción de los estudios en este campo se concentra en los años intermedios del período abarcado, debido a la expansión de la inteligencia artificial generativa producida tras la aparición pública de los grandes modelos de lenguaje. Las investigaciones se caracterizan por proceder de Asia, Norteamérica, Europa e Iberoamérica. Predominan los estudios elaborados en educación secundaria y educación superior, siendo menores los que se centran en educación primaria. Las herramientas que se analizan son modelos generativos y sistemas de

tutores inteligentes y tecnologías aplicativas en matemáticas. Esta diversidad contribuye a la síntesis, ya que permite analizar las brechas cognitivas en diferentes contextos educativos y con el uso de diferentes tecnologías educativas, aspectos que se presentan con mayor profundidad en los resultados de la segunda etapa.

Resultados de la etapa dos. Sistematización de las fuentes bibliográficas

La segunda etapa del proceso investigativo correspondió a la sistematización de las fuentes bibliográficas mediante la construcción de matrices documentales analíticas. En la matriz principal se registró, para cada uno de los 60 estudios incluidos, la autoría, el año de publicación, la procedencia geográfica, el nivel educativo, la herramienta de inteligencia artificial examinada, el diseño metodológico, los procesos cognitivos evaluados, las condiciones pedagógicas del uso y los hallazgos principales. Esta organización estructurada garantizó la homogeneidad de la información extraída y posibilitó la comparación sistemática entre investigaciones desarrolladas en contextos diversos, condición necesaria para trascender la lectura aislada de cada estudio y avanzar hacia la identificación de patrones comunes en la evidencia.

Se ejecutó un proceso de codificación temática de la matriz consolidada en dos puntos. Primeramente, la codificación que se realizó fue descriptiva de los hallazgos de los estudios tomando como criterios de interpretación los componentes del pensamiento matemático contemplados en la revisión teórica: razonamiento, argumentación, metacognición, autorregulación y transferencia. En el segundo momento, se contrastaron los códigos a partir de la comparación por convergencia y divergencia, que generó el descubrimiento de regularidades transversales a los estudios, pero también los matices asociados a la condición del nivel educativo y a la propia situación de uso de la IA: del proceso de reflexión continua surgieron las agrupaciones temáticas que dan forma a las categorías de la revisión.

El desarrollo de la codificación hizo posible establecer cinco categorías analíticas y representativas: la distinción que se da en el rendimiento asistido y el rendimiento no asistido durante la resolución de problemas matemáticos llevada a cabo por los estudiantes, la afectación de los procesos de razonamiento, argumentación y justificación matemática, la dependencia tecnológica y la delegación sistemática del esfuerzo cognitivo, la afectación de la metacognición y de la autorregulación del aprendizaje matemático, y el uso sin mediación pedagógica como rasgo distintivo de las diferencias cognitivas emergentes. No se habla de compartimentos absolutamente independientes uno de otro, sino que se refieren a dimensiones de un único fenómeno, así como constituyen diferentes ejes que estructuran la síntesis interpretativa que se da en la etapa tres.

Resultados de la etapa tres. Síntesis e interpretación de los hallazgos

La tercera etapa del proceso investigativo corresponde a la síntesis e interpretación de los hallazgos, momento en el cual la evidencia sistematizada se organiza en las cinco categorías analíticas emergentes de la codificación temática. Cada categoría se desarrolla como un apartado que integra los resultados convergentes de los estudios revisados y los interpreta a la luz de los componentes del pensamiento matemático definidos en el referente teórico. Con esta organización se responde a los propósitos del estudio, pues las cuatro primeras categorías documentan la naturaleza de las brechas cognitivas identificadas, mientras la quinta establece las condiciones pedagógicas bajo las cuales dichas brechas se acentúan o se atenúan.

Disociación entre el desempeño asistido y el desempeño autónomo en la resolución de problemas matemáticos

La discrepancia entre las puntuaciones obtenidas mediante asistencia por medio del modelo de inteligencia artificial y las logradas sin la misma es el hallazgo más relevante de la revisión realizada en esta categoría. En este sentido, Bastani et al. (2025) constatan que los estudiantes de secundaria que ya habían practicado la resolución de problemas con apoyo del modelo generativo alcanzaban puntuaciones superiores a las de aquellos compañeros que practicaban la resolución de problemas de forma autónoma, y en el caso de estos últimos, la puntuación alcanzaba niveles inferiores también con respecto a otros estudiantes que no disponían de la IA como apoyo. En la misma línea, Darvishi et al. (2024) documentan una disminución de la calidad de las respuestas cuando el acompañamiento tecnológico cesa, mientras que Akintande (2024) subrayan un descenso en la puntuación obtenida con los mismos ejercicios sin uso de la inteligencia artificial.

La evidencia empírica analizada pone de manifiesto, además, que el beneficio inmediato de la asistencia no se traduce en apropiación ni en transferencia del conocimiento. De hecho, los hallazgos de Fan et al. (2024) validan que los estudiantes que utilizaron ChatGPT (en comparación con otros) aumentaron su producción, pero no lograron mejoras significativas en el aprendizaje conseguido con otras modalidades de apoyo, hallazgo que Sánchez-Ruiz et al. (2023), apoyan desde el contexto de las matemáticas universitarias, al observar que el rendimiento asistido no se traduce en la apropiación real de competencias transferibles. En esa misma línea, Kaya y Yavuz (2025) corroboran que en tareas procedimentales la ventaja de estos sistemas flaquea en las actividades de mayor exigencia conceptual que son las que caracterizan el rendimiento matemático autónomo.

Los estudios analizados coinciden en atribuir esta brecha a la disminución del esfuerzo cognitivo provocada por el uso instrumental de la IA. Stadler et al. (2024) concluyen que los modelos de lenguaje disminuyen la carga mental de la tarea, pero producen justificaciones de escasa calidad asociadas a un procesamiento superficial del contenido. En su parte, Abbas et al. (2024) identifican que el corto uso favorece atajos en el trabajo que provocan más procrastinación y disminución de la retención; mientras que Fernandes et al. (2026) observan que el rendimiento asistido se da en paralelo a un mal calibrado, porque los estudiantes sobrevaloran sus logros y aceptan las respuestas del propio sistema sin someterlas a examen crítico.

La síntesis de los estudios analizados comprueba la transversalidad del fenómeno, dejando matices que son importantes. La revisión sistemática de Zhai et al. (2024), de forma complementaria, incide en la idea de que la sobredependencia de los sistemas conversacionales afecta el desarrollo del pensamiento crítico, la toma de decisiones y la forma en la que se puede llegar a enfrentarse a situaciones nuevas de una manera autónoma, abriendo la distancia entre el desempeño con ayuda y el desempeño sin esta. No obstante, en contraposición, la revisión sistemática de Deng et al. (2025) lleva a cabo un análisis mostrando efectos heterogéneos y hasta contradictorios en función del contexto educativo y del propio diseño de la intervención, señalando que la disociación no es una consecuencia inevitable de la tecnología de soporte, sino que es el efecto de las propias condiciones de uso.

A nivel global, se puede afirmar que un desempeño asistido sin autonomía no es sinónimo de aprendizaje matemático tal como han indicado los referentes teóricos desde la transferencia como del andamiaje. La distancia resulta sensible a la rigurosidad del diseño pedagógico, pues Lee et al. (2024), ponen de manifiesto que una orientación con preguntas, pistas y realimentación gradual a partir de soluciones propuestas, requiere un diseño previo de la respuesta del estudiante, favorece la autorregulación y la retención frente a la entrega de las soluciones. Esto desplaza el

problema de la IA en la intencionalidad de su uso y anticipa la condición de diferenciación que se presenta en la quinta categoría analizada.

Debilitamiento de los procesos de razonamiento, argumentación y justificación matemática

El análisis de la literatura académica permite constatar cómo la brecha documentada no sólo se limita al rendimiento observable, sino que también tiene que ver con los procesos que son constitutivos del pensamiento matemático. Tal y como lo evidencian Dilling y Herrmann (2024) al trabajar con la IA para construir demostraciones geométricas, la mayor parte de los estudiantes solicitaban la prueba completa y sólo se limitaban a copiar o regenerar respuestas, con escasas preguntas orientadas a comprender o justificar los procedimientos. Hallazgos similares se reportan en el estudio de Yoon et al. (2024), quienes constatan que los estudiantes siguen las estrategias argumentativas indicadas por la IA aunque no entiendan del todo su procedimiento, lo que, por una parte, disminuye su elaboración de las demostraciones y, por la otra, merma el ejercicio del razonamiento deductivo.

El deterioro de la capacidad validativa es muy importante en la medida en que la IA tiene un carácter falible. Plevris et al. (2023) y Hasib e Islam (2025) demuestran que los diversos modelos conversacionales cometen errores con las operaciones sencillas e, incluso, son más propensos a fallar dado que la inferencia es más compleja, mientras que Wardat et al. (2023) hacen notar que la IA presenta un lenguaje matemático convincente y, aun así, comete errores conceptuales en álgebra, cálculo y geometría. En este contexto, Dahlkemper et al. (2023) muestran que, cuando los estudiantes no tienen un dominio conceptual fuerte, no logran aclarar las respuestas erróneas de IA de las preguntas correctas, puesto que la calidad lingüística crea una ilusión de comprensión que sustituye el juicio fundado.

Los estudios revisados coinciden en atribuir este debilitamiento a este desplazamiento cognitivo derivado del uso intensivo de la IA. Así, Gerlich (2025), en una amplia investigación, muestra la existencia de una relación inversamente proporcional a la frecuencia del uso y al pensamiento crítico, relación que es mediada por la delegación de las tareas intelectuales y acentuada en personas más jóvenes y con menor formación. De hecho, los propios estudiantes dan cuenta del fenómeno, ya que Darwin et al. (2024) y Lo et al. (2024) muestran construcciones desde las que los propios estudiantes afirman que la IA impide el propio análisis y evaluación de pruebas, mientras que Stadler et al. (2024) muestran cómo quienes habían utilizado IA presentaron explicaciones menos sólidas, con menor implicación en procesos de integración e inferencia.

El análisis de las investigaciones, apoyan el alcance del fenómeno y detallan la graduación de este. La revisión de Zhai et al. (2024) de una parte advierte que la aceptación acrítica de los contenidos generados, agravada por las alucinaciones y los sesgos de la IA, tiene como consecuencia que se pierda la elaboración de sus propios argumentos, tal y como Zhai et al. (2024) documentan que una muestra de futuros docentes, no son capaces de detectar respuestas incorrectas de la IA. En una dirección complementaria, Essien et al. (2024) subrayan que las mejoras comprometidas con las inteligencias artificiales solamente se ven justificadas, por así decirlo, por los niveles cognitivos más bajos de recordar, comprender, y aplicar, se ven moderadas en analizar y evaluar y no logran alcanzar la creación de conocimiento original, justo cuando la argumentación matemática se sitúa en la cúspide.

Por otra parte, los datos académicos corroboran que es posible alcanzar la respuesta correcta al mismo tiempo que disminuyen las posibilidades de justificarla y de argumentarla, disociación que hace que la respuesta no tenga una característica formativa, dado que una solución por sí misma, exenta de una argumentación que la avale, no es en consecuencia actividad

matemática. El deterioro, no es una condición de necesidad, ya que Essel et al. (2024) muestran que utilizar la IA en actividades argumentadas e instrumentadas potencia el pensamiento crítico, la apertura crítica y el escepticismo reflexivo respecto a los estudiantes que no adoptan esta tecnología. Esta evidencia, contraria a la deducción del deterioro, anticipa la condición didáctica estudiada en la quinta categoría mostrando el mecanismo que fundamenta la relación del deterioro y la delegación sistemática del esfuerzo cognitivo.

Dependencia tecnológica y delegación sistemática del esfuerzo cognitivo

Las evidencias empíricas que se han estudiado y analizado avalan como práctica habitual, lo que el referente teórico pone de manifiesto como riesgo, dado que dicho corpus documenta patrones de uso por los cuales la IA es la primera opción, no la última alternativa. Taylor et al. (2026) observaron en tiempo real a estudiantes que transcribían de forma directa las versiones de los enunciados, y que, además, rara vez comprobaban las respuestas, validando incluso aquellas que eran erróneas. De la misma manera, Šedlbauer et al. (2024) y Ding et al. (2026) documentan el pronto encuadre del modelo de IA como recurso habitual desde la primera experiencia, con una confianza no cuestionada en el tipo de respuestas obtenidas; mientras que Bastani et al. (2025) registran la tendencia a pedir las soluciones de forma directa, pero reconociendo una menor reflexión y verificación de los procedimientos.

La manera de entender el uso irreflexivo se establece como la que caracteriza este tipo de delegación. Basándose en una amplia muestra de estudiantes universitarios, Hou et al. (2025) identifican este tipo de comportamiento marcado por la aceptación sin análisis de las respuestas que genera la IA, y establecen que el pensamiento crítico sería el principal factor que actuaría contra dicha aceptación. En esta línea, la revisión de Zhai et al. (2024) sostiene que la facilidad que tienen estos sistemas de respuesta para generar soluciones completas favorece los atajos cognitivos en lugar de las soluciones autónomas, utilizando la ayuda como apoyo permanente. La observación de los educadores coincide puesto que Liu y Zhang (2025) dan cuenta de episodios reiterados donde el estudiante acepta sin cuestionar las soluciones que produce la IA, manifestando que la delegación se ha quedado anclada como un hábito.

Por otra parte, los estudios revisados también previenen que la delegación puesta en marcha distorsiona la concepción que tiene el sujeto de su propio aprendizaje. Jia et al. (2025) sugieren que la dependencia de la IA generativa predice una autorreferencia inflada, asociada a un efecto Dunning-Kruger, que se manifiesta en menor rendimiento. En el plano conceptual, Herman y Lara-Steidel (2025) reseñan falsos positivos de aprendizaje para referirse a la situación en la que se aprueba la evaluación sin dar cabida a un entendimiento e inversión cognitiva antes de este punto. Esta apariencia se pierde cuando se cambian las condiciones, dado que Akintande (2024) demuestra que los hábitos de copiar la pregunta y transcribir la respuesta colapsan en el momento en que el secuenciado de las preguntas exige un razonamiento que la IA no puede proporcionar directamente.

La gradación de la delegación complementa el análisis porque la evidencia discrimina niveles de profundización de la externalización. En el nivel procedimental, Assamah et al. (2026) refieren que algunos docentes de matemáticas reconocen que la dependencia de los recursos digitales, reduce la práctica del cálculo mental. En un nivel intermedio, Kim et al. (2025) constatan la delegación de la validación de manera que más de la mitad de los futuros docentes no son capaces de identificar los errores de los problemas generados por la IA. En el nivel más profundo, Ng et al. (2024b) dan cuenta de estudiantes que siguen las instrucciones de la IA como caja negra,

sin entender principios delegados, externalización ya no del cálculo, sino del completo razonamiento.

En términos generales, estos resultados corroboran el mecanismo que vincula las brechas que se han descrito. La externalización cognitiva ya no sería una estrategia instantánea provechosa, sino una rutina que elimina el esfuerzo del cual depende el aprendizaje. La distinción tiene sus consecuencias, porque la disminución de la carga cognitiva ayuda a aprender cuando permite liberar recursos para el entendimiento; no obstante, frena la comprensión cuando suprime el esfuerzo deseable, el cual permite la construcción del conocimiento matemático. Una vez abandonada la delegación, el estudiante va perdiendo de forma progresiva el control consciente de su propio proceso de resolución, deterioro que da cuenta de la dimensión más sutil de las brechas que han sido descritas.

Afectación de la metacognición y de la autorregulación del aprendizaje matemático

La cara más silenciosa de las brechas emergentes en el corpus menoscaba precisamente procesos que la evidencia señala como predictores del desempeño matemático. Así, Zhang et al. (2024) pone a prueba que el conocimiento y la supervisión metacognitivos tienen una correlación directa sobre las tareas de resolución de problemas, permitiendo la selección de estrategias, el control del progreso o el detalle del error, mientras que Sutarto et al. (2022) defienden que la planificación, supervisión y evaluación consciente de las acciones propias es lo que diferencia a los que construyen aprendizajes matemáticos autónomos. En este sentido, el uso no mediado de la IA contamina precisamente ese control consciente, deterioro que la calificación no refleja porque el rendimiento asistido lo oculta.

Por su parte, la calibración entre confianza y desempeño es donde más se hace visible la evidencia. Ganuthula y cols. (2026) formalizan una miscalibración metacognitiva en la que los usuarios intensivos de modelos de lenguaje sobredimensionan su comprensión y dependencia, dado que las respuestas elaboradas cubren los vacíos del aprendizaje, distorsión que se ve amplificada por la sobreconfianza que tienen los propios sistemas, incapaces de realizar un reconocimiento de sus límites. De la misma forma, Liu et al. (2023) advierten que la confianza que se puede depositar por parte de los estudiantes ante la IA implica dejarse llevar ante respuestas incorrectas y conlleva también una menor motivación a querer investigar por su cuenta, por su parte Hou et al. (2025) afirman que una mayor confianza conlleva copiar las soluciones sin ninguna noción metacognitiva.

A esta distorsión se suma una limitación estructural, ya que las tareas que la IA genera espontáneamente rara vez convocan la metacognición. Sawyer y Aga (2025) constatan que tan sólo el 3.1 % de las actividades matemáticas generadas por diferentes IA alcanzó el nivel de hacer matemáticas, precisamente la categoría de demanda cognitiva que requiere planificación, monitoreo y reflexión, concentración que sólo se corrige mediante ingeniería de indicaciones e intervención docente. De manera complementaria, Chee et al. (2024) argumentan que la mentalidad autorreflexiva respecto a la IA es una habilidad que debe desarrollarse de forma intencional, ya que sin ella la verificación se pierde frente a la aceptación automática, comprometiendo la resolución autónoma de problemas.

Sin embargo, el impacto no debe considerarse un fin inevitable, pues la recopilación documental analizada también constata construcciones que hacen retornar al estudiante hacia el ciclo autorregulatorio. Ng et al. (2024a) indican que un chatbot configurado en un entorno de IA exclusivamente orientado hacia la fijación de una meta y de retorno logra mejorar la autorregulación en comparación con el sistema automático predefinido. Olmo-Muñoz et al. (2022)

y Son (2024) demuestran cómo, para los sistemas de tutoría inteligente, los estudiantes deben dar soluciones alternativas a las que proponen los sistemas y que, por tanto, deben controlar ante todo el pensamiento; mientras que Richter et al. (2025) destacan que los agentes conversacionales logra que se reflexione y autoevalúe, aunque la advertencia es que las interacciones con estos sistemas podrían sobredimensionarse haciendo que la forma de resolución no sea posible, sino que se objective.

En síntesis, la variable que inclina la balanza no es la aparición de la IA, sino la inminente posibilidad de salida de la metacognición en la medida que puede ser absorbida por el sistema, o bien, devuelta al estudiante. En este sentido, Ruiz Recio et al. (2026) sostienen que la tecnología actúa como un espejo metacognitivo toda vez que la realimentación inmediata posibilitaba la exteriorización y la rectificación de los errores y, Dúo-Terrón et al. (2026) señalan que el hecho de ubicar al estudiante en la posición del evaluador crítico de los propios modelos contribuye a afianzar la autorregulación del aprendizaje del área. Los mismos autores añaden el descubrimiento de una brecha de formación que privaba a muchos de esas mismas herramientas metacognitivas que se dotan. Esta misma condición que recorre las cuatro brechas constatadas aboca a la última categoría; donde la mediación pedagógica es el elemento cuya diferenciación suscita.

El uso sin mediación pedagógica como condición diferenciadora de las brechas cognitivas

El corpus converge en un descubrimiento condicional que da sentido a las cuatro brechas anteriores, ya que son solo las que aparecen cuando el uso de la IA es libre e instrumental y se atenúan cuando interrumpe un diseño didáctico. La síntesis de Almarashdi et al. (2024) concluye que los condicionales de la IA en matemáticas dependen, en gran medida, de la guía docente, mientras que el uso libre se vincula a la dependencia y a experimentar problemas ante preguntas complejas, y Pepin et al. (2025) avalan que sin andamiaje instruccional se ven debilitados el pensamiento crítico y la autorregulación. Como evidencia directa Noster et al. (2024), documentan que los futuros educadores en ocasiones aceptaron como válidas respuestas incorrectas emitidas por la IA.

La gravedad del problema mencionado puede explicarse, entre otros motivos, por un vacío estructural del propio campo de la investigación. Topali et al. (2025) demuestran que, de los estudios sobre IA en educación básica, casi la mitad no tienen fundamentación pedagógica y no se contextualizan didácticamente, lo cual modifica los efectos cognitivos descritos y puede aumentar las diferencias entre estudiantes. En la misma línea, Panqueban y Huincahue (2024) advierten que la mayoría de las investigaciones en educación matemática no tienen en cuenta un propio marco de la didáctica en el área y atribuyen a la IA el rol de un entorno de trabajo autónomo, obviando el trabajo mediado por el docente que la integración efectiva de la IA requiere.

Bajo un entorno pedagógico específico, la evidencia empírica indica que las brechas no solo se mitigan sino se revierten. Drushlyak et al. (2025) desarrollaron un experimento en el que la IA produjo respuestas erróneas buscando que los docentes las identificaran y argumentaran la mejor forma de corregirlas, logrando así cambios medibles de la calidad del pensamiento crítico en la mayoría de los estudiantes participantes. Pavlova (2024) invirtieron el rol buscando que el estudiante fuera quien articulara las preguntas. Huda et al. (2025) y Jang et al. (2025) presentan un andamiaje para la interacción con la IA contextualizada en una indagación orientada, que favorece la elaboración de la planificación, el monitoreo y la evaluación durante la aprehensión conceptual de las matemáticas.

La intervención consciente del docente sobre el diseño de la actividad es el factor que comparten todas estas experiencias. De un lado, Gouia-Zarrad y Gunn (2024) encuentran que la

IA produce aprendizajes más profundos si está integrada en proyectos de modelado mediado y validado por el profesor, y por otro, Kim et al. (2025) observan que un andamiaje estructurado en actividades de planteamiento de problemas incrementa el reconocimiento y la corrección de errores respecto al uso instrumental. La mediación también es equitativa, ya que Etcuban (2025) y Kolade et al. (2024) considera que sin mediación la eficacia de la IA queda supeditada al conocimiento previo de cada estudiante, con lo cual el uso autónomo de la IA aumenta, y no disminuye las diferencias cognitivas.

En conjunto, la evidencia respalda la tesis que da lugar a esta revisión y que postula que el problema no es la inteligencia artificial sino la ausencia de intencionalidad formativa en su uso. Lo cual sintetizan Kock et al. (2025) al presentar los esquemas de uso que van de lo instrumental a lo reflexivo, lo cual solo deviene cuando está presente la reflexión sobre la práctica propia. Observado desde el andamiaje, la conclusión es clara, el apoyo de la IA es válido cuando es temporario y estructurado y se convierte en brecha cuando es permanente y hay sustitución del pensamiento matemático por la herramienta. La condición que marca la diferencia es la concepción pedagógica de uso.

Discusión

Los resultados que afloran a partir del análisis de la literatura académica reciente muestran una notoria desconexión entre el rendimiento académico que obtienen los estudiantes mediante la inteligencia artificial por un lado y mediante una solución independiente por otro. En este sentido, Bastani et al. (2025) indican que algunos estudiantes rinden mejor que sus compañeros en la práctica apoyada, pero son peores cuando la práctica es independiente, resultado que también recuperan Darvishi et al. (2024) y Akintande (2024) en otros marcos de la enseñanza formal. Esta evidencia sugiere y da fuerza a la idea que el aprendizaje matemático no se agota en una respuesta correcta, sino que debe ir acompañado de comprensión y justificación que la ayuda externa puede sustituir sin garantía de apropiación (Schoenfeld, 1985; OECD, 2023).

En este sentido, la diferencia se hace especialmente evidente cuando se debe solucionar de forma autónoma algún tipo de problema. Al hacer desaparecer el soporte técnico de la IA, se descubre la carga cognitiva. Fan et al. (2024) subrayan que la IA facilita las producciones inmediatas y no ayudar al aprendizaje transferido. Sánchez-Ruiz et al. (2023) exponen evidencias que dan cuenta que los rendimientos ayudados no están vinculados con la apropiación de las competencias matemáticas. Kaya y Yavuz (2025) respaldan que la superioridad del soporte de la tecnología en relación con las actividades propias de la solución de los problemas tiende a desaparecer en los problemas más complejos desde la perspectiva conceptual. La prontitud en la solución de los procedimientos puede aumentar cada vez que la comprensión conceptual que fomenta un aprendizaje matemático auténtico puede disminuir (Kilpatrick et al., 2001).

En efecto, exactamente aquello que se mantiene en esta brecha es el centro neurálgico del pensamiento matemático: el razonamiento, la argumentación y la justificación. Dilling y Herrmann (2024) evidencian que, al construir demostraciones con la IA, los estudiantes o bien piden la prueba completa o bien reproducen la misma. Al respecto, Yoon et al. (2024) comprueban el uso de estrategias argumentativas sin suficiente comprensión. En los casos donde esta capacidad es limitada el riesgo aumenta, pues los estudiantes tienen dificultades en la identificación de las respuestas incorrectas que puede generar IA (Wardat et al., 2023; Dahlkemper et al., 2023). En consecuencia, cuando la respuesta es válida, se constituye en formativa, pero el estudiante no puede explicar ni justificar el porqué de su validez (Cervantes-Barraza et al., 2020).

Esta sustitución de los procesos cognitivos adquiridos puede seguir profundizándose por la dependencia de la tecnología y por la sistematización de la delegación del esfuerzo intelectual.

Taylor et al. (2026) observan estudiantes que transcriben los enunciados en las herramientas de IA sin recuperar la respuesta verificada por esta tecnología, mientras que Šedlbauer et al. (2024) documentan la aparición de la confianza sin cuestionarla en las primeras experiencias de uso. La externalización alcanza incluso a los procedimientos de comprensión, ya que Ng et al. (2024b) muestran estudiantes que siguen las instrucciones de la IA, que consideran como una caja negra, sin comprender los principios que subyacen. Jia et al. (2025) adicionan que la posible dependencia también se puede asociar con un exceso de autoeficacia y un bajo rendimiento académico real.

Igualmente, la delegación sostenida del esfuerzo cognitivo implica procesos metacognitivos esenciales para regular el aprendizaje matemático, ya que Zhang et al. (2024) evidencian que el monitoreo metacognitivo influye en la resolución de problemas al igual que Turgut y Uğurlu (2024), quienes abogan por incluir la planificación, el monitoreo y la evaluación. Esta relación es de especial interés si se tiene en cuenta el desajuste entre percepción y rendimiento identificado por Ganuthula et al. (2026), quienes subrayan que los usuarios intensivos sobreestiman su comprensión y subestiman su dependencia, por su parte Liu et al. (2023) advierten que la seguridad en la IA puede facilitar una aceptación por parte del usuario, ocultando bajo el rendimiento asistido el debilitamiento del control consciente sobre el aprendizaje.

Por otro lado, la evidencia no permite atribuir estas brechas a la sola presencia de la inteligencia artificial, los niveles de las brechas estarán determinados por las condiciones pedagógicas que lo acompañan. Almarashdi et al. (2024) concluyen que los beneficios dependen de la guía del docente, mientras que el uso libre se relaciona con dependencia y dificultades ante preguntas complejas. De igual modo, Pepin et al. (2025) muestran que, sin andamiaje instruccional, el propio pensamiento crítico o autorregulación se ven debilitados. Topali et al. (2025) explican en parte esta situación al señalar la escasa fundamentación pedagógica de muchos estudios. Al contrario, Huda et al. (2025); Drushlyak et al. (2025) ponen de manifiesto que el andamiaje docente podría revertir estas tendencias.

Bajo este enfoque, lo que identifica este trabajo de revisión es la posibilidad de desplazar la proyección de la brecha cognitiva desde la tecnología hacia las formas de uso que se construyen a su alrededor. La evidencia que se sintetiza permite afirmar que dicha brecha no es la consecuencia natural de la inteligencia artificial sino que brota de los procesos cognitivos que se delegan, de la frecuencia de dicha delegación, así como de las condiciones pedagógicas que la rigen. Kock et al. (2025) muestran que los esquemas de uso van desde lo instrumental hasta las prácticas reflexivas con mediación docente; en este sentido, la revisión recupera el vacío señalado por Tlili (2023) y Crompton y Burke (2023), ya que el reto no es la posibilidad de restringir a la IA sino que, como sistema asistido, esta no reemplace el razonamiento que con la educación matemática se debe desarrollar.

Conclusiones

La revisión sistemática permite concluir que el uso de la inteligencia artificial en el aprendizaje de las matemáticas provoca una disociación consistente entre el rendimiento que alcanzan los estudiantes haciendo uso de una asistencia tecnológica y el que logran cuando resuelven problemas de forma autónoma. La evidencia analizada muestra que esta ventaja se concentra sobre tareas procedimentales y se diluye ante exigencias de mayor demanda conceptual. A la par de esta disociación, los estudios muestran lo que denominan como un debilitamiento del razonamiento, la argumentación y la justificación matemática, dado que los estudiantes son propensos a aceptar y reproducir soluciones generadas por la IA, sin someterlas a un examen crítico, lo cual compromete el valor formativo de las respuestas alcanzadas.

Una segunda conclusión importante considera que la dependencia tecnológica puede incrementar la delegación sistemática del esfuerzo cognitivo, una pauta que se va fijando progresivamente como hábito y que puede llegar a niveles profundos de externalización, incluso del razonamiento completo. La metacognición y la autorregulación del aprendizaje se ven fuertemente afectadas porque los estudiantes suelen sobrevalorar la IA en cuanto a su comprensión real y se infravaloran en cuanto a su dependencia. Esta miscalibración es especialmente peligrosa en tanto que el rendimiento asistido puede ocultar vacíos formativos que difícilmente se ponen de manifiesto en la evaluación tradicional.

Sin embargo, la evidencia analizada permite afirmar que estas brechas no representan una consecuencia inevitable de la inteligencia artificial, sino que dependen fundamentalmente de las condiciones pedagógicas en que se lleve a cabo su uso. Los estudios que contemplan mediación docente, realimentación y actividades estructuradas evidencian que el pensamiento crítico, la argumentación y la autorregulación pueden reforzarse incluso con el uso de la IA; mientras el uso libre e instrumental sin ninguna intención formativa- tiende a acentuar la diferencia cognitiva entre estudiantes en función de su conocimiento previo, lo cual ratifica que la mediación pedagógica es un factor central de diferenciación.

Con lo cual se puede concluir que el problema que justifica este estudio no es la inteligencia artificial como tal, sino la falta de una intencionalidad formativa que pueda conducir a la construcción del pensamiento matemático. Esta conclusión supone un marco de referencia de interés para los educadores y las instituciones educativas porque hace ver que la inteligencia artificial tendría que ofrecer abiertamente un andamiaje del razonamiento de forma temporalmente estructurada que fuera retirándose; en vez de ser un recurso que sustituye a la persona. En un futuro próximo se podrían ofrecer diseños pedagógicos que permitan articular esta transición desde el uso instrumental hacia el uso reflexivo, constituyendo así un uso más autónomo.

Referencias bibliográficas

- Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*, 21, Artículo 10. <https://doi.org/10.1186/s41239-024-00444-7>
- Akintande, O. J. (2024). Artificial versus natural intelligence: Overcoming students' cheating likelihood with artificial intelligence tools during virtual assessment. *Future in Educational Research*, 2(2), 147–165. <https://doi.org/10.1002/fer3.33>
- Almarashdi, H. S., Jarrah, A. M., Abu Khurma, O., & Gningue, S. M. (2024). Unveiling the potential: A systematic review of ChatGPT in transforming mathematics teaching and learning. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(12), em2555. <https://doi.org/10.29333/ejmste/15739>
- Assamah, G., Sualisu, A., Gabriel-Wetty, F. K. N., Asomah, R. K., Abonyi, U. K., Ahwireng, D., & Kaura, N. (2026). Integrating ICT in Mathematics Pedagogy: A Contextual Analysis of Pre-Tertiary Teachers' Practices in Ghana. *Human Behavior and Emerging Technologies*, 2026(1), Artículo 5577487. <https://doi.org/10.1155/hbe2/5577487>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, 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>

- Cervantes-Barraza, J. A., Hernandez Moreno, A. y Rumsey, C. (2020). Promoting mathematical proof from collective argumentation in primary school. *School Science and Mathematics*, 120(1), 4-14. <https://doi.org/10.1111/ssm.12379>
- Chee, H., Ahn, S., & Lee, J. (2024). A competency framework for AI literacy: Variations by different learner groups and an implied learning pathway. *British Journal of Educational Technology*, 56(5), 2146-2182. <https://doi.org/10.1111/bjet.13556>
- Crompton, H., and Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20(1), Article 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Dahlkemper, M. N., Lahme, S. Z., & Klein, P. (2023). How do physics students evaluate artificial intelligence responses on comprehension questions? A study on the perceived scientific accuracy and linguistic quality of ChatGPT. *Physical Review Physics Education Research*, 19(1), 010142. <https://doi.org/10.1103/PhysRevPhysEducRes.19.010142>
- Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education*, 210, 104967. <https://doi.org/10.1016/j.compedu.2023.104967>
- Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1), 2290342. <https://doi.org/10.1080/2331186X.2023.2290342>
- 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>
- Dilling, F., & Herrmann, M. (2024). Using large language models to support pre-service teachers mathematical reasoning—an exploratory study on ChatGPT as an instrument for creating mathematical proofs in geometry. *Frontiers in Artificial Intelligence*, 7, 1460337. <https://doi.org/10.3389/frai.2024.1460337>
- Ding, H., Shen, Y., Chen, J., and Wang, P. (2026). More vs. less cognitive offloading from AI assistants: Impacts on novices' collaborative performance and skill development. *Information Processing & Management*, 63, 105046. <https://doi.org/10.1016/j.ipm.2026.105046>
- Drushlyak, M., Lukashova, T., Shamonina, V., & Semenikhina, O. (2025). ChatGPT-based simulation helps to develop the pre-service mathematics teachers' critical thinking. *International Journal of Instruction*, 18(1), 153–172. <https://doi.org/10.29333/iji.2025.1819a>
- Dúo-Terrón, P., Rodríguez-García, J. D., Robles-Martínez, G., & Moreno-Guerrero, A. J. (2026). Artificial intelligence literacy: Scientific impact of LearningML software. *Computer Applications in Engineering Education*, 34(2), e70162. <https://doi.org/10.1002/cae.70162>
- Duval, R. (2006). A cognitive analysis of problems of comprehension in a learning of mathematics. *Educational Studies in Mathematics*, 61(1–2), 103–131. <https://doi.org/10.1007/s10649-006-0400-z>
- Ellis, A. B., Lockwood, E., Tillema, E. y Moore, K. (2022). Generalization across multiple mathematical domains: Relating, forming, and extending. *Cognition and Instruction*, 40(3), 351-384. <https://doi.org/10.1080/07370008.2021.2000989>
- Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based

- conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, 6, 100198. <https://doi.org/10.1016/j.caeai.2023.100198>
- Essien, A., Bukoye, O. T., O'Dea, X., & Kremantzis, M. (2024). The influence of AI text generators on critical thinking skills in UK business schools. *Studies in Higher Education*, 49(5), 865–882. <https://doi.org/10.1080/03075079.2024.2316881>
- Etcuban, J. O. (2025). The use of ChatGPT in addressing Algebra anxiety and promoting confidence. *International Electronic Journal of Mathematics Education*, 20(2), em0825. <https://doi.org/10.29333/iejme/16008>
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology*, 56(2), 489–530. <https://doi.org/10.1111/bjet.13544>
- Fernandes, D., Villa, S., Nicholls, S., Haavisto, O., Buschek, D., Schmidt, A., Kosch, T., Shen, C., & Welsch, R. (2026). AI makes you smarter but none the wiser: The disconnect between performance and metacognition. *Computers in Human Behavior*, 175, 108779. <https://doi.org/10.1016/j.chb.2025.108779>
- Ganuthula, V. R., Kumari, N., & Chakraborty, P. (2026). The curiosity erosion model: A theoretical framework for understanding the impact of large language models on human curiosity. *Human Behavior and Emerging Technologies*, 2026(1), 6725387. <https://doi.org/10.1155/hbe2/6725387>
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. <https://doi.org/10.3390/soc15010006>
- Gouia-Zarrad, R., & Gunn, C. (2024). Enhancing students' learning experience in mathematics class through ChatGPT. *International Electronic Journal of Mathematics Education*, 19(3), em0781. <https://doi.org/10.29333/iejme/14614>
- Hasib, M., & Islam, M. S. (2025). How University students in Bangladesh engage with ChatGPT: A qualitative study. *PLOS ONE*, 20(9), e0333089. <https://doi.org/10.1371/journal.pone.0333089>
- Herman, J., & Lara-Steidel, H. (2025). Artificial Intelligence on Campus: Revisiting Understanding as an Aim of Higher Education. *Educational Theory*, 75(4), 603–625. <https://doi.org/10.1111/edth.70026>
- Hou, C., Zhu, G., & Sudarshan, V. (2025). The role of critical thinking on undergraduates' reliance behaviours on generative AI in problem-solving. *British Journal of Educational Technology*, 56(5), 1919–1941. <https://doi.org/10.1111/bjet.13613>
- Huda, N., Anwar, K., Novferma, & Kurniawan, W. (2025). ChatGPT scaffolding in supporting metacognition for limit concepts in guided inquiry mathematics learning. *Journal of Information Technology Education: Research*, 24, Article 38. <https://doi.org/10.28945/5645>
- Jang, H., Kim, S., & So, H.-J. (2025). Technology-based scaffolding for ill-structured problem-solving: A systematic review and emerging directions with AI. *Learning: Research and Practice*. <https://doi.org/10.1080/23735082.2025.2591403>
- Jia, W., Pan, L., & Neary, S. (2025). Effect of GenAI Dependency on University Students' Academic Achievement: The Mediating Role of Self-Efficacy and Moderating Role of Perceived Teacher Caring. *Behavioral Sciences*, 15(10), Artículo 1348. <https://doi.org/10.3390/bs15101348>

- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeiffer, F., 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, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kaya, D., & Yavuz, S. (2025). Can generative AI and ChatGPT break human supremacy in mathematics and reshape competence in cognitive-demanding problem-solving tasks? *Journal of Intelligence*, 13(4), 43. <https://doi.org/10.3390/jintelligence13040043>
- Kilpatrick, J., Swafford, J., and Findell, B. (Eds.). (2001). Adding it up: Helping children learn mathematics. National Academy Press. <https://doi.org/10.17226/9822>
- Kim, Y. R., Park, M. S., & Joung, E. (2025). Exploring the integration of artificial intelligence in math education: Preservice Teachers' experiences and reflections on problem-posing activities with ChatGPT. *School Science and Mathematics*, 126(1), 9–23. <https://doi.org/10.1111/ssm.18336>
- Kock, Z. J., Salinas-Hernández, U., & Pepin, B. (2025). Engineering students' initial use schemes of ChatGPT as an instrument for learning. *Digital Experiences in Mathematics Education*, 11, 192–218. <https://doi.org/10.1007/s40751-025-00169-w>
- Kolade, O., Owoseni, A., & Egbetokun, A. (2024). Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework. *Heliyon*, 10, (1-10). <https://doi.org/10.1016/j.heliyon.2024.e31533>
- Lee, H.-Y., Chen, P.-H., Wang, W.-S., Huang, Y.-M., & Wu, T.-T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21, Artículo 16. <https://doi.org/10.1186/s41239-024-00447-4>
- Liu, M., Ren, Y., Nyagoga, L. M., Stonier, F., Wu, Z., & Yu, L. (2023). Future of education in the era of generative artificial intelligence: Consensus among Chinese scholars on applications of ChatGPT in schools. *Future in Educational Research*, 1(1), 72-101. <https://doi.org/10.1002/fer3.10>
- Liu, Y., & Zhang, Y. (2025). Research on University Teachers' Perceptions of Gen-AI From the Perspective of Actor-Network Theory. *Complexity*, 2025, Artículo 8885265. <https://doi.org/10.1155/cplx/8885265>
- Lo, C. K., Hew, K. F., & Jong, M. S.-Y. (2024). The influence of ChatGPT on student engagement: A systematic review and future research agenda. *Computers & Education*, 219, 105100. <https://doi.org/10.1016/j.compedu.2024.105100>
- Mohebi, L. (2024). Empowering learners with ChatGPT: Insights from a systematic literature exploration. *Discover Education*, 3, 36. <https://doi.org/10.1007/s44217-024-00120-y>
- National Council of Teachers of Mathematics. (2023). Mathematical processes. <https://www.nctm.org>
- Ng, D. T., Tan, C. W., & Leung, J. K. (2024a). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology*, 55(4), 1328-1353. <https://doi.org/10.1111/bjet.13454>
- Ng, D. T., Xinyu, C., Leung, J. K., & Chu, S. K. (2024b). Fostering students' AI literacy development through educational games: AI knowledge, affective and cognitive

- engagement. *Journal of Computer Assisted Learning*, 40(5), 2049–2064. <https://doi.org/10.1111/jcal.13009>
- Noster, N., Gerber, S., & Siller, H.-S. (2024). Pre-service teachers' approaches in solving mathematics tasks with ChatGPT. *Digital Experiences in Mathematics Education*, 10, 543–567. <https://doi.org/10.1007/s40751-024-00155-8>
- Olmo-Muñoz, J., González-Calero, J. A., Diago, P. D., Arnau, D., & Arevalillo-Herráez, M. (2022). Using intra-task flexibility on an intelligent tutoring system to promote arithmetic problem-solving proficiency. *British Journal of Educational Technology*, 53(6), 1976–1992. <https://doi.org/10.1111/bjet.13228>
- Organisation for Economic Co-operation and Development (OECD). (2023). PISA 2022 assessment and analytical framework: Mathematics, reading, science and financial literacy. OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>
- 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, Article n71. <https://doi.org/10.1136/bmj.n71>
- Panqueban, D., & Huincahue, J. (2024). Artificial intelligence in mathematics education: A systematic review. *Uniciencia*, 38(1), 357–373. <https://doi.org/10.15359/ru.38-1.20>
- Pavlova, N. H. (2024). Flipped dialogic learning method with ChatGPT: A case study. *International Electronic Journal of Mathematics Education*, 19(1), em0764. <https://doi.org/10.29333/iejme/14025>
- Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). A scoping survey of ChatGPT in mathematics education. *Digital Experiences in Mathematics Education*, 11(1), 9–41. <https://doi.org/10.1007/s40751-025-00172-1>
- Plevris, V., Papazafeiropoulos, G., & Jiménez Rios, A. (2023). Chatbots put to the test in math and logic problems: A comparison and assessment of ChatGPT-3.5, ChatGPT-4, and Google Bard. *AI*, 4(4), 949–969. <https://doi.org/10.3390/ai4040048>
- Richter, S., Kishore, S., Piven, I., Dodd, P., & Bate, G. (2025). Chatbots in tertiary education: Exploring the impact of warm and competent avatars on self-directed learning. *British Journal of Educational Technology*, 56(5), 2102–2124. <https://doi.org/10.1111/bjet.13610>
- Ruiz Recio, M. I., Molina-Ayuso, Á., & Adamuz-Povedano, N. (2026). Integrating computational thinking into secondary school mathematics teaching: A literature review. *Review of Education*, 14(1), e70142. <https://doi.org/10.1002/rev3.70142>
- Sánchez-Ruiz, L. M., Moll-López, S., Nuñez-Pérez, A., Moraño-Fernández, J. A., & Vega-Fleitas, E. (2023). ChatGPT challenges blended learning methodologies in engineering education: A case study in mathematics. *Applied Sciences*, 13(10), 6039. <https://doi.org/10.3390/app13106039>
- Sawyer, A. G., & Aga, Z. G. (2025). A mile high and an inch deep: Exploring ChatGPT as a mathematics curriculum development tool. *School Science and Mathematics*, 126(1), 60–74. <https://doi.org/10.1111/ssm.18348>
- Schoenfeld, A. H. (1985). *Mathematical problem solving*. Academic Press.
- Šedlbauer, J., Činčera, J., Slavík, M., & Hartlová, A. (2024). Students' reflections on their experience with ChatGPT. *Journal of Computer Assisted Learning*, 40(4), 1526–1534. <https://doi.org/10.1111/jcal.12967>

- Son, T. (2024). Intelligent Tutoring Systems in Mathematics Education: A Systematic Literature Review Using the Substitution, Augmentation, Modification, Redefinition Model. *Computers*, 13(10), 270-294. <https://doi.org/10.3390/computers13100270>
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386. <https://doi.org/10.1016/j.chb.2024.108386>
- Sutarto, Dwi Hastuti, I., Fuster-Guillén, D., Palacios Garay, J. P., Hernández, R. M., Namaziandost, E., & Hwang, G. (2022). The effect of problem-based learning on metacognitive ability in the conjecturing process of junior high school students. *Education Research International*, 2022(1), 2313448. <https://doi.org/10.1155/2022/2313448>
- Taylor, A. W. R., Patel, F. Y., Smyth, E. K., Gay, S., & Bird, T. (2026). AI in the Classroom: Observing Preclinical Students' Use of ChatGPT During Case-Based Learning at a UK Medical School. *The Clinical Teacher*, 23(2), e70360. <https://doi.org/10.1111/tct.70360>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., and Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), Article 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Topali, P., Haelermans, C., Molenaar, I., & Segers, E. (2025). Pedagogical considerations in the automation era: A systematic literature review of AIED in K-12 authentic settings. *British Educational Research Journal*, 51(6), 2777–2809. <https://doi.org/10.1002/berj.4200>
- Turgut, S. y Uğurlu, M. (2024). A moderated mediation analysis on the relationship between metacognition and mathematical resilience of gifted students. *Psychology in the Schools*, 61(5), 1848-1867. <https://doi.org/10.1002/pits.23141>
- Ulusoy, F. y Girit-Yildiz, D. (2024). The quality of prospective mathematics teachers' dynamic geometry tasks in terms of the coordination between mathematical depth levels and technological actions. *Journal of Computer Assisted Learning*, 40(5), 2261-2277. <https://doi.org/10.1111/jcal.13029>
- United Nations Educational, Scientific and Cultural Organization (UNESCO). (2023). Guidance for generative AI in education and research. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- Wardat, Y., Tashtoush, M. A., AlAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2286. <https://doi.org/10.29333/ejmste/13272>
- Wu, F., Dang, Y., & Li, M. (2025). A systematic review of responses, attitudes, and utilization behaviors on generative AI for teaching and learning in higher education. *Behavioral Sciences*, 15(4), 467. <https://doi.org/10.3390/bs15040467>
- Yoon, H., Hwang, J., Lee, K., Roh, K. H., & Kwon, O. N. (2024). Students' use of generative artificial intelligence for proving mathematical statements. *ZDM—Mathematics Education*, 56(7), 1531–1551. <https://doi.org/10.1007/s11858-024-01629-0>
- 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, Artículo 28. <https://doi.org/10.1186/s40561-024-00316-7>
- Zhang, J., Zhou, Y., Jing, B., Pi, Z., & Ma, H. (2024). Metacognition and mathematical modeling skills: The mediating roles of computational thinking in high school students. *Journal of Intelligence*, 12(6), 55. <https://doi.org/10.3390/jintelligence12060055>