



Facultad de Ciencias Económicas y Empresariales

ICADE

IMPACTO DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EL DESARROLLO COMPETENCIAL DEL ALUMNADO UNIVERSITARIO

UNA REVISIÓN SISTEMÁTICA DE LA LITERATURA

Autor: Nazaret Martínez Carcelén

Director: Isabel Carrero Bosch

Madrid | Marzo 2026

RESUMEN

El objetivo de esta revisión sistemática de la literatura es analizar el impacto del uso de herramientas de inteligencia artificial generativa por parte del alumnado universitario en el desarrollo de competencias cognitivas, metacognitivas, digitales, sociales y comunicativas, así como en las barreras y limitaciones asociadas a su implementación. La revisión se llevó a cabo siguiendo las directrices PRISMA y el marco de cinco etapas de Hopia et al. (2016). La búsqueda, realizada en *Web of Science* en octubre de 2025, arrojó 742 resultados, de los cuales se seleccionaron finalmente 51 estudios tras el proceso de cribado y evaluación.

Los resultados muestran que el impacto de la inteligencia artificial generativa no es uniforme ni intrínsecamente positivo o negativo, sino que depende del diseño pedagógico de la intervención, del grado de estructuración de las tareas y del nivel de formación previa del alumnado. Por otro lado, los efectos más favorables se concentran en las competencias cognitivas y comunicativas, especialmente en tareas de escritura académica, argumentación y pensamiento crítico, mientras que en las competencias metacognitivas, digitales y sociales los resultados son más heterogéneos. Asimismo, se identifican barreras frecuentes vinculadas al uso irreflexivo de la herramienta, la sobreconfianza, la reducción del esfuerzo cognitivo, las alucinaciones y la falta de guía estructurada por parte de los docentes. En cuanto a las limitaciones de los estudios revisados, predominan las de tipo muestral, contextual y metodológico.

Por último, como hallazgo principal, se concluye que la inteligencia artificial generativa puede favorecer el aprendizaje universitario cuando se integra como apoyo guiado dentro de un diseño pedagógico explícito y con una formación previa, pero no cuando sustituye la implicación cognitiva del estudiante, omitiendo etapas del proceso cognitivo.

PALABRAS CLAVE: Inteligencia artificial generativa; educación superior; competencias cognitivas; competencias metacognitivas; competencias digitales; competencias sociales; competencias comunicativas.

ABSTRACT

The aim of this systematic literature review is to analyse the impact of university students' use of generative artificial intelligence tools on the development of cognitive, metacognitive, digital, social and communicative competencies, as well as the barriers and limitations associated with their implementation. The review was conducted following the PRISMA guidelines and the five-stage framework proposed by Hopia et al. (2016). The search, carried out in Web of Science in October 2025, yielded 742 results, of which 51 studies were ultimately selected after the screening and evaluation process.

The findings show that the impact of generative artificial intelligence is neither uniform nor intrinsically positive or negative but rather depends on the pedagogical design of the intervention, the degree of task structuring, and the level of students' prior training. The most favorable effects are concentrated in cognitive and communicative competencies, particularly in academic writing, argumentation and critical thinking tasks, whereas the results for metacognitive, digital and social competencies are more heterogeneous. In addition, frequent barriers were identified, linked to unreflective use of the tool, overconfidence, reduced cognitive effort, hallucinations, and the lack of structured guidance from teachers. Regarding the limitations of the studies reviewed, sample-related, contextual and methodological limitations were the most prevalent.

Finally, the main finding of this review is that generative artificial intelligence can support university learning when it is integrated as guided assistance within an explicit pedagogical design and accompanied by prior training, but not when it replaces students' cognitive engagement by passing stages of the cognitive process.

KEY WORDS: Generative artificial intelligence; higher education; cognitive competencies; metacognitive competencies; digital competencies; social competencies; communicative competencies.

ÍNDICE DE CONTENIDOS

| | |
|---|----|
| ÍNDICE DE GRÁFICOS..... | 6 |
| ÍNDICE DE TABLAS..... | 6 |
| 1. INTRODUCCIÓN..... | 7 |
| 1.1. Contexto y relevancia del estudio | 7 |
| 1.2. Justificación del estudio | 7 |
| 1.3. Preguntas y objetivo de investigación | 8 |
| 1.4. Metodología | 8 |
| 1.4.1. <i>Identificación del problema</i> | 9 |
| 1.4.2. <i>Búsqueda de literatura</i> | 10 |
| 1.4.3. <i>Evaluación de los datos</i> | 10 |
| 1.4.4. <i>Análisis de los datos</i> | 10 |
| 2. PATRONES DE INTERVENCIÓN, EFECTOS EN EL DESARROLLO COMPETENCIAL, BARRERAS Y LIMITACIONES..... | 11 |
| 2.1. Patrones recurrentes en el diseño de los grupos experimental y control 11 | |
| 2.2. Resultados asociados al análisis de competencias | 16 |
| 2.2.1. <i>Competencias cognitivas</i> | 16 |
| 2.2.2. <i>Competencias metacognitivas</i> | 19 |
| 2.2.3. <i>Competencias digitales</i> | 24 |
| 2.2.4. <i>Competencias sociales</i> | 28 |
| 2.2.5. <i>Competencias comunicativas</i> | 31 |
| 2.3. Análisis de las barreras detectadas en las intervenciones | 35 |
| 2.3.1. <i>Distribución general en la muestra</i> | 36 |
| 2.3.2. <i>Análisis por tipología de barrera</i> | 38 |
| 2.4. Análisis de las limitaciones de los estudios incluidos | 43 |
| 2.4.1. <i>Limitaciones muestrales</i> | 43 |
| 2.4.2. <i>Limitaciones contextuales</i> | 44 |
| 2.4.3. <i>Limitaciones de diseño</i> | 44 |
| 2.4.4. <i>Limitaciones de duración o ausencia de seguimiento longitudinal</i> | 44 |
| 2.4.5. <i>Limitaciones instrumentales o de medición</i> | 45 |
| 2.4.6. <i>Limitaciones tecnológicas</i> | 45 |
| 2.4.7. <i>Limitaciones analíticas</i> | 45 |

| | | |
|--------|--|----|
| 2.4.8. | <i>Análisis global de las limitaciones encontradas</i> | 45 |
| 3. | CONCLUSIÓN | 46 |
| 3.1. | Propósito y objetivos del estudio | 46 |
| 3.2. | Principales hallazgos y contraste con otras revisiones de la literatura ... | 47 |
| 3.3. | Contribución del estudio | 49 |
| 3.4. | Limitaciones | 49 |
| 3.5. | Recomendaciones para investigaciones futuras | 50 |
| 3.6. | Recomendaciones prácticas | 50 |
| 3.6.1. | <i>Recomendaciones para instituciones públicas</i> | 51 |
| 3.6.2. | <i>Recomendaciones para el profesorado universitario</i> | 51 |
| 3.6.3. | <i>Recomendaciones para el alumnado universitario</i> | 52 |
| 4. | DECLARACIÓN DE USO DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL GENERATIVA EN TRABAJOS DE GRADO | 56 |
| 5. | BIBLIOGRAFÍA | 58 |

ÍNDICE DE GRÁFICOS

| | |
|--|----|
| Gráfico 1. Representación del proceso de revisión de la literatura. | 10 |
| Gráfico 2. Distribución del rol del estudiante en las intervenciones analizadas. | 13 |
| Gráfico 3. Distribución de las intervenciones según la existencia de formación previa en el uso de GenIA. | 13 |
| Gráfico 4. Frecuencia de los distintos usos de la GenIA en las intervenciones analizadas..... | 15 |
| Gráfico 5. Distribución de las intervenciones según la presencia de grupo de control. | 16 |
| Gráfico 6. Distribución de las intervenciones según su contexto de implementación. . | 17 |
| Gráfico 7. Frecuencia de los principales tipos de barreras identificadas en los estudios analizados. | 38 |
| Gráfico 8. Distribución de los estudios según la presencia de barreras asociadas al uso de la IA. | 38 |
| Gráfico 9. Distribución de los estudios según la presencia de barreras asociadas al uso de la IA. | 39 |
| Gráfico 10. Frecuencia de los principales tipos de limitaciones reportadas en los estudios incluidos. | 44 |

ÍNDICE DE TABLAS

| | |
|---|----|
| Tabla 1. Síntesis del análisis de las competencias cognitivas. | 20 |
| Tabla 2. Síntesis del análisis de las competencias metacognitivas..... | 24 |
| Tabla 3. Síntesis del análisis de las competencias digitales. | 28 |
| Tabla 4. Síntesis del análisis de las competencias sociales. | 32 |
| Tabla 5. Síntesis del análisis de las competencias comunicativas..... | 36 |
| Tabla 6. Síntesis sobre las tipologías de barreras en el uso de la IA generativa por parte del alumnado universitario. | 43 |

1. INTRODUCCIÓN

1.1. Contexto y relevancia del estudio

El uso de herramientas de inteligencia artificial generativa por parte del alumnado ha ido tomando relevancia en los últimos años, debido al rápido desarrollo y difusión de la misma. Tanto es así, que la Organización para la Cooperación y el Desarrollo Económicos (OCDE), ha elaborado un informe llamado “*Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education*”, en el que analiza la investigación emergente que sugiere que **la IA generativa puede favorecer el aprendizaje cuando su uso está guiado por principios docentes claros**. Como muestra de su amplia difusión, recoge una encuesta internacional realizada en 2024 a 3.000 **estudiantes de educación superior** en 16 países, que concluye que el **86%** de ellos **utilizaba IA** en sus estudios y un 54% de manera diaria o semanal (Rong & Chun, 2024).

Esta amplísima difusión del uso de IA generativa por parte del alumnado suscita **preguntas** sobre la **naturaleza** misma del **aprendizaje universitario**, transformando el panorama académico a un ritmo sin precedentes (Zhang, 2025) y cuestionando los procesos que el alumno debe seguir realizando por sí mismos, cuales pueden ser apoyados por la IA, la relación entre autonomía del estudiante y dependencia de la herramienta (Chen et al., 2025a) y el **efecto** de todo ello sobre la **adquisición de competencias** (Chen et al., 2025a; Du et al., 2025; Hou et al., 2025 y Naamati-Schneider & Alt, 2024).

1.2. Justificación del estudio

Con el auge de la IA generativa y su creciente utilización por parte de los estudiantes universitarios, son muchas las investigaciones que han estudiado su impacto en el aprendizaje y, en concreto, en las competencias cognitivas, metacognitivas, digitales, sociales y comunicativas del alumnado. Sin embargo, la **proliferación** de estos **estudios** ha producido una **dispersión de la evidencia**, presentando diferencias relevantes en sus enfoques, diseños y variables analizadas, así como en los resultados.

En este contexto, resulta necesario elaborar un análisis que reúna, ordene y compare los resultados existentes, con el fin de identificar patrones recurrentes, puntos de convergencia y de divergencia que contribuyan a **entender** de manera más **profunda** el **impacto** del uso de **inteligencia artificial generativa** en el **aprendizaje universitario**.

1.3. Preguntas y objetivo de investigación

A partir de estas consideraciones, esta revisión sistemática se articula en torno a las siguientes preguntas: ¿qué **patrones recurrentes** se observan en el **diseño** de las intervenciones con inteligencia artificial generativa en alumnado universitario? ¿qué efectos se identifican sobre las **competencias** cognitivas, metacognitivas, digitales, sociales y comunicativas? ¿qué **barreras** aparecen en las intervenciones analizadas? y ¿qué **limitaciones** presentan los estudios revisados?

Para dar respuesta a estas preguntas, el **objetivo general** de esta revisión consiste en analizar de manera sistemática la literatura académica reciente sobre el impacto del uso de inteligencia artificial generativa en el ámbito académico universitario.

Asimismo, los **objetivos específicos** son, en primer lugar, describir cómo se configuran los grupos experimental y de control y las tipologías de intervención predominantes en la literatura revisada; en segundo lugar, examinar los efectos del uso de la GenIA sobre las competencias cognitivas, metacognitivas, digitales, sociales y comunicativas; en tercer lugar, identificar las barreras observadas en las intervenciones analizadas y, en cuarto lugar, identificar y clasificar las principales limitaciones reconocidas por los estudios.

1.4. Metodología

Se ha llevado a cabo una revisión sistemática de la literatura estructurada de acuerdo con el marco de cinco etapas de Hopia et al. (2016): identificación del problema, búsqueda de literatura, evaluación de los datos, análisis de los datos y presentación de resultados. Con el fin de garantizar el rigor en el procedimiento de búsqueda, cribado y presentación de resultados, se han seguido las directrices PRISMA (Page et al., 2021).

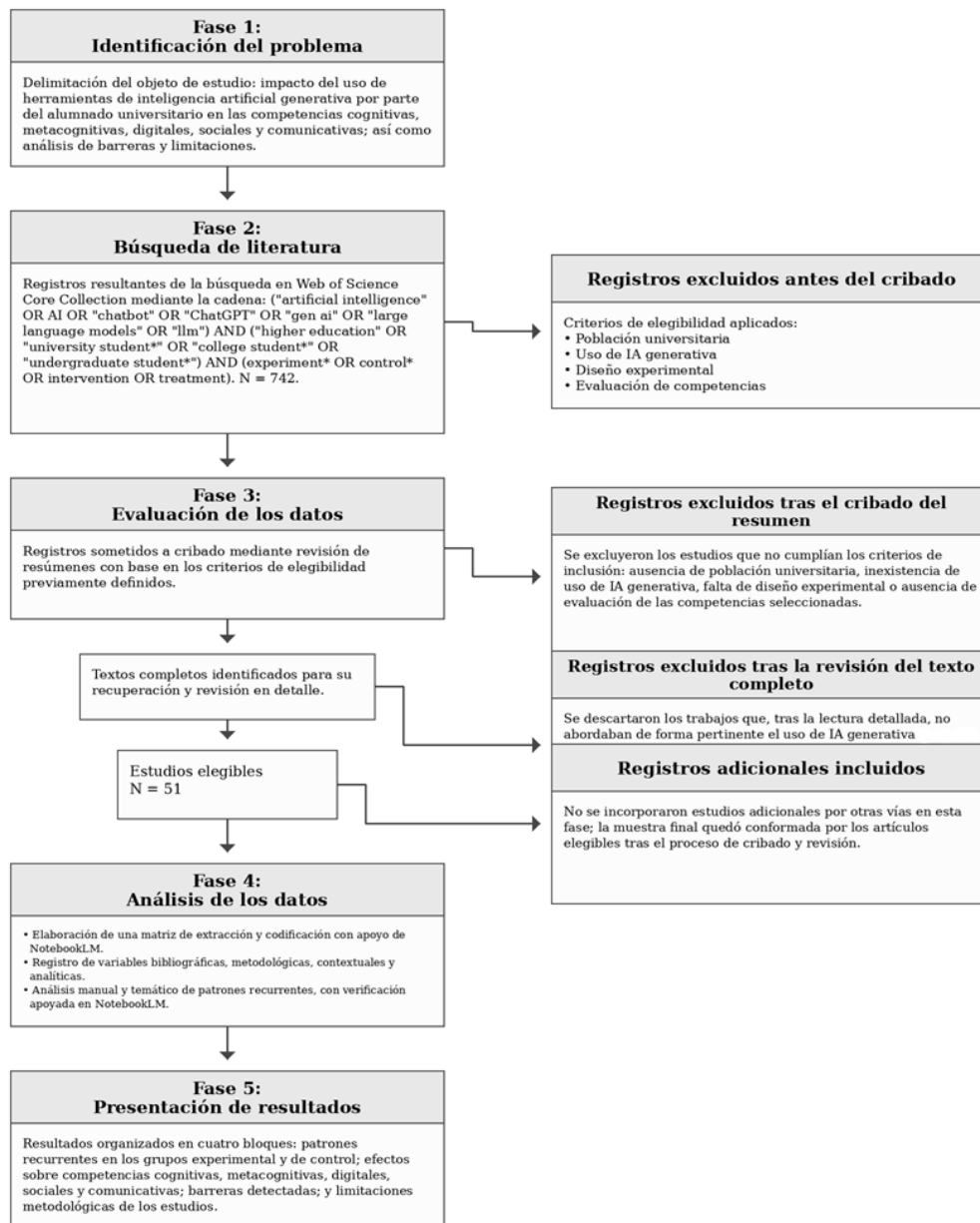


Gráfico 1. Representación del proceso de revisión de la literatura.

1.4.1. Identificación del problema

El propósito de esta revisión es examinar de manera sistemática la literatura académica reciente sobre el uso de herramientas de inteligencia artificial generativa en el contexto universitario; concretamente, en sus efectos sobre las competencias cognitivas, metacognitivas, digitales, sociales y comunicativas del alumnado. Asimismo, se han analizado las barreras detectadas en los estudios y las limitaciones metodológicas recogidas por los mismos.

1.4.2. *Búsqueda de literatura*

Para llevar a cabo la búsqueda sistemática de la literatura se ha seguido el procedimiento PRISMA. En primer lugar, se definieron los **criterios de elegibilidad**. Se delimitó el alcance de la revisión a investigaciones centradas en el estudio del impacto del uso de herramientas de inteligencia artificial generativa por parte de los estudiantes universitarios.

En segundo lugar, se identificaron las **fuentes de información**. La búsqueda se realizó en la *Web of Science (WoS) Core Collection*, reconocida como una de las principales plataformas mundiales para la citación científica y la localización de literatura académica, siendo una fuente de datos completa y autorizada para revisiones sistemáticas (Li et al., 2018). Se aplicaron los siguientes **criterios de inclusión**: artículos que investigaran en población universitaria, el uso de herramientas de inteligencia artificial generativa, que tuvieran un diseño experimental y evaluaran competencias relacionadas con la creatividad, el pensamiento crítico, la colaboración y la comunicación. Se elaboró una **cadena de búsqueda** con palabras clave: ("artificial intelligence" OR AI OR "chatbot" OR "ChatGPT" OR "gen ai" OR "large language models" OR "llm") AND ("higher education" OR "university student*" OR "college student*" OR "undergraduate student*") AND (experiment* OR control* OR intervention OR treatment) La **búsqueda final**, realizada en octubre de 2025, arrojó **742 artículos**.

1.4.3. *Evaluación de los datos*

Una vez obtenidos los resultados iniciales, se llevó a cabo un proceso de cribado progresivo de los estudios identificados. En una primera fase, se revisaron los resúmenes de los artículos seleccionados para valorar que cumplieran con los criterios de elegibilidad previamente definidos. Posteriormente, se realizó una revisión más detallada de los textos resultantes para confirmar su inclusión definitiva en la muestra. La **muestra final** incluyó **51 estudios**.

1.4.4. *Análisis de los datos*

El proceso de análisis de datos se llevó a cabo mediante la elaboración de una **matriz de extracción y codificación** de datos con el apoyo de NotebookLM. Las **variables recogidas** incluyeron el título del estudio, su autoría, resumen, DOI, tipo de estudios, la asignatura, tamaño de la muestra, el contexto y rol del estudiante en la intervención, la

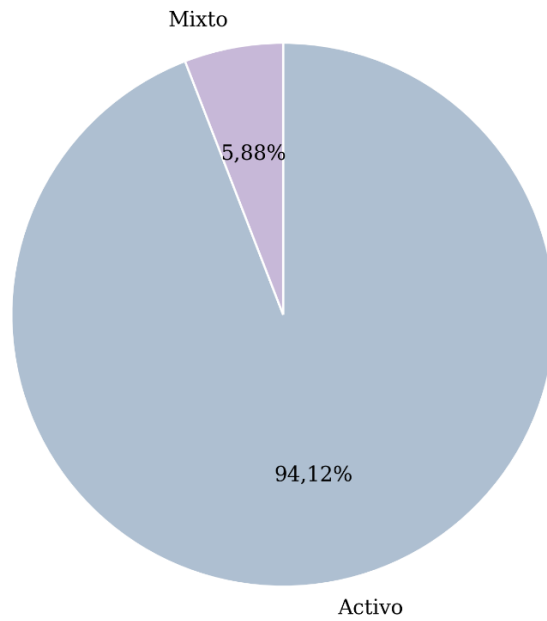
descripción de la intervención en el grupo experimental y de control, la herramienta de IA generativa utilizada, los resultados de competencias cognitivas, metacognitivas, digitales, sociales y comunicativas; las barreras técnicas, docentes y del alumnado; así como las limitaciones detectadas y la aportación principal de cada estudio.

En base a esta matriz, se realizó un **análisis de la información recogida en la matriz**, con el objetivo de localizar patrones recurrentes en el diseño de las intervenciones, la configuración de los grupos experimental y de control, los resultados asociados al desarrollo de competencias (categorizadas en: cognitivas, metacognitivas, digitales, sociales y comunicativas), las barreras detectadas y las limitaciones metodológicas declaradas. Este análisis se realizó verificando con NotebookLM y revisando de manera manual la información de los estudios volcada en la matriz.

2. PATRONES DE INTERVENCIÓN, EFECTOS EN EL DESARROLLO COMPETENCIAL, BARRERAS Y LIMITACIONES

2.1. Patrones recurrentes en el diseño de los grupos experimental y control

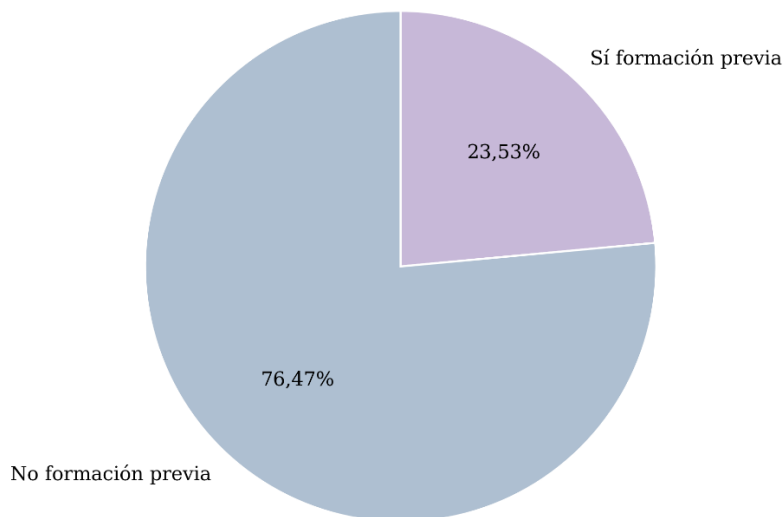
El análisis de las 51 intervenciones revela un **rol predominantemente activo del estudiante**, siendo así en el 94,12% de los casos, permitiendo al alumno tomar la iniciativa y construir el conocimiento con ayuda de la inteligencia artificial generativa (GenAI) pertinente de manera autónoma o recibiendo instrucciones básicas. Por el contrario, solo en el 5,88% de los casos, el rol del estudiante es mixto (Çelik et al., 2024; Le et al., 2024; Mendoza Moheno et al., 2024), combinando elementos de participación activa con elementos de participación pasiva, estando el uso de la IA minuciosamente pautado y dirigido.



Base: N = 51 intervenciones. Elaboración propia a partir de los estudios incluidos en la systematic review.

Gráfico 2. Distribución del rol del estudiante en las intervenciones analizadas.

En segundo lugar, cabe destacar que solo en el 23,53% de los casos los estudiantes recibieron una **formación previa al uso de la GenIA** en el marco de la intervención, lo que sugiere que, en la mayoría de los estudios, la herramienta de inteligencia artificial utilizada fue incorporada sin un entrenamiento previo sobre su funcionamiento, limitaciones o criterios de uso (como la óptima formulación de prompts).



Base: N = 51 intervenciones. Elaboración propia a partir de los estudios incluidos en la systematic review.

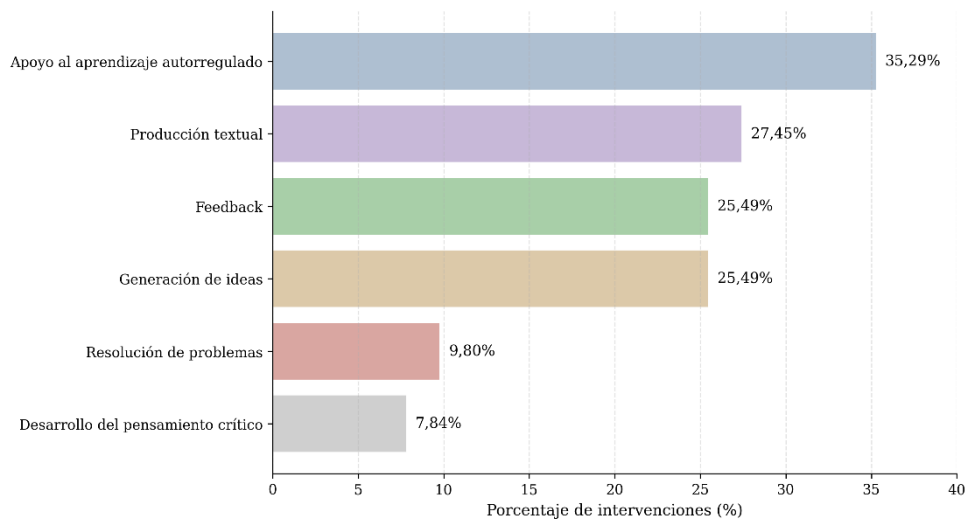
Gráfico 3. Distribución de las intervenciones según la existencia de formación previa en el uso de GenIA.

Por otro lado, el uso concreto de IA Generativa ha sido diferente en función del estudio, pero para su revisión sistematizada, se pueden **catalogar** las **intervenciones** en seis tipos: uso de GenIA como herramienta de apoyo al aprendizaje autorregulado, como herramienta de producción textual, como sistema de retroalimentación o feedback, como sistema de resolución de problemas, como generador de ideas y como herramienta para el desarrollo de pensamiento crítico. La clasificación de las intervenciones en estas seis categorías no impide que en algunas intervenciones el uso de la IA sea múltiple, es decir y a modo ilustrativo, que en una misma intervención se utilice la GenIA como herramienta de apoyo al aprendizaje autorregulado y como herramienta de feedback simultánea o sucesivamente.

En concreto, el **uso más frecuente de GenIA** según estas categorías ha sido como herramienta de apoyo al aprendizaje autorregulado (35,29%). Este dato sugiere que la IA se integra principalmente como herramienta metacognitiva, favoreciendo procesos como la búsqueda de información, planificación, estructuración del contenido y revisión del propio aprendizaje. Este enfoque predominante muestra una especial concentración de estudios enfocados en el uso de la GenIA para el aprendizaje autónomo o autorregulado.

En segundo lugar, destacan tres usos con frecuencias muy similares: para la producción textual (27,45%), como herramienta de feedback (25,49%) y como herramienta generadora de ideas (25,49%). Esta similitud porcentual sugiere un uso de la GenIA pluridimensional, pudiendo ser utilizada funcionalmente para diversas tareas.

Por último, los usos menos frecuentes en los estudios analizados han sido para la resolución de problemas (9,8%) y para el desarrollo del pensamiento crítico (7,84%). Estos porcentajes sugieren que existe un campo de investigación poco explorado, pudiendo ser un objeto interesante para próximos estudios.

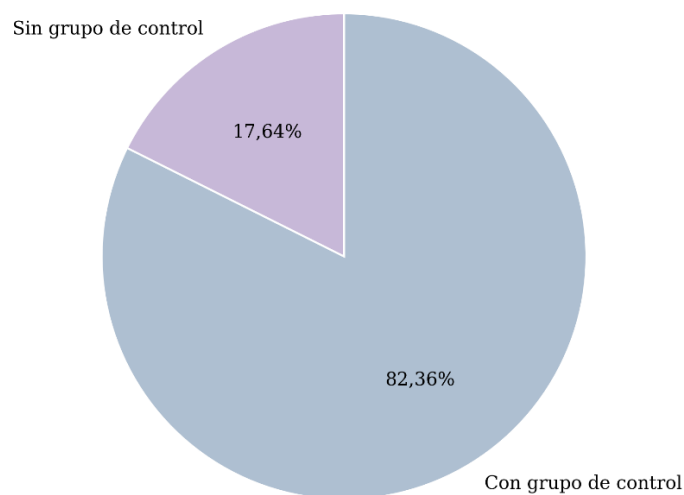


Base: N = 51 intervenciones. Las categorías no son excluyentes, por lo que una misma intervención puede incorporar más de un uso funcional de la GenIA. Elaboración propia.

Gráfico 4. Frecuencia de los distintos usos de la GenIA en las intervenciones analizadas.

Estos resultados permiten identificar tres **patrones estructurales en el conjunto de grupos experimentales**: el predominio del enfoque sobre la teoría del aprendizaje autorregulado, orientado a ser una herramienta de acompañamiento a lo largo del proceso de aprendizaje autónomo; un uso académicamente instrumental de la herramienta, enfocado a tareas de producción y mejora de textos y, por último, una presencia menor de intervenciones enfocadas a habilidades cognitivas de alto nivel como el pensamiento crítico o la resolución de problemas.

En relación con el **diseño metodológico de las intervenciones**, no en todas ellas ha habido presencia de un grupo de control, debido al uso de medidas pretest-postest en el mismo grupo en algunos estudios. En concreto, no hubo grupo de control en el 17,64% de los estudios, realizándose en estos casos pruebas pretest-postest.

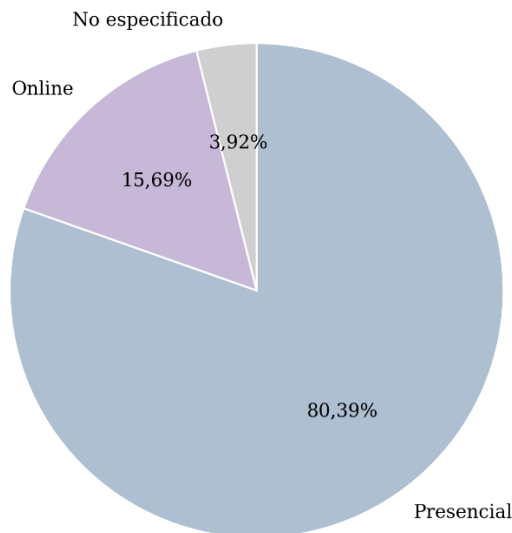


Base: N = 51 intervenciones. En los estudios sin grupo de control se aplicaron diseños pretest-postest en el mismo grupo. Elaboración propia.

Gráfico 5. Distribución de las intervenciones según la presencia de grupo de control.

En cuanto al **tipo de herramienta empleada**, se observan tres categorías, con un predominante uso de ChatGPT como herramienta principal (68,62% de los casos), frente al uso de otras herramientas de GenIA como Gemini o Copilot y al uso combinado de múltiples herramientas de GenIA (Darmawansah et al., 2024; Miao et al., 2025; Oliva-Córdova et al., 2025).

Por último, respecto al **contexto de implementación de la intervención**, hay un claro predominio de intervenciones desarrolladas en entornos presenciales (80,39%), siendo significativamente menor el número de estudios que optaron por la modalidad online (15,69%) y existiendo solo un 3,92% de estudios que no especifican su formato. Este patrón sugiere que a la hora de estudiar el impacto de la GenIA el contexto por el que se está optando es predominantemente presencial, quizá con el fin de ejercer un mayor control sobre las variables intervinientes y reducir inferencias externas no controladas para garantizar la validez interna del estudio.



Base: N = 51 intervenciones. Elaboración propia a partir de los estudios incluidos en la systematic review.

Gráfico 6. Distribución de las intervenciones según su contexto de implementación.

2.2. Resultados asociados al análisis de competencias

2.2.1. Competencias cognitivas

El análisis de los 51 estudios incluidos en la muestra revela que el bloque de **competencias cognitivas** constituye el ámbito en el que se concentra una mayor densidad de resultados y, en términos generales, una orientación mayoritariamente positiva. Concretamente, la mayoría de los artículos que evalúan resultados cognitivos demuestran mejoras en el grupo experimental tras haber integrado herramientas de inteligencia artificial generativa. Este patrón se puede observar especialmente en investigaciones centradas en el rendimiento académico, la comprensión lectora, la producción y calidad textual, la creatividad o el desarrollo del pensamiento crítico.

En particular, en el ámbito de la **escritura académica**, diversos estudios reportan mejoras significativas en la calidad del texto tras la intervención (Hongxia & Razali, 2025 y Tai et al., 2025). En este segundo artículo, si bien es cierto que no se detectan cambios significativos en la disposición a escribir, sí se evidencian mejoras estadísticamente significativas de la calidad argumentativa del texto producido ($p < 0.01$), lo que refuerza la idea de que la herramienta de GenIA actúa como apoyo para estructurar tareas complejas de producción escrita.

Por otro lado, en relación con el **pensamiento crítico**, también se observa un patrón de convergencia consistente. Las investigaciones muestran mejoras significativas en la

aplicación, análisis y evaluación de conceptos y razonamientos, si bien es cierto que se produce en un diseño pre-post, aunque sin grupo de control (Oliva-Córdova et al., 2025). Otro estudio que metodológicamente sí comparan grupo experimental y grupo de control (Chen et al., 2025b) evidencia diferencias significativas a favor del grupo que integró mapas argumentales y ChatGPT, tanto en la puntuación global de pensamiento crítico como en comportamientos de orden superior a través de la identificación mediante análisis de redes epistémicas. Resulta muy interesante resaltar que, en este caso concreto, se puntualiza que en el pretest del grupo de control y del grupo experimental, no hubo una diferencia estadísticamente significativa en las habilidades de pensamiento crítico ($p > 0,05$), por lo que este resultado sugiere que los estudiantes de ambos grupos poseían habilidades de pensamiento crítico similares al inicio del curso. Sin embargo, en el postest, se pudo observar una mejora estadísticamente significativa en las habilidades de pensamiento crítico en los estudiantes del grupo experimental ($p < 0,001$) lo que indica que el uso de ChatGPT y mapas argumentativos mejoró sus habilidades de pensamiento crítico. Del mismo modo, otras investigaciones (Du et al., 2025 y Zhang et al., 2025b) reportan incrementos significativos en la frecuencia de uso de habilidades de pensamiento de orden superior (HOTS) tras la intervención con GenIA, poniendo de manifiesto que la IA puede estimular un razonamiento más profundo al ofrecer múltiples opciones de retroalimentación en tiempo real, lo que favorece que los estudiantes realicen análisis más exhaustivos y revisiones continuos, favoreciendo el desarrollo de estas habilidades (Du et al., 2025). Zhang et al., (2025b) destacan además el papel clave del pensamiento crítico como variable puente para que, con el uso de IA generativa, el impacto en el pensamiento creativo se traduzca en una mejora real sin generar dependencia cognitiva.

Por otro lado, existen también divergencias relevantes a la hora de analizar los resultados de competencias cognitivas, sobre todo en el ámbito del **pensamiento creativo**. Por un lado, si bien se puede observar una mejora en el desempeño creativo objetivo, se revelan dificultades a la hora de calibrar la correlación entre juicio y desempeño, no existiendo correlación entre los juicios de autoevaluación y desempeño real (calidad $r = 0.02$ y originalidad $r = -0.03$), lo que indica que los estudiantes tuvieron dificultades serias a la hora de calibrar sus juicios cuando usaban la herramienta, produciéndose además un sesgo de utilidad, ya que cuanto más útil consideraban ChatGPT, más tendían a sobreestimar la calidad y originalidad de sus resultados, a pesar de que esa percepción de utilidad no correlacionaba con un mejor desempeño real (Urban et al., 2024). Además, de manera

más concreta, Selem et al. (2025), identifican patrones de uso asociados al deterioro del pensamiento creativo a largo plazo, teniendo un efecto amplificador de la productividad a corto plazo, pero conduciendo a una pereza metacognitiva que se refleja en un desempeño deficiente cuando el estudiante debe enfrentar evaluaciones sin asistencia tecnológica, debido a un uso basado en patrones de búsqueda de respuestas directas y omisión de etapas cognitivas.

Asimismo, algunos estudios no encuentran diferencias significativas entre los grupos experimentales y de control. Por ejemplo, del análisis de la capacidad de los estudiantes para sintetizar conocimientos y del grado de elaboración de los argumentos en Zhong et al. (2024), no se demuestran diferencias significativas entre grupos, a pesar de que otras dimensiones sí muestran cambios, lo que subraya que la mejora cognitiva no es uniforme, sino que depende del diseño de la tarea, del contexto y de la forma de integrar la herramienta de IA en el experimento. Del mismo modo, Guo et al., (2023) no observaron mejoras significativas en la complejidad estructural global de los argumentos, aunque sí en la complejidad estructural de los subapartados número de afirmaciones y presencia de datos y garantías de los argumentos.

Desde el punto de vista metodológico, se puede observar un patrón claro: los efectos positivos en competencias cognitivas surgen con mayor consistencia cuando en las intervenciones se integra la IA como apoyo guiado dentro de actividades previamente diseñadas y organizadas por el docente, estableciendo criterios de evaluación explícitos y pautas objetivas de desempeño. Por el contrario, el uso de la GenIA de manera menos estructurada o para tareas más abiertas como las relacionadas con la creatividad, los efectos son más heterogéneos y existe riesgo de dependencia.

| Competencia cognitiva | Resultado principal observado | Dirección del efecto | Matices o limitaciones relevantes | Estudios ejemplificativos | Comentario analítico breve |
|--|--|-------------------------|--|---|--|
| Escritura académica y calidad textual | Mejora en la calidad argumentativa y estructural del texto producido. | Positivo | La herramienta sirve de apoyo para tareas complejas , aunque no se detectan cambios en la disposición a escribir. | Hongxia & Razali (2025); Tai et al. (2025). | El efecto positivo depende de la forma de integración de la IA como andamiaje para la producción escrita. |
| Pensamiento crítico | Incremento en las capacidades de aplicación, análisis y evaluación de conceptos. | Positivo | El éxito depende de un alto nivel de estructuración pedagógica y del uso de apoyos adicionales como mapas conceptuales para estimular un razonamiento profundo. | Oiva-Córdova et al. (2025); Chen et al. (2025b); Du et al. (2025); Zhang et al. (2025). | El pensamiento crítico actúa como variable puente para evitar que la asistencia derive en dependencia cognitiva . |
| Pensamiento creativo | Mejora del desempeño creativo, pero con dificultades de calibración | Mixto | Fuerte dependencia del diseño de la tarea (tareas abiertas); problemas de calibración entre juicio y desempeño real, y sesgo de utilidad. | Urban et al. (2024); Selem et al. (2025). | El uso no estructurado genera pereza metacognitiva y riesgo de deterioro creativo a largo plazo por omisión de etapas cognitivas. |
| Síntesis y elaboración argumentativa | Ausencia de variaciones entre grupo experimental y de control. | No significativo | El resultado depende de la forma de integrar la herramienta y del diseño específico de la actividad propuesta. | Zhong et al. (2024); Guo et al. (2023). | La mejora cognitiva no es automática ni uniforme en todas las dimensiones del aprendizaje. |

Tabla 1. Síntesis del análisis de las competencias cognitivas.

2.2.2. Competencias metacognitivas

El análisis de las **competencias metacognitivas** en los 51 estudios incluidos en la muestra presenta unos resultados mucho más heterogéneos que los observados en el bloque cognitivo. Esta heterogeneidad se explica principalmente porque no todos los estudios

entienden el concepto de “metacognición” de la misma manera (algunos la miden como autorregulación con escalas específicas, otros la incluyen dentro de habilidades más amplias como HOTS y otros la tratan como reflexión cualitativa) y porque los estudios no usan el mismo tipo de diseño (algunos son experimentales con grupo de control, pre-post sin control, y otros utilizan la metacognición como variable moderadora y no como resultado).

En primer lugar, existen **investigaciones que miden procesos metacognitivos de forma explícita y comparativa**. La investigación de Chen et al. (2025a) constituye uno de los ejemplos metodológicamente más sólidos. El estudio emplea un experimento aleatorizado con dos condiciones (ChatGPT vs. experto humano) y analiza el proceso de búsqueda de ayuda mediante datos objetivos (*trace data*, *eye-tracking* y análisis conversacional). Los resultados muestran diferencias estadísticamente significativas en varias etapas del proceso ($p < 0.001$), lo que evidencia que el grupo que utilizó ChatGPT dedicó una proporción significativamente mayor de su proceso a la fase de solicitud directa de ayuda y mostró una presencia menor de etapas de diagnóstico y evaluación crítica de la respuesta. Esto no solo confirma diferencias entre condiciones, sino que parece sugerir que cuando la ayuda proviene de una herramienta de IA, se produce una reconfiguración del proceso de autorregulación del aprendizaje.

En la misma línea, aunque dentro de un marco teórico más amplio, Wei et al. (2025) incluyen **dimensiones directamente vinculadas a procesos metacognitivos**, tales como la regulación de la tarea, la regulación social y la reflexión. El diseño experimental aleatorizado muestra diferencias significativas a favor del grupo que utilizó herramientas de IA generativa tanto en la puntuación global de resolución colaborativa ($F = 10.69$, $p < 0.01$) como en subdimensiones de regulación social ($F = 7.05$, $p < 0.01$) y regulación de la tarea ($F = 4.67$, $p < 0.05$). En este caso, la metacognición no aparece como un concepto aislado a analizar, pero sí medida de forma comparativa, lo que permite establecer un efecto positivo en los procesos reguladores del aprendizaje.

Por otro lado, otros estudios muestran **efectos** más complejos de interpretar, incluso **ambiguos**. En esta línea, Urban et al. (2024) reporta incrementos significativos en autoeficacia y reducción de dificultad percibida, pero no encuentra correlación entre los juicios de autoevaluación y el desempeño real (calidad $r = 0.02$; originalidad $r = -0.03$). Este resultado indica que, aunque el desempeño creativo mejora, el monitoreo

metacognitivo no se ajusta con precisión al rendimiento, produciéndose un sesgo de utilidad que conduce a que los estudiantes sobreestimen la calidad del resultado producido. Por tanto, en este caso, no puede afirmarse una mejora de las competencias metacognitivas con el uso de la herramienta de IA generativa, sino una distorsión en el proceso de autoevaluación. Estos resultados pueden servir de **predictor directo de la integración de errores**, de modo que aquellos estudiantes con menor sesgo metacognitivo integran más información correcta, mientras que aquellos con un exceso de confianza en la herramienta de IA generativa son más propensos a pasar por alto y aceptar alucinaciones de la IA.

En otras intervenciones específicas en las que **el aprendizaje autorregulado actúa como variable dependiente** (Lee et al., 2025), y en las que se utiliza un diseño experimental con ANCOVA, se demuestra que la guía interactiva basada en chatbot mejora significativamente el aprendizaje autorregulado (($F = 4.21$, $p < 0.05$), específicamente en lo que respecta a la previsión, la ejecución y la autorreflexión. Este resultado introduce un matiz relevante: parece que la mejora metacognitiva no se deriva automáticamente del uso de IA, sino del modo en que se integra pedagógicamente con mecanismos de acompañamiento y control del proceso.

Por otra parte, existen estudios en los que la **metacognición** no actúa como variable dependiente principal, **sino como un componente que estructura o modera otras variables** dentro de estudios más amplios. Esto ocurre en Zhang et al. (2025b), en el que el pensamiento metacognitivo funciona como **variable moderadora** dentro de un modelo de mediación. En concreto, en este estudio el término de interacción resulta estadísticamente significativo ($p < 0.001$), indicando que niveles más elevados de pensamiento metacognitivo potencian el efecto del pensamiento creativo sobre el pensamiento crítico. No obstante, dado que en este caso la metacognición no se mide como un resultado independiente del uso de la IA, no es posible afirmar que exista una mejora directa atribuible a la intervención, sino solamente que tiene un **papel amplificador dentro del proceso cognitivo**.

En un segundo bloque de investigaciones, la autorregulación y la reflexión son abordadas desde un **enfoque cualitativo**, sin realizar una medición o comparación experimental entre grupos. Expresión de esta circunstancia son los estudios de Al-Obaydi et al. (2025) y Sedlbauer et al. (2024). El primero desarrolla un diseño sustentado en entrevistas, donde

el aprendizaje autorregulado es parte del marco conceptual, pero no se mide como variable, mientras que en el segundo se realiza un análisis cualitativo de reflexiones escritas por un único grupo de estudiantes, sin grupo de control ni análisis estadístico de inferencias. Asimismo, aunque la investigación de Leahy et al. (2025) sí recoge datos cuantitativos descriptivos sobre percepciones de pensamiento crítico y compromiso, no mide específicamente la metacognición. En casos como estos (Al-Obaydi et al., 2025; Leahy et al., 2025 y Sedlbauer et al., 2025), la evidencia aportada es relevante para comprender la experiencia subjetiva del estudiante, pero no permite establecer conclusiones sobre la mejora de la competencia metacognitiva.

Finalmente, existen estudios con un **diseño comparativo** en los que, a pesar de formarse un grupo experimental y un grupo de control, la **metacognición no constituye el resultado principal**. En el estudio realizado por Mendoza Moheno et al. (2024), se observan mejoras significativas en variables como la confianza en el proceso de aprendizaje ($p = 0.030$) y el trabajo colaborativo ($p = 0.004$), pero no se utilizan escalas psicométricas específicas de autorregulación. Del mismo modo, Makransky et al. (2025), elaboraron un diseño experimental entre grupos y reportan diferencias significativas en conocimiento conceptual y actitudes, pero no miden la metacognición como un concepto independiente.

En conclusión, el bloque de competencias metacognitivas evidencia que **el impacto de la inteligencia artificial generativa depende en gran medida del modo en que se define y mide la variable**. Cuando los procesos metacognitivos se operacionalizan explícitamente y se comparan entre condiciones, pueden observarse diferencias significativas en la regulación del aprendizaje. Sin embargo, en un número considerable de estudios, la metacognición aparece como marco teórico o como percepción subjetiva, sin medición inferencial. Además, algunos resultados revelan efectos ambivalentes (Chen et al., 2025a y Urban et al., 2024), donde la reducción de la carga percibida (descarga metacognitiva) o el aumento de la confianza no se traducen necesariamente en una mejora del proceso o la percepción del propio desempeño, produciéndose incluso una pereza metacognitiva en el proceso autorregulatorio (Kim et al., 2025) fomentando la consecución de ideas rápidas frente a una planificación estratégica y un monitoreo del aprendizaje profundo. Esta heterogeneidad metodológica impide formular una conclusión unívoca y refuerza la necesidad de distinguir entre la mejora del rendimiento, la

percepción del estudiante y el desarrollo efectivo de procesos de aprendizaje autorregulado.

| Forma de abordaje de la metacognición | Forma de medición | Resultado principal observado | Dirección/tipo de evidencia | Estudios ejemplificativos |
|---|---|---|---|--|
| Medición explícita y comparativa | A través de datos objetivos como <i>trace data</i> , <i>eye-tracking</i> y análisis conversacional. | Reconfiguración del proceso autorregulatorio: aumento de la solicitud directa de ayuda y reducción de fases de diagnóstico y evaluación crítica. | Efecto ambivalente | Chen et al. (2025a) |
| Dimensiones de regulación y colaboración vinculadas a la metacognición | La metacognición se analiza como parte de la competencia colaborativa | Efecto positivo significativo en los procesos reguladores del aprendizaje y puntuaciones globales. | Mejora directa | Wei et al. (2025) |
| Monitoreo y precisión del juicio | Relación entre juicios de autoevaluación y desempeño real. | Distorsión en la autoevaluación y sesgo de utilidad; los estudiantes sobreestiman la calidad de su trabajo. | Efecto ambivalente | Urban et al. (2024) |
| Variable dependiente en intervención | Medición de fases de previsión, ejecución y autorreflexión mediante ANCOVA. | Mejora significativa del aprendizaje autorregulado cuando hay mecanismos de acompañamiento y control del proceso. | Mejora condicionada por el diseño pedagógico | Lee et al. (2025) |
| Variable moderadora | Como componente que estructura o potencia la relación entre otras variables. | Niveles elevados de pensamiento metacognitivo potencian el efecto del pensamiento creativo sobre el crítico. | Función moderadora | Zhang et al. (2025) |
| Enfoque cualitativo | A través de entrevistas, reflexiones escritas y percepciones de pensamiento crítico. | Comprensión de la experiencia subjetiva del estudiante durante el aprendizaje. | Evidencia cualitativa no concluyente | Al-Obaydi et al. (2025), Sedlbauer et al. (2024) y Leahy et al. (2025) |
| Resultado no principal | Medición de confianza, trabajo colaborativo o conocimiento conceptual. | Mejoras en variables actitudinales y de conocimiento , pero sin datos de autorregulación. | Sin medición específica de metacognición | Mendoza Moheno et al. (2024) y Makransky et al. (2025) |

2.2.3. Competencias digitales

El análisis de las **competencias digitales** dentro de los 51 estudios incluidos en la muestra revela que se trata de una de las competencias con menor presencia explícita como variable dependiente formalmente medida. A diferencia de las competencias cognitivas y metacognitivas, donde existe una mayor densidad de estudios con instrumentos específicos y análisis inferenciales, en el ámbito digital solo un número reducido de investigaciones evalúa de manera directa la alfabetización digital, la competencia digital o la alfabetización en IA como resultado cuantitativo.

En primer lugar, un estudio que mide la competencia digital es el elaborado por Naamati-Schneider & Alt (2024), en el que se emplea la *Digital Literacy Scale* (DLs), estructurada en tres dimensiones: acceso, análisis y evaluación de la información. El diseño cuasi-experimental compara tres condiciones (*Problem-based learning* con apoyo de ChatGPT (AIPBL), PBL sin IA y *lectura-based learning* como grupo de control tradicional). Los resultados muestran que el grupo AIPBL mejora significativamente únicamente en la dimensión de evaluación ($F = 6.46, p < 0.05$), mientras que el grupo PBL sin IA mejora en las tres dimensiones (acceso, análisis y evaluación). Este resultado es especialmente relevante, pues indica que la integración de la GenIA no es garantía de una mejora global de la alfabetización digital, sino que esta puede concentrarse en determinadas subcompetencias, en este caso, la capacidad de evaluación crítica.

En segundo lugar, la **alfabetización en IA** aparece como variable explícita en el estudio de Hou et al. (2025), en el que se utiliza la *Artificial Intelligence Literacy Scale* (AILS) para medir las variables comprensión y competencia en cuatro subescalas (conciencia, uso, evaluación y ética). Sin embargo, el diseño es transversal y no experimental, por lo que la alfabetización en IA no se analiza como resultado de una intervención, sino como variable predictora dentro de un modelo explicativo. Los resultados muestran que la alfabetización en IA predice significativamente habilidades de pensamiento crítico ($\beta = 0.66, p < 0.001$) y disposición al pensamiento crítico ($\beta = 0.41, p < 0.001$), así como un efecto indirecto significativo sobre el uso reflexivo a través del pensamiento crítico ($\beta =$

0.20, SE = 0.03, $p < 0.001$). En este caso, la competencia digital actúa como un **factor asociado al tipo de uso de la herramienta**, pero no como una mejora tras la intervención.

El análisis de estos resultados confirma los hallazgos ya observados en otros estudios enfocados en competencias cognitivas y metacognitivas. En concreto, los estudios de Al-Abri (2025); Darmawansah et al. (2025); Wu et al. (2025); Zhang (2025) y Zhang et al. (2025b) incorporan formación previa en el uso de IA, obteniendo mejoras significativas al utilizar la herramienta de GenIA. De este modo, los resultados ponen de relieve la importancia de proporcionar una formación inicial adecuada antes del uso de las herramientas de IA, ya que parece favorecer un uso más eficaz y permite que el alumnado aproveche mejor su potencial.

Por otro lado, el estudio de Mendoza Moheno et al. (2024) incorpora la medición de competencias digitales tomando como referencia el marco DigComp 2.2 dentro de un diseño cuasi-experimental con grupo experimental y grupo de control. En este caso se observan mejoras significativas en el grupo experimental en el nivel de conocimiento de ChatGPT ($t = -2.013$, $p = 0.050$) y en la frecuencia de uso ($t = -2.944$, $p = 0.005$), así como diferencias significativas en variables como la confianza en el proceso de aprendizaje y el trabajo colaborativo. Si bien es cierto que estas variables no se centran en habilidades de desempeño digital estrictamente, sino que se sitúan en el plano de conocimiento, el diseño comparativo del experimento permite afirmar que la integración estructurada de la herramienta puede generar cambios significativos en dimensiones asociadas a la competencia digital.

Asimismo, Beninger et al. (2025) evalúan la alfabetización en IA generativa a través de preguntas alineadas con categorías amplias de comprensión, uso y evaluación. Sin embargo, los propios autores reconocen que las preguntas se mantienen en un nivel general y no profundizan en aspectos específicos de dicha competencia. Esto implica que, aunque el estudio aporte evidencia sobre cambios tras la intervención, la medición no descompone la competencia digital en subcompetencias concretas, lo que limita el grado de detalle con el que puede analizarse el impacto de esta intervención y su comparación con escalas validadas más estructuradas.

En conjunto, el bloque de competencias digitales presenta dos **patrones principales**. En primer lugar, cuando se utilizan escalas estructuradas de alfabetización digital (como DLs

o AILS), los resultados muestran mejoras selectivas y no uniformes por dimensiones, concentrándose especialmente en la evaluación crítica más que en el acceso o análisis (Naamati-Schneider & Alt, 2024). En segundo lugar, cuando la competencia digital se conceptualiza como alfabetización en IA, la evidencia proviene mayoritariamente de modelos predictivos o de instrumentos de alcance general, lo que limita la posibilidad de establecer relaciones causales robustas derivadas de la intervención.

Desde el punto de vista metodológico, se observa que la evidencia experimental directa sobre mejora de competencias digitales es más limitada que en otros bloques competenciales. Esto no implica una ausencia de impacto, sino una menor densidad de estudios que operativicen la competencia digital como resultado principal con instrumentos específicos y comparación entre condiciones. En consecuencia, cualquier conclusión sobre el impacto de la GenIA en competencias digitales debe formularse con cautela, distinguiendo entre el incremento de conocimiento o confianza en el uso de la herramienta y el desarrollo efectivo de habilidades digitales estructuradas.

| Forma de abordaje de la competencia digital | Forma de medición | Resultado principal observado | Dirección/tipo de evidencia | Estudios ejemplificativos | Matices o limitaciones relevantes |
|--|--|--|---|----------------------------------|---|
| Evaluación mediante escalas de alfabetización digital estructuradas | Uso de instrumentos como la <i>Digital Literacy Scale</i> (DLs) centrada en acceso, análisis y evaluación. | Mejora significativa limitada a la dimensión de evaluación crítica | Mejora parcial y selectiva. | Naamati-Schneider et al. (2024). | La integración de GenIA no garantiza una mejora global; diseños sin IA mostraron progresos en más dimensiones (acceso y análisis). |
| Alfabetización en IA como variable predictora | Aplicación de la <i>Artificial Intelligence Literacy Scale</i> (AILS) en dimensiones de conciencia, uso, evaluación y ética. | La alfabetización en IA predice de forma significativa habilidades y disposición al pensamiento crítico . | Asociación predictiva. | Hou et al. (2025). | Diseño transversal no experimental; la competencia digital actúa como factor asociado al tipo de uso, pero no como mejora tras la intervención. |
| Dimensiones asociadas en diseños comparativos | Medición de conocimiento y frecuencia de uso bajo marcos de referencia como DigComp 2.2. | Incremento significativo en el nivel de conocimiento de la herramienta y en la frecuencia de uso . | Mejora en dimensiones asociadas. | Mendoza Moheno et al. (2024). | Las variables evaluadas se sitúan en el plano de conocimiento y no en habilidades de desempeño digital estrictas. |
| Medición general de alfabetización en IA | Uso de instrumentos con categorías amplias de comprensión, uso y evaluación sin desagregación técnica. | Identificación de cambios positivos tras la intervención en niveles generales de alfabetización. | Evidencia limitada por medición general. | Beninger et al. (2025). | Los instrumentos no descomponen la competencia en subcompetencias concretas, lo que dificulta la comparación con escalas validadas. |

Tabla 3. Síntesis del análisis de las competencias digitales.

2.2.4. Competencias sociales

En comparación con los bloques anteriores, el número de estudios que miden explícitamente competencias sociales como variable dependiente es reducido. De los 51 artículos incluidos en la muestra, únicamente siete (Darmawansah et al., 2025; Hou et al., 2025; Mendoza Moheno et al., 2024; Saritepeci & Yildiz Durak., 2024; Tai et al., 2025; Wei et al., 2025 y Zhong et al., 2024) miden explícitamente variables asociadas a la colaboración, el trabajo en equipo o el desempeño grupal como resultados cuantificados. Esta delimitación es relevante, ya que un número considerable de intervenciones incorpora dinámicas colaborativas en el diseño de la tarea sin medir la competencia social como variable dependiente (Al-Obaydi et al., 2025; Chen et al., 2025b; De la Puente et al., 2024; Guo et al., 2023; Naamati-Schneider & Alt, 2024 y Wu et al., 2025) lo que obliga a restringir el análisis exclusivamente a los estudios que la operacionalizan de forma explícita.

En cuanto a las convergencias encontradas, cuatro de los estudios reportan **efectos positivos significativos asociados al uso de GenIA en variables sociales** (Darmawansah et al., 2025; Hou et al., 2025; Mendoza Moheno et al., 2024 y Wei et al., 2025). Metodológicamente, el estudio más robusto es el elaborado por Wei et al. (2025), que emplea un diseño experimental aleatorizado con grupo de control y combina autoinforme (*Collaborative Problem-Solving Skills Questionnaire*) con evaluación externa del desempeño creativo del equipo. Los resultados muestran diferencias significativas a favor del grupo con IA tanto en la puntuación general de CPS ($F = 10.69$, $p < 0.01$, $\eta^2 = 0.16$) como en la subdimensión de regulación social ($F = 7.05$, $p < 0.01$), así como en el desempeño creativo evaluado por expertos ($F = 64.51$, $p < 0.001$). En este caso, la mejora percibida por los alumnos se acompaña de una mejora objetiva en el rendimiento grupal.

En esta misma línea, la intervención de Darmawansah et al. (2025), con un diseño cuasi-experimental y grupo de control, muestra un incremento significativo en la **tendencia a la colaboración en el grupo experimental** ($F = 8.076$, $p < 0.01$, $\eta^2 = 0.33$). Asimismo, el estudio de Mendoza Moheno et al. (2024) identifica un aumento significativo en la

confianza para fomentar el trabajo colaborativo ($t = 3.031$, $p = 0.004$), si bien en este caso la variable medida es perceptiva y no mide el desempeño colaborativo de manera objetiva.

Por su parte, Hou et al. (2025) introducen una aproximación distinta, al conceptualizar el **“uso colaborativo” como parte de los comportamientos de dependencia hacia la IA**. Mediante un modelo de *path analysis*, el estudio muestra que las habilidades de pensamiento crítico predicen positivamente el uso colaborativo ($\beta = 0.25$, $SE = 0.03$, $p < 0.001$). Aunque no se trata de un diseño experimental con grupo de control, sí aporta evidencia sobre la asociación entre variables cognitivas individuales y la forma en que los estudiantes integran la IA en dinámicas de trabajo compartido. De este modo, analiza también el efecto positivo indirecto que tiene la alfabetización en IA (competencia digital) en el uso colaborativo.

No obstante, la evidencia respecto de las variables sociales no es unívoca. Tres estudios (Saritepeci & Yildiz Durak, 2024; Tai et al., 2025 y Zhong et al., 2024) **no encuentran efectos significativos en las variables sociales analizadas**. En el estudio de Zhong et al. (2024), pese a contar con tres condiciones experimentales, no se detectan diferencias significativas en la colaboración percibida, ni tampoco en la integración interdisciplinar demostrada en los productos académicos ($\chi^2 = 0.65$, $p = 0.72$). De forma similar, la investigación de Tai et al. (2025), con diseño pre-post de un solo grupo, no encuentra cambios significativos en el factor de involucramiento relacionado con el papel de los compañeros en la comunicación escrita ($t = -0.41$, $p = 0.68$). Finalmente, Saritepeci & Yildiz Durak (2024), concluyen que la subdimensión *collaborative working with diversity* no muestra un incremento significativo atribuible al uso de IA, observándose incluso mayores incrementos brutos en el grupo de control.

Desde el punto de vista metodológico, se observan **tres patrones relevantes**. En primer lugar, existe un predominio claro de escalas de autoinforme tipo Likert para medir la competencia social, incluso en estudios con diseño experimental. Solo en el caso de Wei et al. (2025) se combina la percepción con una evaluación externa sistemática del desempeño del equipo, lo que refuerza la solidez de los hallazgos positivos. En segundo lugar, la mayoría de los estudios con efectos significativos positivos cuentan con grupo de control, lo que sugiere que, cuando se detectan mejoras, tienden a ser respaldadas por diseños comparativos. En tercer lugar, la dirección del efecto es mixta: aproximadamente

la mitad de los estudios identifican mejoras significativas, mientras que el resto no encuentra diferencias atribuibles al uso de la herramienta.

En conjunto, la evidencia sugiere que la GenIA puede potenciar determinadas dimensiones de la colaboración cuando la intervención está estructurada y la herramienta se integra como apoyo regulativo en tareas cooperativas. Sin embargo, no existe una evidencia consistente de que su mera introducción en entornos grupales garantice mejoras en las competencias sociales. La dirección del efecto parece depender, de nuevo, del diseño e instrucciones del experimento y del modo en que se operacionaliza y evalúa la competencia. La mejora, por tanto, no parece derivarse del uso de la herramienta en sí, sino de su inserción en marcos pedagógicos que obligan a la coordinación, la regulación compartida y la corresponsabilidad en la tarea. En ausencia de dichas condiciones, la IA actúa como recurso instrumental individual dentro de un contexto grupal, sin traducirse necesariamente en un desarrollo medible de la competencia social.

| Tipo de evidencia sobre competencia social | Resultado principal observado | Dirección/tipo de evidencia | Estudios ejemplificativos | Matices o limitaciones relevantes |
|---|---|--|---|---|
| Efectos positivos significativos en variables sociales | Incremento en regulación social, desempeño creativo grupal y confianza para el trabajo colaborativo. Pensamiento crítico predice positivamente la integración de IA en el uso colaborativo | Positiva significativa directa. Positiva indirecta. | Darmawansah et al. (2025); Mendoza Moheno et al. (2024); Wei et al. (2025). Hou et al. (2025). | Predominio de diseños experimentales o cuasi-experimentales con grupo de control y autoinforme subjetivo de la evaluación. Solo el estudio de Wei et al. combina la percepción del alumno con una evaluación externa sistemática. Hoş et al. No emplean un diseño experimental comparativo, sino un análisis de relaciones entre variables – conceptualiza el uso colaborativo como un comportamiento de dependencia hacia la IA. |
| Ausencia de efectos significativos en variables sociales | Inexistencia de diferencias estadísticas entre grupos o tras la intervención; incluso mejores resultados brutos en grupos de control. | No significativa | Saritepeci et al. (2024); Tai et al. (2025); Zhong et al. (2024). | Estudios que analizan dimensiones como <i>collaborative working with diversity</i> o involucramiento comunicativo. |

Tabla 4. Síntesis del análisis de las competencias sociales.

2.2.5. Competencias comunicativas

El bloque de competencias comunicativas constituye, tras las cognitivas, una de las competencias con mayor densidad de evidencia empírica dentro de los 51 estudios analizados. A diferencia de lo observado en las competencias digitales y sociales, donde la medición explícita como resultado es más limitada, en este caso se identifican catorce investigaciones que evalúan el desempeño comunicativo mediante rúbricas externas, pruebas estandarizadas o métricas automatizadas, más allá del autoinforme perceptivo.

En términos generales, la mayor parte de estos catorce estudios son **favorables a la integración de herramientas de Inteligencia Artificial Generativa**, si bien con matices relevantes según el diseño y el modo de uso.

En el ámbito de la **expresión escrita**, la convergencia es especialmente sólida. Diversos estudios experimentales y cuasi-experimentales con grupo de control reportan **mejoras significativas** en la calidad textual tras la intervención con IA (Darmawansah et al., 2025; De la Puente et al., 2024; Guo et al., 2023; Hongxia & Razali, 2025; Lo et al., 2025; Miao et al., 2025; Urban et al., 2025 y Zhang, 2025). Así ocurre en Lo et al. (2025), cuyo procedimiento experimental consiste en un ensayo controlado aleatorizado y que muestra una mejora significativa en la calidad de la revisión de ensayos ($p = 0.003106$), con una diferencia media de 3.113 puntos a favor del grupo que utilizó IA. De forma similar, Hongxia & Razali (2025), en un diseño cuasi-experimental con grupo control, evidencian mejoras significativas en escritura académica ($t = 3.993$, $p = 0.000$), observándose incrementos en dimensiones como contenido y estructura.

También en el ámbito de la **escritura argumentativa**, Tai et al. (2025), mediante un diseño pre-post intragrupo, reportan mejoras significativas en la puntuación total de escritura ($t_{(47)} = -6.30$, $p < 0.01$, $d = 1.20$), mientras que Polakova & Ivenz (2024) identifican mejoras en gramática, estructura y concisión a través de la recogida de datos tanto cuantitativos como cualitativos. En esta misma línea, Zhang (2025) demuestra, en un diseño experimental con tres condiciones, que el grupo de uso guiado de IA obtiene mejoras altamente significativas en calidad textual medida mediante *Grammarly for Education Platform Performance Score* ($F_{(2,253)} = 789.53$, $p < 0.001$), lo que introduce un matiz importante y que se repite una vez más: la mejora comunicativa es mayor cuando el uso de la herramienta está pedagógicamente estructurado.

Por otro lado, **el efecto no es uniforme en todos los contextos**. El estudio de Chowdhury Niloy et al. (2023), con un amplio diseño experimental con grupo de control, identifica un deterioro significativo en la puntuación global de creatividad cuando el uso de IA es exclusivo ($p = 0.000$, $r = -0.352$), especialmente en originalidad ($p = 0.000$, $r = -0.977$). Este resultado constituye una divergencia relevante y sugiere que la dependencia desmesurada de la herramienta puede afectar negativamente dimensiones expresivas vinculadas a la creatividad en la escritura, y, por tanto, en la comunicación.

En relación con la **expresión oral y la argumentación**, también se observa una **tendencia mayoritariamente positiva**. Darmawansah et al. (2025), en un diseño cuasi-experimental con grupo control, reportan mejoras significativas en el desempeño del habla argumentativa ($F = 6.428, p = 0.01, \eta^2 = 0.33$). De manera similar, Guo et al. (2023) identifican mejoras significativas en la complejidad estructural de los argumentos en debates asistidos por chatbot, concretamente en el número de afirmaciones ($F_{(1,42)} = 8.600, p = 0.005$) y en la presencia de datos y garantías ($F_{(1,42)} = 5.185, p = 0.028$). No obstante, no se observaron diferencias significativas en otros elementos como contraargumentos o refutaciones, ni en el nivel global de complejidad estructural ($F_{(1,42)} = 0.075, p = 0.785$), lo que sugiere que el impacto se concentra en la expansión estructural del discurso, pero no necesariamente en su mejora global. Asimismo, Karagöl et al. (2025) muestran un impacto sustancial en habilidades de habla académica en presentaciones orales ($F_{(1,57)} = 608.295, p < 0.001, \eta^2 \text{ parcial} = 0.914$), reforzando la evidencia favorable en contextos de comunicación oral estructurada.

Por otra parte, estudios como De la Puente et al. (2024) muestran mejoras significativas en **habilidades de argumentación** dentro de textos elaborados con apoyo de IA ($F_{(1,90)} = 18.92, p < 0.001$), lo que parece sugerir que la herramienta puede favorecer no solo la fluidez, sino también la calidad estructural y la integración de la información en el discurso.

Por último, constituye un caso particularmente interesante el presentado por Fisher et al. (2025), cuyo ensayo controlado aleatorizado muestra efectos selectivos **según la carrera cursada por los estudiantes**. Mientras que en los estudiantes de marketing se observan mejoras significativas en la mayoría de indicadores de creatividad escrita ($p < 0.001$) en *management* la mejora se limita a la fluidez ($F = 18.69, p < 0.001$) y la elaboración ($F = 41,34, p < 0.001$), sin alcanzar significatividad en originalidad ($p = 0.06$). Este resultado introduce una dimensión contextual que sugiere que el impacto del uso de herramientas de inteligencia artificial generativa puede depender del tipo de tarea y del dominio disciplinar en el que estén especializados los sujetos.

De este modo, desde el punto de vista metodológico, el bloque comunicativo presenta **tres patrones distintivos**. En primer lugar, existe una mayor presencia de diseños experimentales con grupo de control en comparación con los bloques social y digital, lo que refuerza la robustez de las conclusiones y una mayor homogeneidad en los resultados que la identificada en el ámbito metacognitivo, pues, mientras en este último la

heterogeneidad deriva en gran medida de la diversidad en la operacionalización del experimento, en el caso de las competencias comunicativas la mayoría de los estudios emplean instrumentos relativamente estandarizados (rúbricas de escritura, escalas de argumentación y métricas automatizadas), lo que facilita la comparación entre grupo de control y experimental.

En segundo lugar, se observa un uso frecuente de rúbricas externas y evaluaciones objetivas del desempeño textual u oral, lo que reduce la dependencia exclusiva de la autopercepción del sujeto. No obstante, esta convergencia positiva no puede interpretarse como un efecto automático del uso de la herramienta, sino que, al igual que en el bloque cognitivo, donde los efectos más sólidos aparecen cuando la IA se integra como apoyo estructurado y de andamiaje dentro de una tarea diseñada pedagógicamente, en el ámbito comunicativo los resultados más positivos se observan en intervenciones con una guía explícita, criterios de evaluación definidos y comparación entre grupos. Por el contrario, cuando el uso por parte del estudiante es sustitutivo del proceso creativo, pueden emerger efectos adversos en dimensiones como la originalidad, lo que conecta con los riesgos de dependencia ya identificados en los bloques cognitivo y metacognitivo.

En conjunto, **la evidencia sugiere que la inteligencia artificial generativa puede potenciar de manera significativa el desarrollo de competencias comunicativas**, especialmente en tareas estructuradas de escritura académica, revisión de textos y argumentación oral, siempre que su integración esté acompañada de criterios pedagógicos claros. Sin embargo, el impacto depende del grado de regulación del uso y del tipo de tarea, pues al estar las competencias comunicativas especialmente vinculadas a la creatividad, son sensibles al deterioro de las dimensiones creativas cuando se produce una externalización excesiva del proceso expresivo. Cuando se desplaza el proceso expresivo del estudiante o se reduce su implicación cognitiva en tareas creativas, el impacto puede volverse selectivo o incluso negativo. Este patrón refuerza una conclusión común a los bloques analizados: el efecto de la GenIA no depende exclusivamente de la herramienta, sino del diseño pedagógico que estructura su uso y del grado de implicación activa que se exige al estudiante en el proceso.

| Tipo de evidencia comunicativa | Forma de medición | Resultado principal observado | Dirección/tipo de evidencia | Estudios ejemplificativos | Matices o limitaciones relevantes |
|---|--|---|-------------------------------------|--|---|
| Expresión escrita | Diseños experimentales /cuasi-experimentales ; métricas como <i>Platform Performance Score</i> . | Incremento sólido en la calidad global de los textos, contenido y organización estructural. | Mejora significativa | Darmawansah et al. (2025); De la Puente et al. (2024); Guo et al. (2023); Hongxia & Razali (2025); Lo et al. (2025); Miao et al. (2025); Urban et al. (2025); Zhang et al. (2025). | Los resultados más potentes se vinculan a un uso guiado y pedagógicamente estructurado . |
| Escritura argumentativa | Diseño pre-post intragrupo y diseño experimental | Mejora en la puntuación global de escritura, así como en gramática, estructura, concisión y en la calidad textual. | Mejora significativa | De la Puente et al. (2024); Lo et al. (2025); Polakova & Ivenz (2024); Zhang et al. (2025). | El uso de la IA favorece no solo la fluidez, sino la calidad estructural y cohesión del discurso. La mejora es mayor cuando el uso de IA está estructurado. |
| Expresión oral y argumentación | Diseño cuasi-experimental y experimental | Aumento en el desempeño del habla argumentativa, complejidad estructural de argumentos y en la presencia de afirmaciones y datos. | Mejora parcial/selectiva | Darmawansah et al. (2025); Guo et al. (2023); Karagöl et al. (2025). | Se mejora la expansión estructural (datos/garantías), pero no la complejidad global ni los contraargumentos. |
| Dimensiones creativas y originalidad | Diseños experimentales con grupo de control. | Deterioro significativo de la creatividad y la originalidad del mensaje. | Efecto negativo o divergente | Chowdhury Niloy et al. (2023). | La dependencia de la IA puede afectar negativamente a la comunicación. |
| Efectos según dominio disciplinar | Ensayos controlados aleatorizados aplicados a diferentes áreas de conocimiento. | Variabilidad de resultados: mejoras globales en unas áreas y restringidas en otras. | Mejora parcial/selectiva | Fisher et al. (2025). | En marketing hay mejoras amplias; en <i>management</i> la originalidad no alcanza significatividad. |

Tabla 5. Síntesis del análisis de las competencias comunicativas.

2.3. Análisis de las barreras detectadas en las intervenciones

El presente análisis se basa exclusivamente en la información contenida en la matriz de codificación relativa a las barreras técnicas, docentes y del alumnado, únicamente considerando como barreras aquellas que los propios estudios reportan de forma explícita. Por otro lado, no se han incluido limitaciones metodológicas generales, pues serán analizadas en el siguiente apartado; salvo en aquellos casos en que dichas limitaciones

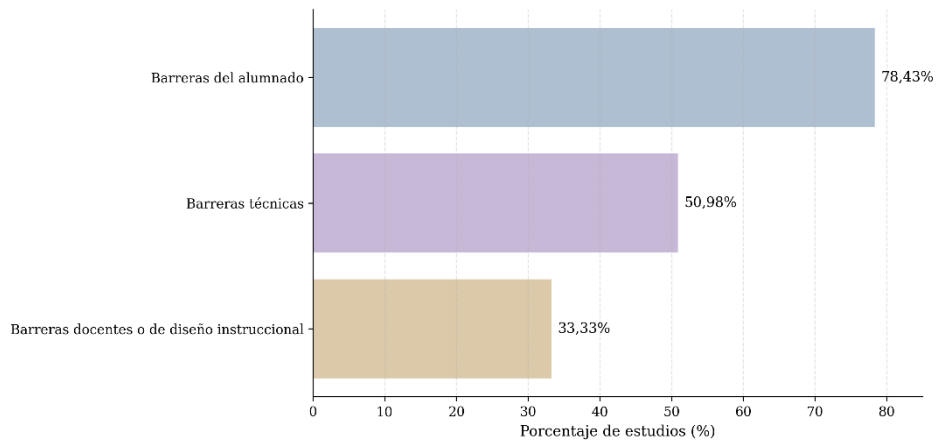
estén formuladas en el artículo como problema derivado directamente del uso de la inteligencia artificial generativa.

2.3.1. Distribución general en la muestra

Sobre la muestra final de 51 estudios, 26 (50,98 %) reportan barreras técnicas (Al-Abri, 2025; Alnemrat et al., 2025; Beninger et al., 2024; Bykov & Gladchenko, 2024; Chen et al., 2025; Darmawansah et al., 2025; de la Puente et al., 2024; de Vicente-Yagüe-Jara et al., 2023; Fisher et al., 2025; Guo et al., 2023; Hongxia & Razali, 2025; Hou et al., 2025; Karagöl et al., 2025; Leahy et al., 2025; Lo et al., 2025; Moheno et al., 2024; Naamati-Schneider & Alt, 2024; Polakova & Ivenz, 2024; Tai et al., 2025; Urban et al., 2024; Urban et al., 2025; Wei et al., 2025; Wu et al., 2025; Zhang, 2025; Zhang et al., 2025a; Zhang et al., 2025b).

En segundo lugar, 17 estudios (33,33 %) identifican barreras docentes o vinculadas al diseño instruccional (Al-Abri, 2025; Alnemrat et al., 2025; Beninger et al., 2025; Bykov & Gladchenko, 2024; Chen et al., 2025; Darmawansah et al., 2025; Guo et al., 2023; Hongxia & Razali, 2025; Hou et al., 2025; Karagöl et al., 2025; Leahy et al., 2025; Moheno et al., 2024; Naamati-Schneider & Alt, 2024; Urban et al., 2025; Wei et al., 2025; Zhang, 2025; Zhang et al., 2025b).

Las barreras del alumnado constituyen la categoría más frecuente: 40 estudios (78,43 %) las reportan explícitamente (Al-Abri, 2025; Alnemrat et al., 2025; Beninger et al., 2024; Bykov & Gladchenko, 2024; Çelik et al., 2024; Chan et al., 2024; Chen et al., 2024; Chen et al., 2025a; Chen et al., 2025b; Chowdhury Niloy et al., 2024; Darmawansah et al., 2025; de la Puente et al., 2024; de Vicente-Yagüe-Jara et al., 2023; Du et al., 2025; Fisher et al., 2025; Guo et al., 2023; Hongxia & Razali, 2025; Hou et al., 2025; Kim et al., 2025; Klyshbekova & Abbott, 2024; Leahy et al., 2025; Lee et al., 2025; Lo et al., 2025; Miao et al., 2025; Mendoza Moheno et al., 2024; Naamati-Schneider & Alt, 2024; Polakova & Ivenz, 2024; Saritepeci & Yildiz Durak, 2024; Selem et al., 2025; Tai et al., 2025; Urban et al., 2024; Urban et al., 2025; Wang et al., 2025; Wei et al., 2025; Wu et al., 2025; Yusof, 2025; Zhang, 2025; Zhang et al., 2025b; Zhong et al., 2024).

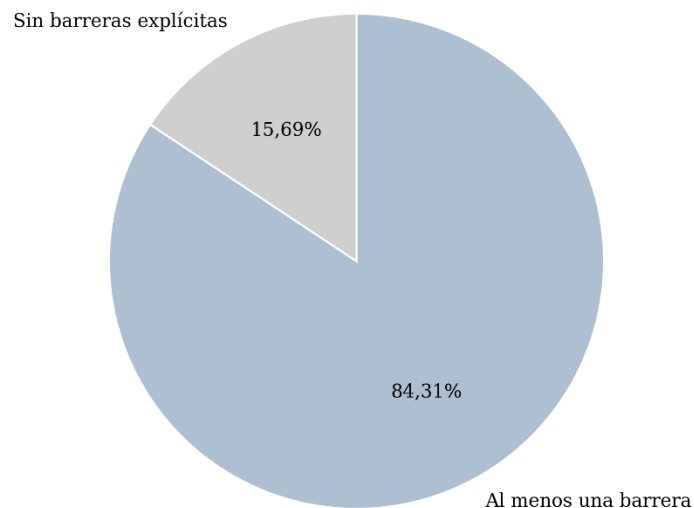


Base: N = 51 estudios. Las categorías no son excluyentes, por lo que un mismo estudio puede registrar simultáneamente barreras técnicas, docentes y del alumnado. Elaboración propia.

Gráfico 7. Frecuencia de los principales tipos de barreras identificadas en los estudios analizados.

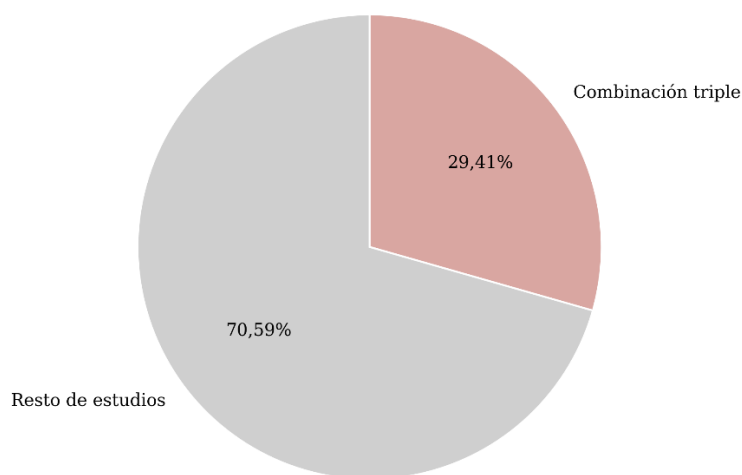
En total, 43 estudios (84,31 %) reportan al menos un tipo de barrera, mientras que 8 (15,69 %) no incluyen ninguna referencia explícita a obstáculos derivados del uso de la IA. Asimismo, el solapamiento es elevado: 15 estudios (29,41 %) combinan barreras técnicas, docentes y del alumnado, lo que sugiere una fuerte conexión entre el diseño de las instrucciones dadas en el experimento y el comportamiento del estudiante.

De este modo, en un principio y solo con estos datos, se puede observar que la mayor parte de los estudios reportan barreras en su investigación.



Base: N = 51 estudios. Elaboración propia a partir de los estudios incluidos en la systematic review.

Gráfico 8. Distribución de los estudios según la presencia de barreras asociadas al uso de la IA.



Base: N = 51 estudios. Elaboración propia a partir de los estudios incluidos en la systematic review.

Gráfico 9. Distribución de los estudios según la presencia de barreras asociadas al uso de la IA.

2.3.2. *Análisis por tipología de barrera*

2.3.2.1. Barreras técnicas

En primer lugar, las **barreras técnicas** identificadas se construyen principalmente en torno a **problemas de precisión, riesgo de alucinaciones y calidad inconsistente de las respuestas**. Este patrón aparece en estudios como Chen et al. (2025), Hou et al. (2025), Urban et al. (2024) y Urban et al. (2025), donde se advierte explícitamente del riesgo de incorporar información incorrecta o superficial y la adopción de sesgos.

En otros trabajos (Lo et al., 2025; Wei et al., 2025 y Zhang, 2025) la **limitación técnica** no se concibe como un fallo intrínseco a la herramienta, sino **como consecuencia contextual**: la calidad depende del tipo de tarea, del nivel del estudiante o del modo de interacción. En estos casos, la barrera aparece más como un riesgo condicionado por el contexto de la intervención que como una deficiencia propia de la herramienta.

En tercer lugar, se encuentran referencias a **limitaciones funcionales o estructurales** (Leahy et al.; 2025), pero son menos frecuentes y suelen aparecer en estudios de implementación curricular amplia (Alnemrat et al., 2025; Beninger et al., 2024; Naamati-Schneider & Alt, 2024), donde se señala la necesidad de competencias digitales previas o formación específica.

2.3.2.2. Barreras docentes o de diseño instruccional

Las **barreras docentes** se centran en la ausencia de guía, el uso no estructurado y la integración superficial de la herramienta. En estudios como Hou et al. (2025), Urban et al. (2025) y Wei et al. (2025), se sugiere que el problema no reside en la herramienta de GenIA en sí, sino en la **falta de mecanismos que pauten y obliguen al estudiante a mantener control sobre el proceso.**

En otros casos (Darmawansah et al., 2025; Guo et al., 2023; Mendoza Moheno et al., 2024), la herramienta demuestra efectos positivos cuando está integrada en tareas estructuradas, lo que refuerza la idea de que **la barrera no es tecnológica sino pedagógica.**

De este modo, la **conexión con los resultados metacognitivos** es clara: cuando el diseño obliga a justificar, revisar o contrastar la producción, disminuye el riesgo de uso irreflexivo y de pereza metacognitiva. Del mismo modo, en competencias comunicativas, estudios como Lo et al. (2025) muestra mejoras cuando el uso está pautado, mientras que los riesgos emergen cuando se permite una delegación total del proceso. En concreto, es muy interesante la idea que defienden Lo et al. (2025) de que **la IA debe ser colaboradora con el intelecto humano para fomentar y no sustituir las capacidades humanas.**

2.3.2.3. Barreras del alumnado

Las **barreras del alumnado** constituyen la tipología más frecuente en la muestra, apareciendo en 40 de los 51 estudios analizados (78,43 %). A diferencia de las barreras técnicas o docentes, que giran en torno a aspectos funcionales o de diseño, este bloque se estructura principalmente en la regulación y estrategias del propio estudiante.

El primer patrón identificado es la **dependencia excesiva o el uso irreflexivo de la herramienta.** Esta barrera ha ido surgiendo a lo largo del presente análisis; sin embargo, es necesario precisar que no todos los estudios que aluden a la dependencia la analizan como variable empírica. En algunos casos, constituye una variable formalmente medida; pero en otros, solamente una advertencia interpretativa derivada de los resultados observados.

El caso con un resultado más robusto en su medición empírica es el análisis de Hou et al. (2025), quienes desarrollan y validan la *Reliance Behaviour Scale*, que incluye

explícitamente la subescala “*thoughtless use*”. La dependencia no es una interpretación, sino que se cuantifica el comportamiento, demostrando que la confianza en la IA predice significativamente el uso irreflexivo ($\beta = 0.47$, $p < 0.001$). De forma similar, Chen et al. (2025a) miden el incremento del “*Executive help-seeking*” en el grupo que utiliza IA, interpretando este patrón como externalización del monitoreo y sobredependencia, buscando los alumnos resultados ejecutables o respuestas directas. En este caso, la dependencia no es una variable del estudio pero sí se deriva de un indicador empírico significativo.

Por otro lado, en Urban et al. (2025), la sobredependencia no se mide como variable autónoma. El estudio evalúa la precisión metacognitiva, sesgo o creencias epistémicas, y la dependencia aparece en el análisis como una consecuencia del aumento de autoeficacia o de la confianza epistémica en la herramienta. Una lógica similar se observa en Chowdhury Niloy et al. (2024), donde el deterioro experimental de la originalidad en escritura creativa se explica parcialmente como posible efecto de dependencia, y en Fisher et al. (2025), donde la sobredependencia se formula como advertencia pedagógica al final de la discusión de la intervención producida.

El segundo patrón identificado es la **sobreconfianza en la herramienta de GenIA**, también llamada mala calibración. Urban et al. (2024) muestran que el incremento de autoeficacia puede no estar alineado con el desempeño real, ya que la sobreconfianza en sus habilidades puede impactar negativamente en el resultado; mientras que Urban et al. (2025) identifican que tanto el sesgo metacognitivo como la confianza epistémica en la IA son predictores de la integración de información errónea. Aunque estos estudios no miden la variable de dependencia como tal, sí documentan que existen desajustes entre percepción y rendimiento que pueden favorecer comportamientos de uso acrítico.

El tercer patrón es la **reducción del esfuerzo cognitivo o la externalización estratégica de procesos de monitoreo**. Este fenómeno aparece explícitamente en Chen et al. (2025a), donde se observa una alteración en el proceso de búsqueda de ayuda, saltándose etapas fundamentales como la evaluación de la información, realizando preguntas operativas con el objetivo de obtener resultados ejecutables o información lista para copiar en lugar de diagnosticar sus dudas proactivamente; así como en Chowdhury Niloy et al. (2024), donde el uso exclusivo de IA se asocia a descensos significativos en la creatividad. Por su parte, Fisher et al. (2025) advierten también del riesgo de que la herramienta impacte negativamente en las capacidades de los estudiantes si no se alinean las necesidades

tanto del método como de los estudiantes, incluso cuando el desempeño mejora en determinadas disciplinas.

Finalmente, algunos estudios mencionan **ansiedad o inseguridad vinculadas al uso de IA** en contextos comunicativos o de evaluación pública, como Karagöl et al. (2025) o Yusof (2025). En estos casos, la barrera no se asocia a dependencia, sino que se trata de una condición preexistente (ansiedad al hablar en público y ante los exámenes) y que pueden modular la interacción con la herramienta.

Un aspecto especialmente relevante es que estas barreras afectivas emergen como consecuencia del uso de la IA incluso en estudios donde el rendimiento mejora significativamente. Por ejemplo, Lo et al. (2025) y Zhang (2025) documentan que la ansiedad puede ser causada o agravada por el uso de la herramienta de GenIA.

Por otro lado, también investigaciones como Lo et al. (2025), Wei et al. (2025) y Zhang (2025) documentan mejoras objetivas en calidad textual, creatividad o desempeño colaborativo, pero simultáneamente advierten del riesgo de uso superficial o delegación excesiva en la herramienta. Esto sugiere que **la mejora en el producto final no excluye la presencia de pérdida de control del proceso metacognitivo sin una estrategia pedagógica clara.**

En conjunto, las barreras del alumnado no apuntan a que el uso de IA por el alumnado conlleve un deterioro lineal del aprendizaje, sino a una **necesidad de estructuración** de los mecanismos de control, monitoreo y atribución de agencia.

En síntesis, la evidencia muestra que las barreras se concentran mayoritariamente en el plano metodológico y educativo más que en el estrictamente técnico, pues el impacto negativo no parece inherente a la herramienta, sino condicionado por el modo de integración pedagógica y por el nivel de autorregulación del estudiante. La coexistencia de mejora en desempeño y riesgo de dependencia no apunta a un deterioro lineal del aprendizaje, sino a una reconfiguración del proceso cognitivo: la IA redistribuye la carga, altera las fases de monitoreo y obliga a redefinir qué se entiende por competencia autónoma en entornos mediados tecnológicamente.

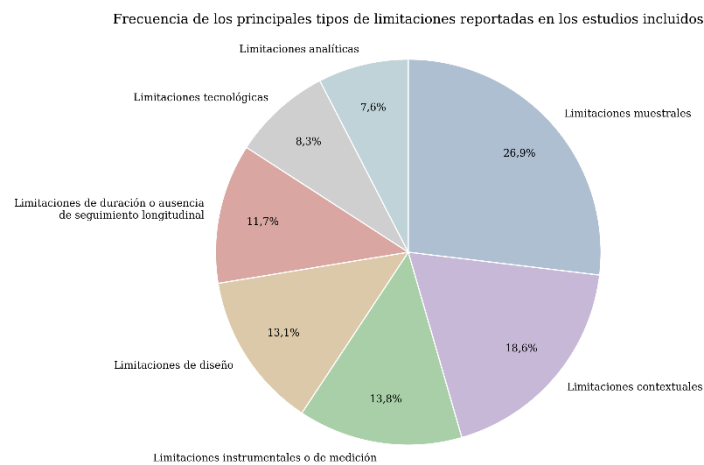
| Tipo de barrera | Subtipologías o patrones identificados | Manifestaciones concretas | Estudios ejemplificativos |
|------------------------------|---|---|---|
| Barreras técnicas | <p>Déficit de fiabilidad y calidad: Problemas intrínsecos de precisión y veracidad.</p> <p>Riesgo condicionado: Limitaciones derivadas del contexto de interacción.</p> <p>Limitaciones estructurales: Carencias en la base funcional o competencial.</p> | <p>Alucinaciones, incorporación de información incorrecta o superficial y adopción de sesgos.</p> <p>Dependencia de la calidad según el tipo de tarea, nivel del usuario o modo de interacción.</p> <p>Necesidad de competencias digitales previas o formación específica.</p> | <p>Chen et al. (2025), Hou et al. (2025), Urban et al. (2024, 2025).Lo et al.(2025), Wei et al. (2025), Zhang (2025).Leahy et al. (2025), Alnemrat et al. (2025), Beninger et al. (2024), Naamati-Schneider y Alt (2024).</p> |
| Barreras docentes | <p>Ausencia de guía y estructuración: Integración superficial o no pautada de la GenIA.</p> <p>Falta de mecanismos de control: Diseños que no fomentan la supervisión humana.</p> | <p>Falta de instrucciones que obliguen al estudiante a mantener el control sobre el proceso.</p> <p>Emergencia de riesgos ante la delegación total del proceso en la IA.</p> | <p>Hou et al. (2025), Urban et al. (2025), Wei et al. (2025). Darmawansah et al. (2025), Guo et al. (2023), Mendoza Moheno et al. (2024).</p> |
| Barreras del alumnado | <p>Dependencia excesiva o uso irreflexivo: Externalización del monitoreo (<i>thoughtless use</i>).</p> <p>Sobreconfianza: Desajuste entre percepción y desempeño real.</p> <p>Reducción del esfuerzo cognitivo: Búsqueda de resultados ejecutables y "listos para copiar".</p> | <p>Variables medidas empíricamente: Escalas de comportamiento de dependencia y aumento del <i>Executive help-seeking</i>.</p> <p>Advertencias interpretativas: Riesgos de pérdida de originalidad y desajustes de autoeficacia.</p> <p>Salto de etapas de evaluación; ansiedad causada o agravada por la interacción con la IA.</p> | <p>Hou et al. (2025), Chen et al. (2025a), Urban et al. (2024, 2025), Chowdhury Niloy et al. (2024), Fisher et al. (2025).Karagöl et al. (2025), Yusof (2025), Lo et al. (2025), Zhang (2025).</p> |

Tabla 6. Síntesis sobre las tipologías de barreras en el uso de la IA generativa por parte del alumnado universitario.

2.4. Análisis de las limitaciones de los estudios incluidos

Para el análisis de las limitaciones de los estudios incluidos, únicamente se han considerado aquellas limitaciones que los propios autores reconocen de forma explícita en los apartados de discusión, conclusiones o secciones específicas de limitaciones. Por tanto, no se han reinterpretado resultados como limitaciones si los autores no las formulan como tales, ni se han inferido debilidades metodológicas adicionales más allá de lo expresamente declarado en los textos originales.

A partir del contenido de las limitaciones reconocidas, se ha optado en este análisis por categorizar **siete grandes limitaciones**: muestrales, contextuales, de diseño, de duración o ausencia de seguimiento longitudinal, instrumentales o de medición, tecnológicas y analíticas.



Base: N = 51 estudios. Las categorías no son excluyentes, por lo que los porcentajes no representan partes cerradas de un mismo total, sino la frecuencia de aparición de cada tipo de limitación en la muestra analizada. Elaboración propia.

Gráfico 10. Frecuencia de los principales tipos de limitaciones reportadas en los estudios incluidos.

2.4.1. Limitaciones muestrales

Las **limitaciones relativas al tamaño o composición de la muestra** aparecen en 39 de los 51 estudios, lo que representa el 76,47 %. Esta limitación incluye muestras pequeñas, no probabilísticas, homogéneas o centradas en una única titulación o perfil académico. Entre los estudios que reconocen este tipo de limitaciones se encuentran, entre otros, Bykov y Gladchenko (2024), Chen et al. (2025b), Çelik et al. (2024), Chowdhury Niloy et al. (2024), Darmawansah et al. (2025), de la Puente et al. (2024), de Vicente-Yagüe-Jara et al. (2023), Du et al. (2025a), Du et al. (2025b), Guo et al. (2023), Hongxia & Razali (2025), Huang & Mizumoto (2024), Karagöl et al. (2025), Kim et al. (2025), Le et al. (2025), Lee et al. (2025), Lo et al. (2025), Mendoza Moheno

et al. (2024), Polakova & Ivenz (2024), Šedlbauer et al. (2024), Usher & Amzalag (2025), Wang et al. (2025), Wu et al. (2025) y Yusof (2025).

Esta limitación es estructural y afecta principalmente a la validez externa y a la capacidad de generalización de los resultados.

2.4.2. Limitaciones contextuales

En 27 estudios (52,94 %) se identifican **limitaciones derivadas del contexto específico** de aplicación, como una única universidad, un solo país, una disciplina concreta o una tarea específica. Esta dimensión aparece, por ejemplo, en Çelik et al. (2024), Hongxia & Razali (2025), Karagöl et al. (2025), Le et al. (2025), Lo et al. (2025), Oliveira et al. (2025), Selem et al. (2025) y Wang et al. (2025).

Estas limitaciones son también de carácter estructural en cuanto condicionan la extrapolación de los hallazgos, pero también tienen un componente contextual claro vinculado a la implementación concreta del estudio.

2.4.3. Limitaciones de diseño

Las **limitaciones vinculadas al diseño metodológico** se reconocen en 19 estudios, lo que equivale al 37,25 %. Incluyen ausencia de grupo de control (Al-Abri, 2025 y Tai et al., 2025), diseños pretest-postest de un solo grupo (Wu et al., 2025), ausencia de control de variables de fondo, sin aislar el componente del chatbot (Guo et al., 2023), configuraciones experimentales de laboratorio con baja validez ecológica (Urban et al., 2024; Le et al., 2025), o limitaciones en la asignación de participantes (Bykov y Gladchenko, 2024).

Se trata de limitaciones estructurales que afectan a la validez interna y a la solidez inferencial de los resultados.

2.4.4. Limitaciones de duración o ausencia de seguimiento longitudinal

En 17 estudios (33,33 %) los autores reconocen **intervenciones de corta duración o ausencia de seguimiento longitudinal**. Esto se observa, entre otros, en Chen et al. (2025b), Çelik et al. (2024), de la Puente et al. (2024), Hongxia & Razali (2025), Huang (2024), Lee et al. (2025), Polakova & Ivenz (2024), Tai et al. (2025) y Wang et al. (2025).

Esta limitación es estructural, en la medida en que impide evaluar la estabilidad temporal de los efectos y la posible aparición de fenómenos como la dependencia o el efecto novedad.

2.4.5. Limitaciones instrumentales o de medición

En 20 estudios (39,22 %) se reconocen **limitaciones relacionadas con los instrumentos empleados**. Incluyen el uso de autoinformes (Zhang, 2025; Oliveira et al., 2025), ausencia de triangulación metodológica (Le et al., 2025), evaluación mediante un único corrector (Alnemrat et al., 2025), insuficiente validación de instrumentos adaptados (Le et al., 2025), o dificultades para medir procesos cognitivos internos (Chen et al., 2025a).

Estas limitaciones son estructurales, ya que afectan directamente a la fiabilidad y validez de las medidas utilizadas para evaluar competencias.

2.4.6. Limitaciones tecnológicas

En 12 estudios (23,53%) se identifican **limitaciones vinculadas específicamente a la herramienta o versión de IA utilizada**. Incluyen el uso de un único sistema (Guo et al., 2023), pruebas con versiones concretas del modelo (Klyshbekova & Abbott, 2024), restricciones de acceso en determinados países (Hongxia & Razali, 2025), o limitaciones derivadas del tipo de herramienta empleada (Fabus et al., 2025).

Estas limitaciones son en parte contextuales, pero también específicas del estado evolutivo de la tecnología.

2.4.7. Limitaciones analíticas

En 11 estudios (21,57 %) se reconocen **limitaciones relacionadas con la profundidad del análisis o la falta de determinados controles estadísticos o conceptuales**. Entre ellas se incluyen ausencia de corrección por comparaciones múltiples (Miao et al., 2025), falta de análisis diferenciado por niveles iniciales de competencia (Chen et al., 2025b), o insuficiente exploración de mecanismos explicativos (Le et al., 2025).

Se trata de limitaciones estructurales que afectan a la interpretación y alcance de los resultados.

2.4.8. Análisis global de las limitaciones encontradas

Un análisis global de los estudios denota que el tipo de limitación más frecuente es la muestral (76,47 %), seguida por las limitaciones instrumentales (39,22 %) y de diseño

(37,25 %). Por su parte, las limitaciones tecnológicas (23,53 %) aparecen con menor frecuencia que las metodológicas clásicas.

De este modo, **predominan claramente las limitaciones metodológicas tradicionales como muestra, diseño y medición**, frente a aquellas específicas del uso de inteligencia artificial; lo cual demuestra que, al igual que en las barreras analizadas, la mayor parte de ellas son extrínsecas a la herramienta de GenIA. Incluso en estudios que abordan explícitamente el riesgo de sobredependencia tecnológica (Le et al., 2025 y Wei et al., 2025), las limitaciones reconocidas se centran más en la robustez del diseño y la medición, que en la herramienta en sí misma.

Además, en términos generales, los estudios experimentales tienden a reconocer tanto limitaciones internas (diseño, medición) como externas (muestra, contexto), mientras que los estudios cuasiexperimentales o exploratorios enfatizan más las restricciones muestrales y contextuales, lo que es coherente con el diseño de la intervención.

Se observan varios **patrones significativos**. En primer lugar, numerosos estudios que reportan resultados positivos en términos de mejora competencial reconocen simultáneamente limitaciones relevantes de muestra o duración (por ejemplo, Chen et al., 2025b; Hongxia & Razali, 2025 y Wang et al., 2025), lo que introduce una cautela metodológica coherente con la magnitud de los hallazgos.

En segundo lugar, algunos estudios reconocen riesgos asociados al uso de IA, como la sobredependencia o influencia del efecto novedad (Lee et al., 2025; Wei et al., 2025; Le et al., 2025), sin que ello implique necesariamente una debilidad tecnológica estructural, sino más bien una limitación en el diseño longitudinal.

Por último, en varios trabajos las limitaciones afectan directamente a la interpretación de los resultados competenciales, especialmente cuando se emplean autoinformes o no se evalúa de manera objetiva la calidad del desempeño (Alnemrat et al., 2025; Oliveira et al., 2025 y Zhang, 2025).

3. CONCLUSIÓN

3.1. Propósito y objetivos del estudio

Esta revisión sistemática se planteó con el **objetivo** de analizar la literatura reciente sobre el impacto del uso de herramientas de inteligencia artificial generativa por parte del alumnado universitario, teniendo en cuenta no solo su efecto sobre las competencias estudiadas, sino también las condiciones de las intervenciones, las barreras detectadas en su implementación y las limitaciones de los estudios. Desde este enfoque, el **propósito** ha sido tratar de esclarecer en qué condiciones la integración de la GenIA parece asociarse a resultados más favorables, ambiguos o problemáticos.

3.2. Principales hallazgos y contraste con otras revisiones de la literatura

El análisis de los 51 estudios revisados revela que la integración de la GenIA en el ámbito de educación universitaria no constituye una mejora automática del aprendizaje, sino que su **impacto es complejo y condicionado por el diseño pedagógico y el nivel de agencia** del estudiante (Makranksy et al., 2025; Zhang, 2025). El 94,12% de los estudios presenta un **rol activo del alumno**, quien, principalmente, suele emplear la herramienta para el aprendizaje autorregulado (Huang, 2024; Le et al., 2025; Lee et al., 2025 y Makranksy et al., 2025) y la producción textual (De Vicente-Yagüe-Jara et al., 2023; Fabus et al., 2025; Klyshbekova & Abbott, 2024 y Sedlbauer et al., 2024). Por otra parte, resulta relevante que una **gran mayoría de las intervenciones carezcan de formación previa** sobre el funcionamiento de la herramienta de GenIA, lo cual impacta de manera directa en los resultados medidos sobre las competencias, especialmente las digitales.

Respecto del **análisis competencial**, los resultados son especialmente favorables en las competencias cognitivas y comunicativas, en las que se observa un impacto positivo en la calidad de los textos ((Hongxia & Razali, 2025 y Tai et al., 2025), pensamiento crítico (Chen et al., 2025b; Oliva-Córdova et al., 2025 Du et al., 2025 y Zhang et al., 2025b) y la expresión oral (Darmawansah et al., 2025). No obstante, existe ambigüedad en lo que respecta al pensamiento creativo (Selem et al., 2025) y las competencias cognitivas, pues se ha detectado un sesgo de utilidad mediante el cual, los estudiantes sobreestiman la calidad de su trabajo debido a la consideración de ChatGPT como útil (Urban et al., 2024); lo que puede derivar en una pereza metacognitiva (Kim et al., 2025 y Selem et al., 2025), omitiéndose etapas clave del aprendizaje. Por tanto, la GenIA actúa como una **variable puente** (Zhang et al., 2025b): si se integra como un **andamiaje guiado**, potencia las

habilidades de orden superior, pero si se utiliza de manera **desestructurada**, fomenta la **dependencia**, erosionando la creatividad y la capacidad de autopercepción del trabajo realizado (Kim et al., 2025; Selem et al., 2025 y Urban et al., 2024).

Estos hallazgos respecto de las competencias se sitúan en una **línea coincidente con otras revisiones de la literatura** que, pese a no analizar exactamente las mismas competencias ni diseños de intervención, subrayan también el impacto de la inteligencia artificial generativa en la educación. En concreto, Chiu et al. (2023) analiza, entre otras, las competencias de colaboración, creatividad y habilidades de autorregulación, obteniendo resultados coherentes con esta revisión: las plataformas de feedback en tiempo real generan mejoras en las competencias de colaboración, capacidad de resolución de problemas, habilidades comunicativas y la capacidad de aprendizaje autodirigido. Sin embargo, sus resultados positivos en creatividad son divergentes respecto de esta revisión sistemática. Por otro lado, subrayan que la falta de familiaridad con las tecnologías de IA dificulta su integración en las universidades.

Cebrián et al. (2024), coincide con la necesidad de desarrollar una alfabetización de la IA en los estudiantes previa a su empleo para que puedan construir su propio análisis crítico de las herramientas. Resalta también la importancia de fomentar la participación del alumnado en su proceso de aprendizaje, para que mantenga su capacidad de decisión, criterio y responsabilidad, lo cual concuerda con el análisis extraído en esta revisión de la necesidad de un andamiaje guiado. Además, esta misma revisión sistemática reporta también mejoras en la producción escrita con el uso de GenIA.

Las **barreras** identificadas en los estudios son más recurrentes en el caso del alumnado (78,43%), destacando el uso irreflexivo (Hou et al., 2025), la sobreconfianza (Urban et al., 2024) y la reducción del esfuerzo cognitivo (Chen et al., 2025a). Asimismo, estas barreras pueden ser agravadas por las barreras técnicas, como las alucinaciones de la IA (Chen et al., 2025; Hou et al., 2025; Urban et al., 2024; y Urban et al., 2025), y por las barreras docentes, como la falta de mecanismos de control que obliguen al estudiante a mantener la supervisión sobre el resultado final (Hou et al., 2025 y Wei et al., 2025). En esta línea, la revisión sistemática de Chiu et al. (2023) desarrolla el concepto de “*black box*” o caja negra, para describir la falta de comprensión docente sobre el funcionamiento de la IA, y Cebrián Cifuentes et al. (2024) apunta la falta de precisión de ChatGPT, con una media del 63,41%.

Asimismo, las limitaciones de los estudios muestran los **puntos metodológicos en los que las próximas investigaciones deben hacer hincapié**: el predominio de muestras pequeñas y no probabilísticas (76,47%); las limitaciones contextuales, en particular de disciplina, debido a los resultados dispares en las distintas especialidades en el estudio de Fisher et al. (2025), cuestión en la que merece la pena profundizar; así como en intervenciones con seguimiento longitudinal que midan el impacto a largo plazo del uso de la herramienta. En línea con esta última limitación, la revisión de Chiu et al. (2023) apunta en el mismo sentido, resaltando la falta de perspectivas educativas en la investigación sobre aplicaciones de IA en educación, estando centrada en el campo de la ingeniería.

3.3. Contribución del estudio

La relevancia de estos hallazgos, en línea con los de otros estudios similares, se basa en que la evidencia revisada subraya que el **éxito o fracaso del uso de la GenIA** en el ámbito de la educación universitaria no es intrínseco a la herramienta, sino que **depende** de variables como el **grado de estructuración de la intervención, la guía docente y la formación previa para una óptima alfabetización en IA que promueva un uso eficaz de la herramienta**. En este sentido, es fundamental comprender que la **mejora del resultado final no debe confundirse con un desarrollo competencial efectivo** si no existe un seguimiento del proceso, pues esto genera dependencia, pereza cognitiva y descarga metacognitiva, perdiendo el estudiante el control sobre su propio aprendizaje. Por tanto, es necesario **evitar** un uso instrumental y sustitutivo de las herramientas de GenIA, y **proponer** un modelo de colaboración, en el que la herramienta refuerce y no reemplace la implicación cognitiva y responsabilidad del estudiante en su aprendizaje.

3.4. Limitaciones

Como todo estudio, este trabajo presenta **limitaciones**. En primer lugar, la búsqueda se ha llevado a cabo en una única base de datos, lo que ha podido dejar fuera estudios relevantes contenidos en otras bases. En segundo lugar, la delimitación mediante criterios de inclusión específicos reduce el alcance de la muestra revisada, y la heterogeneidad de los estudios incluidos (diseño, instrumentos, definiciones, intervenciones...) dificulta una comparación directa entre resultados y la posibilidad de formular unas conclusiones más homogéneas. Por otra parte, el análisis de los estudios se ha realizado de manera manual con apoyo de NotebookLM, lo que dificulta el grado de exhaustividad del examen

realizado, pudiendo haberse producido errores de categorización o interpretación. Finalmente, el análisis está ligado a la información declarada por los autores, e inevitablemente, a sus barreras y limitaciones.

Todas estas limitaciones no invalidan la aportación realizada por esta revisión, pero aconsejan una interpretación crítica y fundamentada de su alcance.

3.5. Recomendaciones para investigaciones futuras

Partiendo de las limitaciones detectadas en los estudios analizados, resulta necesario avanzar hacia investigaciones con **diseños longitudinales** que observen efectos sostenidos en el tiempo más allá del impacto inmediato, así como muestras y contextos más amplios, que incorporen diversas disciplinas.

Del mismo modo, se requieren **instrumentos de medición** que posibiliten la comparación de resultados y diferencia con mayor precisión entre resultado final, procesos de autorregulación, alfabetización en IA y el desarrollo efectivo de competencias. También es necesaria la **profundización en el estudio de variables como la guía docente estructurada o pautada**, la disciplina académica y el nivel de formación previa del estudiante.

Por último, para futuras investigaciones, es especialmente interesante el estudio de la **inteligencia artificial generativa como variable puente**, haciendo un estudio comparativo entre grupos con y sin formación previa, con el objetivo de medir el efecto del andamiaje en cada intervención, la dependencia a la herramienta desarrollada y la autopercepción calibrada de los estudiantes.

3.6. Recomendaciones prácticas

Tras este análisis sistemático, se puede concluir que existe un impacto evidente en la educación actual y futura del alumnado. En este contexto, es responsabilidad de las instituciones, tanto gubernamentales como educativas, ofrecer una formación que invite a los alumnos a fomentar las capacidades que potencien su rendimiento presente, para convertirse en un futuro en los profesionales que este mundo necesita.

A partir de los hallazgos obtenidos, resulta posible trasladar las conclusiones del estudio a recomendaciones concretas dirigidas a los principales actores implicados en el uso de la inteligencia artificial generativa en la educación superior.

3.6.1. Recomendaciones para instituciones públicas

En primer lugar, a nivel institucional y regulatorio, es necesario **desarrollar un marco institucional** que oriente su uso ético y responsable, enseñando a los alumnos cuando y de qué manera se han de utilizar estas herramientas. En términos prácticos, esto implica:

- **Desarrollar guías oficiales de uso de IA en educación superior**, que establezcan qué se considera un uso legítimo, qué prácticas deben declararse y cómo, y qué conductas constituyen fraude académico.
- **Incorporar la alfabetización en IA como una competencia transversal obligatoria**, especialmente en los primeros cursos para sentar las bases necesarias para un buen uso de la herramienta, enfocándose en aspectos como la verificación de información, detección de errores y uso crítico de la herramienta.
- **Financiar programas de formación docente aplicada**, centrados tanto en el funcionamiento técnico de la IA como en su integración pedagógica en las distintas asignaturas.
- **Garantizar un acceso equitativo a herramientas de IA**, del mismo modo que muchas universidades facilitan a su alumnado el acceso a bases de datos y recursos académicos de pago, a fin de evitar que la capacidad económica del estudiante condicione la calidad de las herramientas utilizadas y, con ello, sus oportunidades de aprendizaje.
- **Adaptar los sistemas de evaluación universitaria**, incentivando metodologías que ponderen el proceso, el razonamiento y la capacidad crítica, reduciendo la importancia del resultado final sobre las calificaciones.

3.6.2. Recomendaciones para el profesorado universitario

Desde la **perspectiva del profesorado**, es clave dejar atrás la aceptación o prohibición de las herramientas de inteligencia artificial generativa y fomentar la integración universitaria de la GenIA desde un diseño pedagógico. Por este motivo, se proponen las siguientes recomendaciones:

- **Valorar el perfil del alumnado y las competencias que conviene desarrollar prioritariamente en cada asignatura**, puesto que la evidencia analizada indica que el impacto de la IA no es uniforme y puede variar según el área de estudio, el nivel previo del estudiante y el tipo de capacidad que se pretenda fortalecer, como la creatividad, el pensamiento crítico o la autorregulación.

- **Introducir una formación inicial específica sobre el uso de la IA en la propia asignatura**, de forma que el alumnado conozca desde el principio para qué tareas puede utilizarla, qué usos no son admisibles y cómo emplearla de manera crítica sin sustituir su propio proceso de aprendizaje
- **Diseñar actividades por fases diferenciadas, en las que el uso de IA esté delimitado**. Por ejemplo: una primera entrega sin IA, una segunda con apoyo de la herramienta y una tercera centrada en la revisión crítica de los resultados generados.
- **Exigir trazabilidad del uso de la IA**, mediante la inclusión de un anexo en los trabajos donde el estudiante deba detallar los prompts utilizados, respuestas obtenidas y decisiones tomadas.
- **Evaluar el proceso, no solo el resultado**, incorporando borradores y reflexiones o comparativas sobre el contenido entregado.
- **Incluir defensas orales**, que permitan verificar la comprensión real del estudiante sobre el contenido entregado.
- **Definir explícitamente qué usos de la IA están permitidos para cada actividad**. Por ejemplo: permitir su uso para revisión gramatical y formal, pero no para la elaboración íntegra del contenido.
- **Utilizar rúbricas que valoren la capacidad crítica**, incluyendo criterios como la verificación de información, detección de errores de la IA o la calidad de la argumentación propia.

3.6.3. Recomendaciones para el alumnado universitario

Por último, **desde la perspectiva del estudiante**, la evidencia analizada sugiere que el uso efectivo de la IA requiere un cambio en la forma de interactuar con la herramienta, por lo que se recomienda:

- **Utilizar la IA como herramienta de apoyo y no como sustituto**: para generar ideas, revisar textos o contrastar argumentos, pero no para saltar etapas del proceso de aprendizaje.
- **Verificar de manera sistemática la información generada y las fuentes**: la trazabilidad de la información es clave para contrastar los datos, referencias y no introducir información errónea.
- **Reformular siempre en palabras propias los contenidos generados**, para asegurar la comprensión real de la materia.

- **Evitar la aceptación automática de las respuestas**, adoptando una actitud crítica frente al contenido generado por la herramienta.
- **Formular preguntas que requieran análisis, comparación y reflexión, y no limitarse a solicitar respuestas cerradas o inmediatas**, utilizando la IA como una herramienta de diálogo que permita profundizar en los contenidos y estimular una comprensión más elaborada.
- **Mantener un análisis crítico sobre tu propio aprendizaje al utilizar la IA**, para identificar cuando su uso mejora el aprendizaje y cuando se reduce a evitar el esfuerzo.

Con el fin de sintetizar y facilitar la comprensión de las recomendaciones expuestas, se incluyen a continuación una serie de infografías que recogen las principales recomendaciones prácticas del estudio para los agentes implicados.

Uso de GenIA en la educación superior

RECOMENDACIONES PARA LAS INSTITUCIONES PÚBLICAS

Las instituciones deben establecer un marco claro que garantice un uso responsable, equitativo y pedagógicamente adecuado de la IA en la universidad.



NORMATIVA DE USO

Definir qué usos son legítimos, cuáles deben declararse y qué prácticas constituyen fraude académico.



ALFABETIZACIÓN EN IA

Formación obligatoria desde primeros cursos en verificación de información, detección de errores y uso crítico.





FORMACIÓN DOCENTE

Impulsar programas que combinen uso técnico de la IA e integración pedagógica en asignaturas.



ACCESO EQUITATIVO

Garantizar acceso institucional a herramientas de IA para evitar desigualdades económicas.





EVALUACIÓN UNIVERSITARIA

Fomentar sistemas que prioricen el proceso, el razonamiento y la capacidad crítica.

Universidad Pontificia Comillas

IMPACTO DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EL DESARROLLO COMPETENCIAL DEL ALUMNADO UNIVERSITARIO: UNA REVISIÓN SISTEMÁTICA DE LA LITERATURA

RECOMENDACIONES PARA EL PROFESORADO

El impacto de la IA depende del diseño de la asignatura. Su uso debe integrarse de forma guiada para reforzar el aprendizaje, no sustituirlo

- **PERFIL DEL ALUMNADO**
Adaptar el uso de la IA según el nivel, el área y las competencias a desarrollar.
- **FORMACIÓN INICIAL**
Explicar desde el inicio usos permitidos, límites y riesgos de la herramienta.
- **ACTIVIDADES POR FASES**
Combinar trabajo sin IA, uso guiado y revisión crítica de resultados.
- **TRAZABILIDAD DEL USO**
Exigir anexo con prompts, respuestas y decisiones del estudiante.
- **EVALUACIÓN DEL PROCESO**
Incluir borradores, versiones intermedias y reflexión sobre el trabajo.
- **DEFENSAS ORALES**
Comprobar la comprensión real del estudiante.
- **USOS PERMITIDOS**
Definir claramente qué se puede y qué no se puede hacer con IA en cada tarea.
- **RÚBRICAS CRÍTICAS**
Evaluar verificación de información, detección de errores y calidad argumentativa.

Universidad Pontificia Comillas

IMPACTO DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EL DESARROLLO COMPETENCIAL DEL ALUMNADO UNIVERSITARIO: UNA REVISIÓN SISTEMÁTICA DE LA LITERATURA

RECOMENDACIONES PARA EL ALUMNADO

El uso eficaz de la IA requiere una actitud activa, crítica y consciente del propio proceso de aprendizaje.



IA COMO APOYO, NO SUSTITUTO

Utilizarla para generar ideas o revisar, no para evitar el aprendizaje.

VERIFICACIÓN DE INFORMACIÓN

Comprobar datos, fuentes y contenidos antes de utilizarlos.



REFORMULAR EN PALABRAS PROPIAS

Asegurar la comprensión real de la materia.

ACTITUD CRÍTICA

No aceptar automáticamente las respuestas generadas.



PREGUNTAS DE ANÁLISIS

Plantear preguntas que exijan reflexión y no respuestas inmediatas.

AUTOCONTROL DEL APRENDIZAJE

Identificar cuándo la IA ayuda a aprender y cuándo sustituye el esfuerzo.



Universidad Pontificia Comillas

IMPACTO DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EL DESARROLLO COMPETENCIAL DEL ALUMNADO UNIVERSITARIO: UNA REVISIÓN SISTEMÁTICA DE LA LITERATURA

4. DECLARACIÓN DE USO DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL GENERATIVA EN TRABAJOS DE GRADO

ADVERTENCIA: Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Nazaret Martínez Carcelén, estudiante de ADE y Derecho de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado “Impacto de la inteligencia artificial generativa en el desarrollo competencial del alumnado universitario: una revisión sistemática de la literatura” declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación [el alumno debe mantener solo aquellas en las que se ha usado ChatGPT o similares y borrar el resto. Si no se ha usado ninguna, borrar todas y escribir “no he usado ninguna”]:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
3. **Interpretador de código:** Para realizar análisis de datos preliminares.
4. **Estudios multidisciplinares:** Para comprender perspectivas de otras comunidades sobre temas de naturaleza multidisciplinar.
5. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
6. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
7. **Generador previo de diagramas de flujo y contenido:** Para esbozar diagramas iniciales.
8. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.

9. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 16 de marzo de 2026.

Firma:

A handwritten signature in black ink, appearing to read 'Mazatuniz'.

5. BIBLIOGRAFÍA

Los artículos marcados con un asterisco (*) son aquellos que forman parte de la muestra de esta revisión sistemática.

- *Al-Abri, A. (2025). Exploring ChatGPT as a virtual tutor: A multi-dimensional analysis of large language models in academic support. *Education and Information Technologies*. Advance online publication. <https://doi.org/10.1007/s10639-025-13484-x>
- *Al-Obaydi, L. H., Pikhart, M., & Hossain, M. K. (2025). ChatGPT and the development of core language skills: An exploratory study of EFL college students. *Contemporary Educational Technology*, 17(4), ep591. <https://doi.org/10.30935/cedtech/17242>
- *Alnemrat, A., Aldamen, H., Almashour, M., Al-Deaibes, M., & AlSharefeen, R. (2025). AI vs. teacher feedback on EFL argumentative writing: A quantitative study. *Frontiers in Education*, 10, 1614673. <https://doi.org/10.3389/feduc.2025.1614673>
- *Beninger, S., Reppel, A., Stanton, J., & Watson, F. (2025). Facilitating generative AI literacy in the face of evolving technology: Interventions in marketing classrooms. *Journal of Marketing Education*, 47(2), 112–125. <https://doi.org/10.1177/02734753251316569>
- *Bykov, I., & Gladchenko, I. (2024). How students develop creative skills and critical thinking with ChatGPT: Experimental results. *European Journal of Contemporary Education*, 13(4), 638–648. <https://doi.org/10.13187/ejced.2024.4.638>
- Cebrián Cifuentes, S., Cano Moya, I., Villalba Martínez, R. & Guerrero Valverde, E. (2024). Revisión sistemática sobre el uso de la IA en la educación superior. *Revista científico profesional de la pedagogía y psicopedagogía*, (9), 117-129.
- *Chan, S. T. S., Lo, N. P. K., & Wong, A. M. H. (2024). Enhancing university level English proficiency with generative AI: Empirical insights into automated feedback and learning outcomes. *Contemporary Educational Technology*, 16(4), ep541. <https://doi.org/10.30935/cedtech/15607>

- *Chen, A., Xiang, M., Zhou, J., Jia, J., Shang, J., Li, X., Gašević, D., & Fan, Y. (2025). Unpacking help-seeking process through multimodal learning analytics: A comparative study of ChatGPT vs human expert. *Computers & Education*, 226, 105198. <https://doi.org/10.1016/j.compedu.2024.105198>
- *Chen, X., Jia, B., Peng, X., Zhao, H., Yao, J., Wang, Z., & Zhu, S. (2025). Effects of ChatGPT and argument map (AM)-supported online argumentation on college students' critical thinking skills and perceptions. *Education and Information Technologies*, 30, 17623–17658. <https://doi.org/10.1007/s10639-025-13471-2>
- Chiu, T. K. J., (2023). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- *Chowdhury Niloy, A., Akter, S., Sultana, N., Sultana, J., & Ur Rahman, S. I. (2024). Is ChatGPT a menace for creative writing ability? An experiment. *Journal of Computer Assisted Learning*, 40 (2), 919–930. <https://doi.org/10.1111/jcal.12929>
- *Çelik, F., Yangın Ersanlı, C., & Arslanbay, G. (2024). Does AI simplification of authentic blog texts improve reading comprehension, inferencing, and anxiety? A one-shot intervention in Turkish EFL context. *International Review of Research in Open and Distributed Learning*, 25(3). <https://doi.org/10.19173/irrodl.v25i3.7779>
- *Darmawansah, D., Rachman, D., Febiyani, F., & Hwang, G. J. (2025). ChatGPT-supported collaborative argumentation: Integrating collaboration script and argument mapping to enhance EFL students' argumentation skills. *Education and Information Technologies*, 30 (3), 3803–3827. <https://doi.org/10.1007/s10639-024-12986-4>
- *De la Puente, M., Torres, J., Blanco Troncoso, A. L., Hernández Meza, Y. Y., & Marquez Carrascal, J. X. (2024). Investigating the use of ChatGPT as a tool for enhancing critical thinking and argumentation skills in international rel

ations debates among undergraduate students. *Smart Learning Environments*, 11 (1), 55. <https://doi.org/10.1186/s40561-024-00347-0>

*De Vicente-Yagüe-Jara, M. I., López-Martínez, O., Navarro-Navarro, V., & Cuéllar-Santiago, F. (2023). Escritura, creatividad e inteligencia artificial: ChatGPT en el contexto universitario. *Comunicar*, 31(77), 47–57. <https://doi.org/10.3916/C77-2023-04>

*Du, X., Du, M., Zhou, Z., & Bai, Y. (2025). Facilitator or hindrance? The impact of AI on university students' higher order thinking skills in complex problem solving. *International Journal of Educational Technology in Higher Education*, 22 (1), 39. <https://doi.org/10.1186/s41239-025-00534-0>

*Fabus, J., Garbarova, M., Vartiak, L., & Mitkova, L. (2025). The use of AI tools in managerial education. *Management Theory and Studies for Rural Business and Infrastructure Development*, 47(2), 251–259. <https://doi.org/10.15544/mts.2025.19>

*Fisher, R., Gogan, T., Williams, J., Laferriere, R., Campbell, G., Gunasekara, A., & Nguyen, J. (2025). Can ChatGPT enhance business student creativity? Evidence from a randomised controlled trial. *Studies in Higher Education*. Advance online publication. <https://doi.org/10.1080/03075079.2025.2515512>

*Guo, K., Zhong, Y., Li, D., & Chu, S. K. W. (2023). Effects of chatbot-assisted in-class debates on students' argumentation skills and task motivation. *Computers & Education*, 203, 104862. <https://doi.org/10.1016/j.compedu.2023.104862>

*Hongxia, H., & Razali, A. B. (2025). Impact of ChatGPT on English academic writing ability and engagement of Chinese EFL undergraduates. *International Journal of Instruction*, 18(2), 323–346. <https://doi.org/10.29333/iji.2025.18218a>

Hopia, H., Latvala, E., Liimatainen, L. (2016). Reviewing the methodology of an integrative review. *Scand J Caring Sci*, 30(4), 662-669. <https://doi.org/10.1111/scs.12327>

- *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, 1919–1941. <https://doi.org/10.1111/bjet.13613>
- *Huang, H. (2024). Promoting students' creative and design thinking with generative AI-supported co-regulated learning: Evidence from digital game development projects in healthcare courses. *Educational Technology & Society*, 27(4), 487–502. [https://doi.org/10.30191/ETS.202410_27\(4\).TP04](https://doi.org/10.30191/ETS.202410_27(4).TP04)
- *Huang, J., & Mizumoto, A. (2024). The effects of generative AI usage in EFL classrooms on the L2 motivational self system. *Education and Information Technologies*. Advance online publication. <https://doi.org/10.1007/s10639-024-13071-6>
- *Karagol, E., Yıldırım Bilgen, D., & Korkmaz, C. B. (2025). The impact of AI applications on pre-service teachers' public speaking anxiety and academic speaking skill in the context of oral presentations: A mixed-methods study. *Teaching and Teacher Education*, 166, 105180. <https://doi.org/10.1016/j.tate.2025.105180>
- *Kim, S., So, H.-J., & Park, K. (2025). Supporting learner agency in collaborative writing with generative AI. *British Journal of Educational Technology*. Advance online publication. <https://doi.org/10.1111/bjet.70015>
- *Klyshbekova, M., & Abbott, P. (2024). ChatGPT and assessment in higher education: A magic wand or a disruptor? *Electronic Journal of e-Learning*, 22(2), 30–45. <https://doi.org/10.34190/ejel.21.5.3114>
- *Le, H., Shen, Y., Li, Z., Xia, M., Tang, L., Li, X., Jia, J., Wang, Q., Gašević, D., & Fan, Y. (2025). Breaking human dominance: Investigating learners' preferences for learning feedback from generative AI and

- human tutors. *British Journal of Educational Technology*, 56, 1758–1783. <https://doi.org/10.1111/bjet.13614>
- *Leahy, K., Ozer, E., & Cummins, E. P. (2025). AI-ENGAGE: A multicentre intervention to support teaching and learning engagement with generative artificial intelligence tools. *Education Sciences*, 15(7), 807. <https://doi.org/10.3390/educsci15070807>
- *Lee, Y.-F., Hwang, G.-J., & Chen, P.-Y. (2025). Technology-based interactive guidance to promote learning performance and self-regulation: A chatbot-assisted self-regulated learning approach. *Educational Technology Research and Development*, 73, 2279–2304. <https://doi.org/10.1007/s11423-025-10478-x>
- Li, K., Rollins, J. & Yan, E. (2018). Web of Science use in published research and review papers 1997-2017: a selective, dynamic, cross-domain, content-based analysis. *Scientometrics*, 115 (1), 1-20. <https://doi.org/10.1007/s11192-017-2622-5>
- *Lo, N., & Wong, A. (2025). Retracted: The impact of generative AI on essay revisions and student engagement. *Computers and Education Open*, 9, 100249. <https://doi.org/10.1016/j.caeo.2025.100249>
- *Makransky, G., Shiwalia, B. M., Herlau, T., & Blurton, S. (2025). Beyond the “wow” factor: Using generative AI for increasing generative sense-making. *Educational Psychology Review*, 37 (3). <https://doi.org/10.1007/s10648-025-10039-x>
- *Mendoza Moheno, J., Aubert Hernández Calzada, M., & Ortega-Mohedano, J. (2024). ChatGPT enters the classrooms: Student perceptions of the incorporation of artificial intelligence tools in the teaching of Economics and Business. *Educational Media International*, 61(4), 352–367. <https://doi.org/10.1080/09523987.2024.2436737>
- *Miao, C., Zhang, S., & Liu, K. (2025). Comparing AI-assisted and traditional tactical instruction: A crossover experimental study among male college students. *Acta Psychologica*, 259, 105431. <https://doi.org/10.1016/j.actpsy.2025.105431>

- *Naamati-Schneider, L., & Alt, D. (2024). Beyond digital literacy: The era of AI-powered assistants and evolving user skills. *Education and Information Technologies*, 29, 21263–21293. <https://doi.org/10.1007/s10639-024-12694-z>
- OECD (2026), *OECD Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education*, OECD Publishing, Paris, <https://doi.org/10.1787/062a7394-en>.
- *Oliva-Córdova, L. M., Alvarez-Icaza, I., & George-Reyes, C. E. (2025). Evaluation of generative AI use to foster critical thinking in higher education. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 20, 237–243. <https://doi.org/10.1109/RITA.2025.3597848>
- *Oliveira, L., Tavares, C., Strzelecki, A., & Silva, M. (2025). Prompting minds: Evaluating how students perceive generative AI's critical thinking dispositions. *Electronic Journal of e-Learning*, 23(2), 1–18. <https://doi.org/10.34190/ejel.23.2.3986>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T.C., Mulrow, C. D. et al. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372 (71). <https://doi.org/10.1136/bmj.n71>
- *Pellas, N. (2023). The effects of generative AI platforms on undergraduates' narrative intelligence and writing self-efficacy. *Education Sciences*, 13(11), 1155. <https://doi.org/10.3390/educsci13111155>
- *Polakova, P., & Ivenz, P. (2024). The impact of ChatGPT feedback on the development of EFL students' writing skills. *Cogent Education*, 11(1), 2410101. <https://doi.org/10.1080/2331186X.2024.2410101>
- Rong, H., & Chun, C. (2024). *Digital Education Council global AI student survey 2024*. Digital Education Council. <https://www.digitaleducationcouncil.com/post/digital-education-council-global-ai-student-survey-2024>
- *Saritepeci, M., & Yildiz Durak, H. (2024). Effectiveness of artificial intelligence integration in design-based learning on design thinking mindset, creative and

reflective thinking skills: An experimental study. *Education and Information Technologies*, 29, 25175–25209. <https://doi.org/10.1007/s10639-024-12829-2>

- *Sedlbauer, 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>
- *Salem, K. M., Saghier, E. G., Khreis, S. H. A., & Tan, C. C. (2025). ChatGPT-related risk patterns and students' creative thinking toward tourism statistics course: Pretest and posttest quasi-experimentation. *Journal of Hospitality & Tourism Education*. Advance online publication. <https://doi.org/10.1080/10963758.2025.2456638>
- *Tai, H.-Y., Lin, M.-F., & Chen, Y.-S. (2025). Incorporating ChatGPT into genre-based instruction for argumentative writing among EFL college students. *International Journal of Applied Linguistics*. Advance online publication. <https://doi.org/10.1111/ijal.12777>
- *Urban, M., Brom, C., Lukavský, J., Děchtěrenko, F., Hein, V., Svacha, F., Kmoníčková, P., & Urban, K. (2025). “ChatGPT can make mistakes. Check important info.” Epistemic beliefs and metacognitive accuracy in students' integration of ChatGPT content into academic writing. *British Journal of Educational Technology*, 56, 1897–1918. <https://doi.org/10.1111/bjet.13591>
- *Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education*. Advance online publication. <https://doi.org/10.1016/j.compedu.2024.105031>
- *Usher, M., & Amzalag, M. (2025). From prompt to polished: Exploring student–chatbot interactions for academic writing assistance. *Education Sciences*, 15(3), 329. <https://doi.org/10.3390/educsci15030329>
- *Wang, S., Tao, X., Ma, H., Li, F., & Wu, C. (2025). EEG assessment of artificial intelligence-generated content impact on student creative performance

- and neurophysiological states in product design. *Frontiers in Psychology*, 16, 1508383. <https://doi.org/10.3389/fpsyg.2025.1508383>
- *Wei, X., Wang, L., Lee, L.-K., & Liu, R. (2025). The effects of generative AI on collaborative problem-solving and team creativity performance in digital story creation: An experimental study. *International Journal of Educational Technology in Higher Education*, 22 (1), 23. <https://doi.org/10.1186/s41239-025-00526-0>
- *Wu, B., He, Y.-N., Song, Y., & Li, H.-H. (2025). Fostering critical thinking in higher education: An intelligent dialogue-based approach empowered by conversational AI. *Interactive Learning Environments*. Advance online publication. <https://doi.org/10.1080/10494820.2025.2538750>
- *Yusof, I. J. (2025). ChatGPT-assisted retrieval practice and exam scores: Does it work? *Journal of Information Technology Education: Research*, 24, 8. <https://doi.org/10.28945/5474>
- *Zhang, M. (2025). Optimizing academic engagement and mental health through AI: An experimental study on LLM integration in higher education. *Frontiers in Psychology*, 16, 1641212. <https://doi.org/10.3389/fpsyg.2025.1641212>
- *Zhang, Y., Lai, X., Yi, S., & Lu, Y. (2024). Does ChatGPT-based reading platform impact foreign language paper reading? Evidence from a quasi-experimental study on Chinese undergraduate students. *Education and Information Technologies*, 30, 9737–9754. <https://doi.org/10.1007/s10639-024-13190-0>
- *Zhang, Y., Tian, H., & Lu, J. (2025). The relationship between higher-order thinking and problem-solving skills development among pre-service teachers using generative AI: An analysis based on moderated mediation. *BMC Psychology*, 13, 1094. <https://doi.org/10.1186/s40359-025-03404-6>
- *Zhong, T., Zhu, G., Hou, C., Wang, Y., & Fan, X. (2024). The influences of ChatGPT on undergraduate students' demonstrated and perceived interdisciplinary

learning. *Education and Information Technologies*. Advance online
publication. <https://doi.org/10.1007/s10639-024-12787-9>