

Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.  
ISSN 2707-2207 / ISSN 2707-2215 (en línea), enero-febrero 2026,  
Volumen 10, Número 1.

[https://doi.org/10.37811/cl\\_rcm.v10i1](https://doi.org/10.37811/cl_rcm.v10i1)

## **INTELIGENCIA ARTIFICIAL EN EL AULA: PERCEPCIONES ÉTICAS Y ACTITUDINALES DE LOS ESTUDIANTES UNIVERSITARIOS**

**GENERATIVE ARTIFICIAL INTELLIGENCE (GENAI) IN  
PHYSICS EDUCATION: A SYSTEMATIC LITERATURE  
REVIEW**

**Angel Leonel Ortiz Herrera**

Universidad Michoacana de San Nicolás de Hidalgo

**Karina Mariela Figueroa Mora**

Universidad Michoacana de San Nicolás de Hidalgo

## Inteligencia artificial generativa (IAG) en la enseñanza de la Física: una revisión sistemática de literatura

Angel Leonel Ortiz Herrera<sup>1</sup>

[angel.ortiz@umich.mx](mailto:angel.ortiz@umich.mx)

<https://orcid.org/0009-0005-8848-9561>

Universidad Michoacana de San Nicolás de Hidalgo  
Ciudad Morelia, Michoacán  
México

Karina Mariela Figueroa Mora

[karina.figueroa@umich.mx](mailto:karina.figueroa@umich.mx)

<https://orcid.org/0000-0002-4680-5950>

Universidad Michoacana de San Nicolás de Hidalgo  
Ciudad Morelia, Michoacán México

### RESUMEN

La IAG ha comenzado a incorporarse en la enseñanza de la Física, pero la evidencia empírica disponible sigue siendo fragmentaria. Este estudio presenta una revisión sistemática de literatura, complementada con un análisis bibliométrico descriptivo, sobre aplicaciones de IAG en contextos formales de enseñanza de la Física entre 2018 y 2025. La búsqueda se realizó en Web of Science, Scopus, IEEE Xplore, ERIC y OpenAlex, siguiendo las directrices PRISMA 2020. Se extrajeron datos sobre tipo de herramienta, contenidos físicos abordados, nivel educativo, diseño metodológico y resultados de aprendizaje. Los hallazgos muestran un campo emergente, con crecimiento marcado a partir de 2023, concentrado en experiencias que utilizan modelos de lenguaje de gran escala como apoyo a la programación de simulaciones, la realización de prácticas de laboratorio y la visualización de fenómenos cuánticos en entornos inmersivos. Con este estudio, se lograron identificar lagunas en torno a evaluaciones controladas a gran escala, impactos sostenidos en el aprendizaje y marcos pedagógicos y éticos explícitos para integrar la IAG en el currículo de Física.

**Palabras clave:** inteligencia artificial generativa; enseñanza de la Física; modelos de lenguaje de gran escala; revisión sistemática de literatura; educación STEM.

---

<sup>1</sup> Autor principal

Correspondencia: [angel.ortiz@umich.mx](mailto:angel.ortiz@umich.mx)

# Generative Artificial Intelligence (GenAI) in Physics Education: A Systematic Literature Review

## ABSTRACT

Generative artificial intelligence is beginning to permeate physics education, yet empirical evidence on its classroom use remains scarce and fragmented. This paper reports a systematic literature review, complemented by a descriptive bibliometric analysis, on applications of GenAI in formal physics teaching between 2018 and 2025. A structured search was carried out in Web of Science, Scopus, IEEE Xplore, ERIC and OpenAlex, following PRISMA 2020 guidelines. After screening 44 initial records and applying predefined inclusion and exclusion criteria, we obtained a thematic corpus of 18 documents and a final set of six empirical studies examined in depth. Data were extracted on the type of GenAI tool, physics content, educational level, methodological design and reported learning outcomes. The results portray an emerging field with marked growth from 2023 onwards, dominated by small-scale implementations of large language models used to support programming-based simulations, introductory laboratory investigations and the visualization of quantum phenomena in immersive environments. Across studies, GenAI tends to foster localized gains in conceptual understanding, step-by-step reasoning and student motivation; however, these findings rest on exploratory designs, small convenience samples and a lack of longitudinal evidence. The review identifies persistent gaps regarding rigorous experimental and quasi-experimental evaluations, long-term impacts on physics learning, teachers' pedagogical appropriation of GenAI, and explicit ethical frameworks for its responsible integration into physics curricula.

**Keywords:** generative artificial intelligence; physics education; large language models; systematic literature review; STEM education.

*Artículo recibido 20 diciembre 2025  
Aceptado para publicación: 22 enero 2026*



## INTRODUCCIÓN

La enseñanza de la Física posee un impacto formativo amplio en la educación científica, al contribuir al desarrollo del razonamiento analítico, la comprensión de modelos matemáticos y la interpretación de fenómenos naturales fundamentales. No obstante, su implementación didáctica implica desafíos significativos, derivados tanto del alto nivel de abstracción conceptual de muchos contenidos como de la complejidad asociada a la preparación de materiales educativos que integren modelación, experimentación y múltiples representaciones. Diversos estudios han señalado que el diseño de experiencias de aprendizaje efectivas en Física exige una cuidadosa coordinación entre objetivos, actividades y recursos, lo que incrementa la carga pedagógica para el profesorado y puede limitar la adopción de enfoques innovadores si no se cuenta con apoyos adecuados (Holmes et al., 2017; Meltzer & Thornton, 2012; Wieman, 2014).

La IAG ha adquirido un papel cada vez más relevante en la educación gracias a su capacidad para producir textos, explicaciones, simulaciones y soluciones contextualizadas en tiempo real (Lo, 2023; Tlili et al., 2023). Este crecimiento está directamente asociado al desarrollo reciente de los modelos de lenguaje de gran escala (LLM), cuya arquitectura ha ampliado el rango de tareas académicas susceptibles de automatización o apoyo a procesos cognitivos complejos (Kasneci et al., 2023; Latif et al., 2024; Liang et al., 2023). Tales avances prolongan una trayectoria previa de investigación sobre inteligencia artificial (IA) en educación, centrada en tutorías inteligentes, aprendizaje adaptativo y sistemas de recomendación (Luckin & Holmes, 2016; VanLehn, 2011).

Diversos organismos internacionales han subrayado que la rápida incorporación de tecnologías basadas en IA exige nuevas competencias, además de marcos éticos y pedagógicos que garanticen un uso responsable en los sistemas educativos (Lo, 2023; UNESCO, 2021, 2023; Zawacki-Richter et al., 2019). En este proceso de transformación, la enseñanza de la Física constituye un campo particularmente significativo debido a su elevado nivel de abstracción, su dependencia de la modelación matemática y la necesidad de múltiples formas de representación para abordar fenómenos que suelen ser contraintuitivos (Radianti et al., 2020). Investigaciones previas han mostrado que tecnologías avanzadas pueden favorecer la comprensión de modelos físicos, mejorar la visualización de fenómenos complejos y fortalecer la construcción de razonamientos científicos (Holmes et al., 2017; Liang et al., 2023;



Meltzer & Thornton, 2012; Wieman, 2014); en concordancia, estudios recientes han documentado el uso directo de modelos generativos en actividades de enseñanza de la Física, como es el caso del empleo de ChatGPT para apoyar explicaciones y resolución de problemas conceptuales (Avila et al., 2024).

Aunque la literatura reconoce el potencial de la IAG para apoyar la educación científica, su aplicación específica en la enseñanza de la Física permanece limitada y metodológicamente heterogénea (Bitzenbauer, 2023; Liang et al., 2023). En escenarios emergentes como este, los análisis bibliométricos han permitido identificar tendencias y configuraciones temáticas que contribuyen a la consolidación del campo (Archambault et al., 2009; Donthu et al., 2021; Zupic & Čater, 2015).

De manera complementaria, revisiones recientes en educación científica han identificado oportunidades y desafíos específicos en la integración de IA en la enseñanza de las ciencias, destacando la necesidad de marcos analíticos más consistentes (Lee et al., 2025).

Paralelamente, revisiones recientes destacan la necesidad de sintetizar la evidencia disponible para orientar decisiones basadas en datos sobre el uso de IAG en contextos educativos (Lo, 2023; Zhai, 2024), así como la importancia de disponer de panoramas sistemáticos más amplios que integren el papel de la IA en áreas STEM (Xu & Ouyang, 2022).

En este contexto, la presente revisión sistemática examina la evolución reciente de la literatura científica sobre IAG aplicada a la enseñanza de la Física. Su propósito es ofrecer una síntesis fundamentada que permita comprender el estado del conocimiento, identificar vacíos y orientar futuras líneas de investigación para una integración pedagógica pertinente y responsable de estas tecnologías.

En coherencia con este objetivo, el presente estudio se estructura como una revisión sistemática de literatura, complementada por un análisis bibliométrico descriptivo que permite caracterizar la evolución reciente del campo y contextualizar los hallazgos empíricos.

## **METODOLOGÍA**

Para asegurar un proceso riguroso y transparente en la identificación, selección y síntesis de la evidencia sobre el uso de IAG en la enseñanza de la Física, se desarrolló una revisión sistemática siguiendo las directrices del protocolo PRISMA 2020 (Page et al., 2021). Este enfoque permitió estructurar cada fase del proceso bajo criterios claros y reproducibles, y resultó especialmente pertinente para un campo

emergente como la IAG aplicada a la educación en Física, donde las aproximaciones conceptuales y metodológicas suelen ser variadas (Gough et al., 2017).

Esta revisión adopta un enfoque metodológico dual: por un lado, sigue el protocolo PRISMA 2020 para la identificación, selección y síntesis sistemática de estudios empíricos; por otro, incorpora un análisis bibliométrico descriptivo derivado de los registros iniciales para examinar tendencias de producción, distribución geográfica y concentración temática.

### **Preguntas de Investigación**

La revisión se organizó a partir de cuatro preguntas que orientan el alcance del estudio y guían el análisis de la evidencia:

1. ¿Qué tipos de aplicaciones de IAG se han explorado en la enseñanza de la Física durante el periodo 2018–2025?
2. ¿Qué enfoques metodológicos se han utilizado para evaluar estas aplicaciones en distintos contextos educativos?
3. ¿Qué efectos, beneficios y limitaciones reportan los estudios empíricos que integran herramientas generativas en actividades de laboratorio, visualización o razonamiento conceptual?
4. ¿Qué vacíos persisten en la literatura y qué líneas de investigación requieren mayor desarrollo?

Estas preguntas permiten articular de manera coherente los objetivos del estudio, los criterios de selección del corpus y la síntesis desarrollada en las secciones posteriores, en consonancia con las recomendaciones metodológicas para revisiones sistemáticas en educación y tecnología (Xiao & Watson, 2019).

### **Diseño de la revisión**

El estudio se estructuró siguiendo las directrices del protocolo PRISMA 2020, lo que permitió organizar de manera transparente cada etapa del proceso. Este procedimiento incluyó la búsqueda sistemática en cinco bases de datos, la depuración y normalización de registros, el cribado temático, la revisión a texto completo, la evaluación de calidad y la síntesis tanto cuantitativa como cualitativa.



El proceso de cribado (títulos, resúmenes y lectura a texto completo) fue realizado por un solo revisor, dado el tamaño reducido del corpus y la claridad de los criterios de inclusión y exclusión. Esta decisión es metodológicamente válida en revisiones con alcances temáticos acotados; no obstante, se aplicaron procedimientos de verificación interna para asegurar la consistencia y coherencia en las decisiones de selección.

Este diseño resultó especialmente pertinente para un campo emergente como la IAG aplicada a la educación en Física, donde la literatura suele ser amplia y heterogénea. Su implementación permitió integrar diferentes tipos de evidencia y asegurar un análisis ordenado y reproducible, en consonancia con las recomendaciones recientes para revisiones sistemáticas en educación y tecnología (Booth et al., 2016; Page et al., 2021).

En la fase de lectura a texto completo se excluyeron doce estudios que, si bien guardaban relación general con la IA en educación, no cumplían simultáneamente los criterios temáticos y metodológicos definidos para esta revisión. Entre ellos se encontraban trabajos centrados en contextos de educación STEM o informática sin un foco específico en contenidos de Física, estudios teóricos o de revisión sin datos empíricos de aula, investigaciones orientadas principalmente a la aceptación o formación del profesorado sin reporte de resultados de aprendizaje del alumnado, propuestas de plataformas gamificadas que no incorporaban de forma explícita modelos de IAG y un preprint no sometido todavía a revisión por pares. La exclusión de estos trabajos permitió delimitar un corpus final de seis estudios empíricos estrictamente alineados con las preguntas de investigación, centrados en experiencias de uso de IAG en la enseñanza de contenidos de Física.

### **Fuentes de información**

La búsqueda bibliográfica se realizó entre julio y noviembre de 2025 en cinco bases de datos reconocidas por su cobertura amplia y complementaria en educación, ciencia y tecnología: Web of Science – Core Collection, Scopus, IEEE Xplore, ERIC y OpenAlex. Esta combinación permitió recuperar literatura reciente y reducir posibles sesgos de indexación. Cada plataforma aportó perspectivas distintas, especialmente en áreas vinculadas con educación STEM, tecnología educativa e investigación en ciencias físicas.



El uso de múltiples bases de datos coincide con recomendaciones metodológicas para estudios interdisciplinarios (Harzing & Alakangas, 2016; Martín-Martín et al., 2018), y con revisiones recientes sobre IAG en educación que subrayan la importancia de integrar fuentes diversas para obtener una visión más completa del campo (Wang et al., 2024).

### **Estrategia de búsqueda**

La estrategia de búsqueda se organizó a partir de tres bloques conceptuales que ayudaron a delimitar los términos relacionados con la IAG, la enseñanza de la Física y el uso de simulaciones o actividades interactivas. A partir de esta estructura se formularon las ecuaciones específicas para cada plataforma, procurando mantener una terminología equivalente entre motores de búsqueda.

A modo de ejemplo, la ecuación utilizada en Scopus fue la siguiente:

("generative AI" OR "generative artificial intelligence" OR GenAI OR "large language model" OR ChatGPT) AND ("physics education" OR "teaching physics" OR "learning physics" OR "physics classroom" OR "physics simulation") AND ("game-based learning" OR "educational simulation" OR "virtual laboratory" OR "interactive learning")

Esta organización permitió recuperar estudios que integraran modelos generativos en procesos de enseñanza de la Física, tanto en contextos experimentales como conceptuales. Los grupos conceptuales y los términos empleados se presentan en la Tabla 1.

Dado que cada base de datos utiliza operadores y reglas de búsqueda particulares, fue necesario ajustar manualmente la sintaxis para mantener la correspondencia entre plataformas. Los mismos descriptores principales se conservaron en todos los casos, modificando únicamente elementos técnicos como comillas, conectores booleanos o símbolos propios de cada sistema. Este procedimiento aseguró una recuperación de documentos uniforme y comparable.

**Tabla 1.** Términos de búsqueda empleados en las bases de datos

<b>Conceptual category</b>	<b>Terms used</b>
GenAI and generative models	“generative AI”, “generative artificial intelligence”, GenAI, “large language model”, ChatGPT
Physics education	“physics education”, “teaching physics”, “learning physics”, “physics classroom”, “physics simulation”
Simulation and interactive activities	“game-based learning”, “educational simulation”, “virtual laboratory”, “interactive learning”

---

### **Criterios de inclusión y exclusión**

Los criterios de selección se definieron para garantizar la pertinencia temática del corpus y mantener coherencia metodológica a lo largo del proceso. Se tomaron como referencia lineamientos para revisiones sistemáticas en educación y tecnología, que recomiendan establecer procedimientos de filtrado claros y reproducibles ante la diversidad documental existente (Snyder, 2019).

Para asegurar la comparabilidad entre los estudios analizados, se adoptó un rango temporal de 2018 a 2025, periodo en el que comienzan a difundirse los modelos de lenguaje y surgen sus primeras aplicaciones educativas. El límite superior corresponde al cierre de la búsqueda sistemática. Asimismo, se incluyeron únicamente documentos redactados en inglés, dado que este idioma concentra buena parte de la producción científica internacional y facilita la integración de los hallazgos.

En cuanto a los criterios de exclusión, se descartó la literatura gris —como tesis, preprints e informes técnicos— por no contar con procesos formales de revisión por pares y presentar una variabilidad metodológica considerable. También se excluyeron estudios centrados únicamente en el desarrollo técnico de modelos de IA sin una aplicación educativa explícita, ya que no respondían a los objetivos de esta revisión.

Finalmente, no se consideraron repositorios abiertos como Google Scholar o arXiv debido a la diversidad y naturaleza preliminar de muchos de sus contenidos, que no siempre han sido sometidos a evaluación editorial. Para garantizar la consistencia del corpus, la revisión se centró en bases de datos con mecanismos formales de selección e indexación académica, lo que permitió trabajar con registros más estables y comparables.

Estos criterios permitieron delimitar un conjunto documental coherente y metodológicamente comparable para el análisis posterior.

### **1.1. Proceso de selección (PRISMA 2020)**

El proceso de identificación y selección de los estudios siguió las etapas establecidas por el protocolo PRISMA 2020, lo que permitió mantener una trazabilidad clara durante el cribado. La búsqueda inicial recuperó 44 registros procedentes de las cinco bases de datos seleccionadas. Tras eliminar los duplicados con base en coincidencias de DOI, se conformó un conjunto de 37 documentos.

Posteriormente, se realizó un cribado temático a partir de la revisión de títulos, resúmenes y palabras clave, con lo que se identificaron 18 estudios vinculados con la enseñanza de la Física y el uso de IAG. La lectura a texto completo permitió delimitar un conjunto final de seis estudios empíricos, que constituyen el corpus analizado en profundidad.

Este flujo de selección aseguró que los estudios incluidos guardaran coherencia con los objetivos y respondieran directamente a las preguntas de investigación planteadas.

### **Procesamiento, normalización y extracción**

Una vez definido el corpus final de estudios, se llevó a cabo el procesamiento y preparación de la información con el propósito de garantizar consistencia, comparabilidad y solidez analítica. Este trabajo se centró en depurar y armonizar los registros seleccionados, revisando de manera sistemática la uniformidad de los metadatos asociados a cada estudio, tales como autorías, afiliaciones institucionales, países de procedencia, tipos de estudio y niveles educativos, con el fin de conformar una base de datos coherente para el análisis posterior.

De manera paralela, se realizó una revisión detallada del contenido de los estudios incluidos para confirmar su pertinencia temática, el tipo de uso de IAG reportado y su vínculo explícito con la enseñanza o el aprendizaje de la Física. Como parte de este proceso, se elaboró un tesoro controlado a

partir de los títulos, resúmenes y palabras clave del corpus, lo que permitió unificar variantes terminológicas, agrupar sinónimos y organizar los conceptos en categorías temáticas consistentes, fortaleciendo así la coherencia semántica del análisis.

Con la información ya normalizada, se procedió a la extracción sistemática de los datos y al análisis descriptivo y temático mediante RStudio y el paquete Bibliometrix, complementados con procedimientos desarrollados específicamente para este estudio. Este enfoque permitió la obtención de indicadores de producción científica, la construcción de matrices de coocurrencia y la generación de visualizaciones temáticas, en consonancia con prácticas consolidadas en el análisis bibliométrico aplicado al ámbito educativo (Aria & Cuccurullo, 2017).

Adicionalmente, los seis estudios incluidos fueron sometidos a una evaluación de calidad metodológica utilizando la matriz presentada en la Tabla 2, con el objetivo de contextualizar la solidez de la evidencia disponible. Esta evaluación no se empleó como criterio de exclusión, sino como un recurso interpretativo que permitió identificar fortalezas y limitaciones en el conjunto de estudios. En términos generales, se observó mayor consistencia en la claridad de los objetivos y en la descripción de las intervenciones, mientras que las principales debilidades se relacionaron con tamaños muestrales reducidos, ausencia de grupos de comparación y escasa información sobre procedimientos de validez y confiabilidad. Estos aspectos fueron considerados al interpretar los resultados y al delimitar el alcance de las conclusiones.

**Tabla 2.** Criterios de calidad metodológica (Gast & Ledford, 2024)

Criterio	Descripción
Propósito	Claridad del objetivo y coherencia con el diseño.
Diseño	Alineación entre metodología y tipo de intervención.
Contexto	Descripción del entorno educativo y participantes.
Datos	Calidad y suficiencia de la recolección de datos.

Validez	Uso de instrumentos válidos y confiables.
Rigor analítico	Procedimientos claros y replicables.
Limitaciones	Reconocimiento de restricciones del estudio.

### **Síntesis y análisis de la información**

La síntesis de la evidencia se desarrolló mediante un enfoque integrado que permitió relacionar los hallazgos cuantitativos del mapeo bibliométrico con el análisis cualitativo de los seis estudios empíricos del corpus final. Este enfoque mixto resulta adecuado en campos emergentes, donde la literatura combina reportes exploratorios, intervenciones pedagógicas y desarrollos tecnológicos diversos.

En un primer nivel, el análisis bibliométrico permitió caracterizar la producción científica reciente, describir patrones de publicación por año, distribución por fuentes y países, y reconocer las tendencias temáticas predominantes relacionadas con el uso de IAG en la enseñanza de la Física. Estos indicadores ofrecieron un panorama general del desarrollo del campo y facilitaron la identificación de áreas emergentes.

En un segundo nivel, el análisis empírico se centró en examinar las aplicaciones didácticas reportadas, los tipos de herramientas generativas empleadas y los efectos documentados en el aprendizaje, especialmente en actividades de laboratorio, visualización y razonamiento conceptual. Esta doble aproximación permitió combinar una visión general del campo con interpretaciones más contextualizadas, en línea con las orientaciones propuestas para la síntesis interpretativa en investigación educativa (Thomas & Harden, 2008).

### **RESULTADOS**

Los resultados de la revisión sistemática permiten describir la evolución reciente de la investigación sobre IAG en la enseñanza de la Física y caracterizar las tendencias metodológicas y temáticas presentes en los estudios incluidos. Para ello, los hallazgos se organizan en dos niveles: primero, una caracterización general del campo basada en indicadores bibliométricos que muestran patrones de producción, distribución geográfica y concentración por fuentes; y segundo, un análisis detallado del corpus empírico, en el que se examinan las aplicaciones reportadas, los enfoques pedagógicos y tecnológicos empleados y los efectos observados en el aprendizaje. Esta estructura facilita combinar una

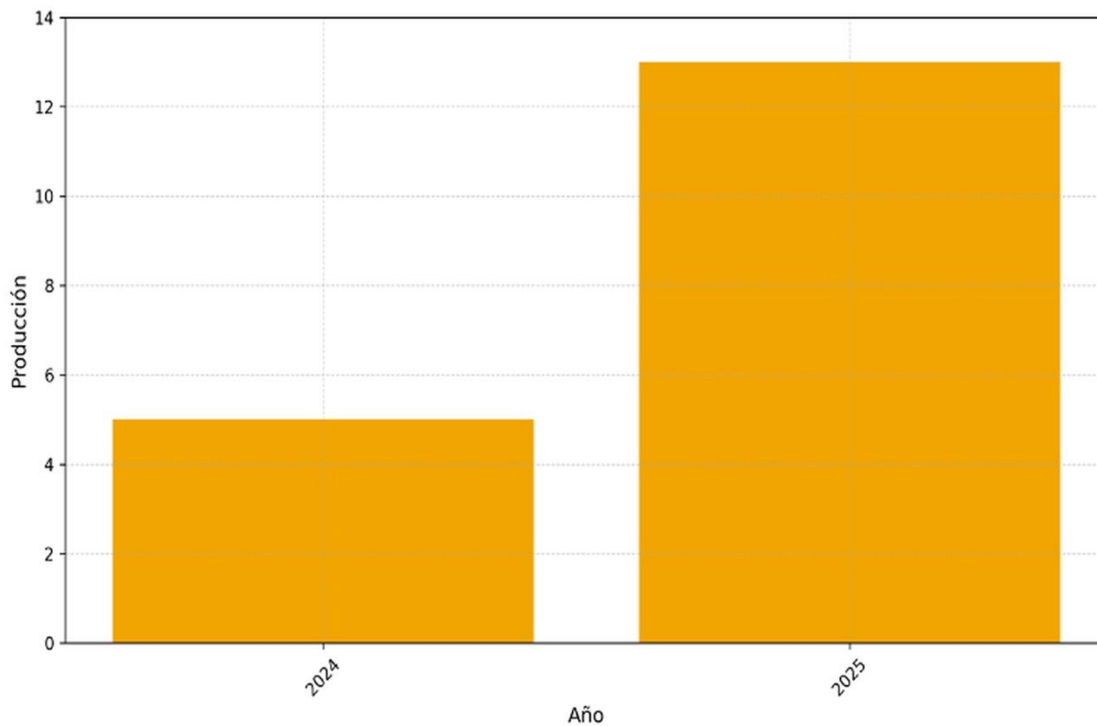


visión panorámica del desarrollo del área con una síntesis profunda de las contribuciones empíricas más relevantes.

### Caracterización general del campo

La producción científica relacionada con la aplicación de IAG en la enseñanza de la Física muestra un aumento notable en los dos años más recientes, como se observa en la Figura 1. Durante 2024 se identificaron las primeras contribuciones relevantes, mientras que en 2025 se registró un incremento sustancial en el número de estudios pertinentes. Este comportamiento confirma el carácter emergente del campo y refleja la rápida incorporación de modelos generativos en entornos educativos durante este periodo.

**Figura 1.** Producción anual de publicaciones sobre IAG en la enseñanza de la Física (n=18).



**Nota.** La figura muestra un aumento marcado en las publicaciones entre 2024 y 2025.

En cuanto a las fuentes de publicación, la Tabla 3 reúne las revistas y congresos en los que se concentra la mayor parte de los estudios recuperados. De los 18 trabajos que integran el corpus, 15 proceden de los canales incluidos en dicha tabla, mientras que los tres restantes provienen de fuentes únicas que no se incorporaron en la síntesis por su carácter aislado. Esta distribución evidencia que la producción se dispersa entre espacios editoriales vinculados con la tecnología educativa, la informática aplicada y la didáctica de las ciencias. La ausencia de un canal claramente dominante es coherente con un campo aún en consolidación, donde las contribuciones no convergen en una plataforma editorial específica.

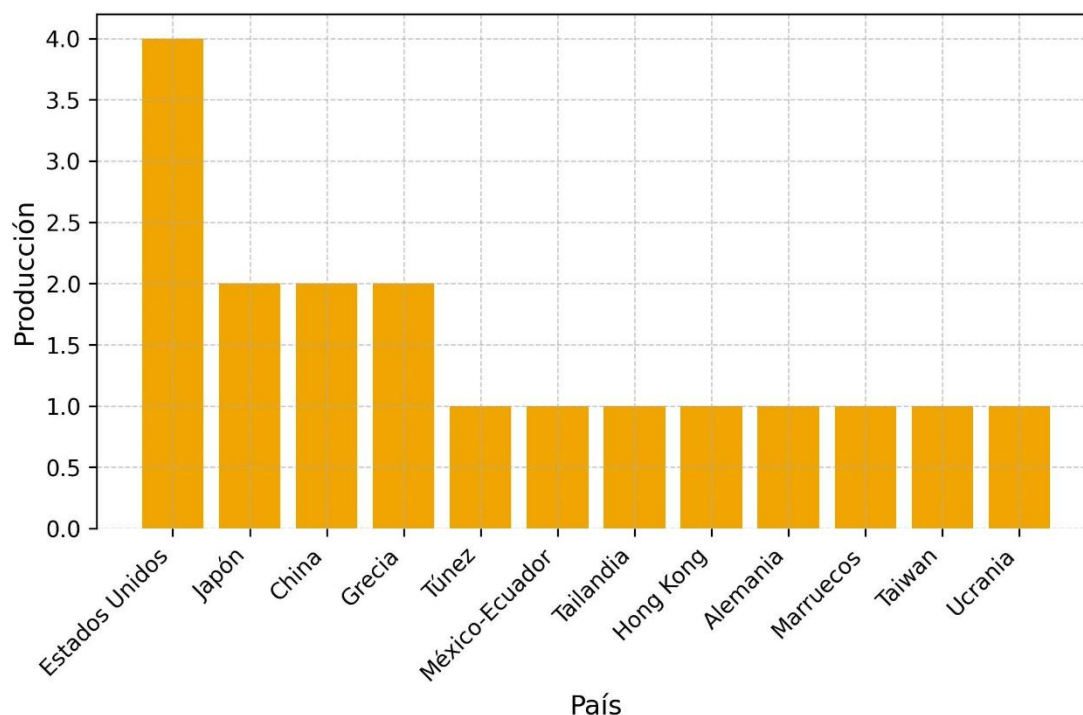
**Tabla 3.** Fuentes de publicación representadas en el corpus temático

Revista	Documentos
Physical Review Physics Education Research (PRPER)	4
Journal Of Baltic Science Education (JBSE)	2
ITICSE 2024 – Conference Proceedings	2
Computer Applications In Engineering Education	1
WIPSCE 2023 – Primary & Secondary Computing Education	1
European Journal Of Education Studies	1
Grani (ΓΡΑΗΙ)	1
Jurnal Penelitian Dan Pembelajaran IPA	1
Eurographics 2025 – Education Papers	1
International Journal Of Information And Education Technology	1

La distribución geográfica de los estudios, presentada en la Figura 2, muestra aportes procedentes de distintas regiones, con una concentración destacada en Asia —especialmente en China, Japón y Taiwán— y, en menor medida, en Norteamérica y Europa. También se identifican contribuciones puntuales desde África y otros contextos emergentes, lo que evidencia que el interés por integrar IAG en la enseñanza de la Física se desarrolla en entornos educativos heterogéneos. Esta dispersión geográfica contribuye a la diversidad de enfoques metodológicos observados.



**Figura 2.** Distribución geográfica de los estudios sobre IAG aplicados a la enseñanza de la Física (n = 18).



**Nota.** Se incluyen países con contribuciones procedentes del corpus temático de 18 artículos.

### **Análisis temático a partir del tesoro**

El tesoro controlado elaborado durante la fase de normalización permitió identificar patrones conceptuales recurrentes en los estudios y organizar los términos en tres ejes temáticos principales que estructuran la diversidad de aplicaciones reportadas.

#### **a) Integración didáctica de la IAG**

Reúne estudios que emplean modelos generativos como apoyo en distintas tareas educativas, ya sea para ofrecer retroalimentación automatizada, orientar la realización de experimentos, proporcionar tutorías personalizadas o generar explicaciones paso a paso que facilitan la comprensión de conceptos físicos.

#### **b) Visualización, simulación y entornos interactivos**

Incluye investigaciones que combinan herramientas generativas con recursos como la realidad aumentada, los simuladores de laboratorio, los juegos serios o distintos ambientes interactivos, con el propósito de enriquecer la visualización de fenómenos y la experimentación virtual en contextos formativos.

### c) Experiencia y percepción estudiantil

Abarca estudios que analizan cómo el alumnado percibe el uso educativo de la IAG, considerando dimensiones como la aceptación tecnológica, la utilidad percibida y las actitudes hacia su incorporación en actividades de aprendizaje de la Física.

Estos tres ejes permiten comprender el campo desde una perspectiva integrada que articula usos pedagógicos, recursos tecnológicos y la experiencia estudiantil, ofreciendo una base común para interpretar la variedad de enfoques presentes en el corpus empírico.

### **Síntesis del corpus empírico**

El análisis detallado del corpus empírico permitió identificar un conjunto de patrones metodológicos, tecnológicos y educativos que caracterizan la investigación reciente sobre IAG en la enseñanza de la Física. La Tabla 4 sintetiza los seis estudios incluidos, diferenciando entre aquellos que conforman el núcleo analítico —por integrar de manera directa herramientas de IAG en actividades de aprendizaje disciplinar— y los que forman parte del corpus contextual, cuyo aporte se orienta a la comprensión de percepciones, usos generales de IA o entornos STEM relacionados.

En los estudios del núcleo analítico (A1–A3), la IAG se incorpora como un recurso de apoyo para la resolución de problemas, el trabajo de laboratorio y la visualización de fenómenos complejos. Estas propuestas coinciden en considerar la IA como un complemento a la instrucción docente y no como un sustituto de la interacción pedagógica, enfatizando la importancia de la mediación experta para evitar interpretaciones imprecisas o explicaciones incompletas.

Desde el punto de vista metodológico, predomina el uso de enfoques exploratorios y diseños descriptivos aplicados en pequeña escala, tanto en demostraciones controladas como en intervenciones de corto alcance. Solo uno de los estudios adopta un diseño experimental formal, mientras que los demás se centran en evaluar la viabilidad pedagógica, la experiencia de uso o las percepciones del estudiantado. Este tipo de aproximaciones es habitual en etapas tempranas de incorporación tecnológica, cuando el objetivo principal es identificar potenciales beneficios antes de avanzar hacia evaluaciones comparativas más amplias.

Los trabajos del corpus contextual (A4–A6) aportan una perspectiva complementaria al documentar tanto las percepciones estudiantiles sobre herramientas generativas como su integración en entornos



inmersivos o prototipos educativos generales. Aunque su relación con la Física es más indirecta, estas investigaciones contribuyen a comprender la disposición del estudiantado, la aceptación de la tecnología y los desafíos asociados a su incorporación en actividades STEM.

En conjunto, los hallazgos muestran un campo en consolidación, caracterizado por experiencias innovadoras pero aún fragmentadas. Las evidencias sugieren que la IAG puede favorecer procesos de reflexión, construcción de explicaciones y apoyo cognitivo en tareas disciplinares, aunque persisten limitaciones relacionadas con la precisión de las respuestas, la necesidad de supervisión y el reducido alcance de los diseños empíricos. Esto refuerza la pertinencia de avanzar hacia estudios más sistemáticos y comparativos que permitan evaluar con mayor profundidad el impacto real de estas tecnologías en el aprendizaje de la Física.



**Tabla 4.** Características del corpus empírico final

ID	Título del Artículo (autor/año)	País	Tipo de Estudio	IAG utilizada	Relación con Física	Tipo de corpus
A1	<i>Application of Generative AI-Based Programming to Physics Learning</i> (Nishiwaki, 2025)	Japón	Exploratorio	Modelo generativo para apoyo en programación	Física general mediante resolución computacional	Núcleo
A2	<i>PhysicsAssistant: An LLM-Powered Interactive Learning Robot for Physics Lab Investigations</i> (Latif et al., 2024)	EE.UU.	Empírico Exploratorio	LLM integrado en robot asistente	Laboratorios de Física (mediciones, voltaje, corriente)	Núcleo
A3	<i>AI Support Meets AR Visualization for Alice and Bob: Personalized Learning in a Quantum Cryptography Lab</i> (Coban et al., 2025)	Alemania	Experimental	LLM + Realidad Aumentada	Física cuántica y criptografía	Núcleo
A4	<i>Exploring Collaborative Immersive VR</i>	China	Empírico Exploratorio	No emplea LLM directamente	Contenidos de Física mediante	Contextual



*Serious Games for  
Enhancing Learning  
Motivation in  
Physics  
Education* (Sun  
et al., 2024)

juego serio en  
VR

A 5	<i>Mapping Moroccan Students’ Perceptions of AI: A Case Study on Generative AI (ChatGPT)</i> (Tafhi et al., 2025)	Marruecos	Estudio de percepción( Cuantitativo descriptivo– exploratorio)	ChatGPT	STEM; vínculo indirecto con Física	Contextua l
A 6	<i>Development of an AI-Supported Chatbot as an Interactive Learning Platform in STEM Education</i> (Fekets, 2025)	Taiwan	Diseño y validación	ChatGPT / LLM	STEM Física incluida en el currículo	Contextua l



## **Análisis de los estudios empíricos**

Antes de examinar cada investigación de forma individual, resulta necesario situar los alcances del conjunto empírico. Los seis estudios incluidos abordan la integración de herramientas de IAG en actividades vinculadas con la enseñanza y el aprendizaje de la Física, aunque difieren en sus objetivos, metodologías y escenarios de aplicación. En esta sección se analizan sus características principales, los resultados que reportan y las aportaciones que ofrecen al campo, con el propósito de identificar patrones comunes y matices distintivos que permitan comprender mejor el papel de la IAG en contextos educativos de Física.

### **A1 – Application of Generative AI-Based Programming to Physics Learning (Nishiwaki, 2025)**

El trabajo de Nishiwaki examina el potencial de los modelos generativos como apoyo en la programación utilizada para estudiar fenómenos físicos. El artículo se centra en mostrar cómo estas herramientas pueden acompañar el proceso de construcción y depuración de código, facilitando que estudiantes y docentes desarrollen simulaciones accesibles sin requerir una formación avanzada en programación. El estudio parte de la premisa de que las barreras técnicas suelen limitar el acceso a actividades computacionales en cursos de Física, por lo que una herramienta que genere código funcional o ayude a corregirlo puede hacer estas prácticas más inclusivas.

El enfoque adoptado por el autor es fundamentalmente exploratorio y se orienta a evaluar la viabilidad pedagógica y técnica del uso de modelos generativos para apoyar la resolución computacional de problemas. Más que presentar un experimento formal con una muestra definida, el artículo ofrece una demostración detallada de cómo la IAG puede producir fragmentos de código en diferentes lenguajes, señalar errores, sugerir mejoras y permitir que el usuario revise paso a paso la lógica de la simulación. A lo largo del texto se ilustran ejemplos de interacción con el modelo, destacando tanto sus aciertos como las situaciones en las que las respuestas requieren ajustes o validación manual.

Los resultados muestran que la IAG puede reducir la complejidad inicial asociada a la programación y ayudar a que el usuario avance con mayor fluidez en la construcción de modelos numéricos. El sistema funciona como un apoyo para identificar inconsistencias y proponer alternativas, pero no sustituye la necesidad de comprender el fenómeno físico ni de verificar el comportamiento del código. El artículo



subraya que la revisión crítica sigue siendo imprescindible, pues la herramienta puede generar soluciones incompletas o que requieren afinación para producir resultados correctos.

Entre las limitaciones reconocidas figura la dependencia del usuario para validar el código y la necesidad de contar con criterios disciplinares para evaluar si la simulación refleja adecuadamente el fenómeno estudiado. A pesar de ello, el autor argumenta que los modelos generativos tienen un valor significativo como recurso educativo, al permitir que las actividades de programación en Física sean más accesibles y fomenten la exploración autónoma. La contribución central del estudio radica en mostrar la viabilidad de integrar IAG como un apoyo flexible en la enseñanza de la Física computacional, enfatizando que su utilidad depende de una interacción crítica y reflexiva por parte del usuario.

#### **A2 – PhysicsAssistant: An LLM-Powered Interactive Learning Robot for Physics Lab Investigations** (Latif et al., 2024)

El estudio presenta *PhysicsAssistant*, un robot interactivo diseñado para apoyar a estudiantes durante actividades de laboratorio en cursos introductorios de Física. El sistema integra un modelo de lenguaje de gran escala con sensores y módulos de respuesta que permiten mantener un diálogo en tiempo real con el estudiante mientras este manipula equipamiento o registra mediciones. El propósito principal es explorar cómo una herramienta de este tipo puede acompañar la realización de prácticas básicas como mediciones de voltaje, corriente o resistencia ofreciendo orientación inmediata y comentarios formativos durante el procedimiento.

A diferencia de investigaciones que se limitan a demostrar capacidades técnicas, este estudio adopta una aproximación empírica centrada en la experiencia de uso. Los autores describen la implementación del robot en un entorno educativo y recogen impresiones del estudiantado sobre la utilidad del sistema para aclarar dudas, verificar procedimientos y reforzar conceptos que suelen generar confusión en el laboratorio. Aunque el estudio no reporta un diseño experimental comparativo, sí proporciona evidencia directa sobre la interacción entre los usuarios y el robot, lo que permite identificar patrones de uso, aciertos y limitaciones.

De acuerdo con los resultados presentados, el robot facilita que los estudiantes avancen con mayor seguridad en actividades que requieren precisión procedimental. La posibilidad de recibir retroalimentación instantánea ayuda a corregir errores simples antes de que afecten las mediciones y



favorece un ambiente de trabajo más autónomo. Los participantes valoran especialmente la capacidad del sistema para explicar pasos concretos de la práctica, interpretar instrucciones y resolver dudas sin necesidad de interrumpir al docente.

El estudio también identifica limitaciones importantes. En ocasiones, el modelo de lenguaje puede ofrecer respuestas breves o demasiado generales, lo que obliga al estudiante a reformular la consulta para obtener una explicación más precisa. Además, la utilidad del robot depende de la calidad de la pregunta formulada y de la capacidad del estudiante para interpretar la respuesta en el contexto del procedimiento experimental. Los autores reconocen que todavía es necesario mejorar la adecuación de las respuestas y la integración del robot con instrumentos de laboratorio más avanzados.

En conjunto, *PhysicsAssistant* muestra el potencial de los modelos generativos para enriquecer la experiencia en laboratorios de Física, especialmente en etapas tempranas de formación. La contribución principal del estudio radica en demostrar que los LLM pueden funcionar como un acompañante flexible durante la ejecución de prácticas experimentales, ofreciendo apoyo inmediato sin desplazar la supervisión docente. El trabajo abre una línea de exploración relevante sobre cómo combinar interacción conversacional, asistencia procedural y manipulación de equipamiento en entornos educativos reales.

### **A3 – AI Support Meets AR Visualization for Alice and Bob: Personalized Learning Based on Individual ChatGPT Feedback in an AR Quantum Cryptography Experiment for Physics Lab Courses (Coban et al., 2025)**

El estudio de Coban y colaboradores se sitúa en un laboratorio de Física cuántica centrado en criptografía cuántica, donde se combina un experimento de “Alice y Bob” en realidad aumentada con retroalimentación personalizada generada por un modelo de lenguaje de gran escala. El objetivo central es analizar cómo este tipo de feedback, integrado en un entorno inmersivo, influye tanto en el aprendizaje conceptual como en los procesos atencionales de los estudiantes mientras trabajan en un experimento sobre el protocolo BB84.

El diseño del estudio es claramente experimental y se apoya en un esquema de tipo *crossover*. Participaron 21 grupos de estudiantes en un curso de laboratorio de Física, distribuidos en dos condiciones alternadas: con retroalimentación de ChatGPT y sin retroalimentación. A lo largo del curso, en cuatro momentos clave se plantearon preguntas de comprensión relacionadas con el experimento, y



los estudiantes respondieron de forma verbal. En determinadas fases, el sistema incorporaba la respuesta al modelo de lenguaje, que generaba comentarios formativos comparando lo dicho por el estudiante con una respuesta esperada y guiando, mediante diálogo, hacia explicaciones más completas o precisas. Además, se empleó registro de movimiento ocular para estudiar cómo cambiaba la distribución de la atención visual de los participantes sobre los distintos elementos del entorno (visualizaciones AR, componentes experimentales reales, etc.).

Los resultados muestran que la retroalimentación basada en ChatGPT mejora de forma significativa el desempeño en las preguntas planteadas, en comparación con las situaciones en las que el estudiantado no recibe este apoyo. El análisis cuantitativo de las puntuaciones indica que quienes trabajaron con feedback formativo proporcionado por la IAG alcanzaron explicaciones más completas y precisas. Al mismo tiempo, los datos de seguimiento ocular revelan que la presencia de feedback modifica la atención visual de manera sistemática: cuando las preguntas se centran en el modelo conceptual, la atención tiende a dirigirse hacia las visualizaciones aumentadas que representan los estados y bases cuánticas; cuando las preguntas se orientan a los procedimientos experimentales, aumenta la fijación visual sobre los componentes físicos del montaje. En conjunto, los autores interpretan estos patrones como evidencia de que la IAG puede orientar la atención hacia los elementos más relevantes para la tarea en cada momento.

El estudio también pone de manifiesto algunas limitaciones. Aunque el número de grupos participantes es razonable para un curso de laboratorio, se trata de un contexto muy específico —un experimento de criptografía cuántica con infraestructura AR y registro de mirada—, lo que dificulta extrapolar los resultados a otros tipos de prácticas o a contextos institucionales con menos recursos. Además, la efectividad del feedback depende de la formulación de las preguntas y de la calidad de las respuestas iniciales del estudiantado; cuando las intervenciones de ChatGPT son demasiado generales, la guía puede resultar menos útil y requerir ajustes por parte del docente.

Pese a estas restricciones, el trabajo de Coban et al. aporta una contribución relevante al mostrar que la integración de modelos generativos en experimentos de Física cuántica no solo impacta en el rendimiento en pruebas de comprensión, sino que también influye en la manera en que los estudiantes distribuyen su atención visual sobre el entorno de aprendizaje. El estudio ilustra una línea de trabajo



prometedora, donde la retroalimentación personalizada basada en IAG se combina con entornos inmersivos para apoyar procesos de construcción de significado en contenidos de alta abstracción como la criptografía cuántica.

#### **A4 – Exploring Collaborative Immersive Virtual Reality Serious Games for Enhancing Learning Motivation in Physics Education (Sun et al., 2024)**

El estudio de Sun y colaboradores presenta el diseño y la evaluación de un *serious game* inmersivo en realidad virtual (IVR) orientado a reforzar la motivación y la participación estudiantil en contenidos de Física. La propuesta integra mecánicas colaborativas dentro de un entorno virtual que simula tres tipos de experimentos: lentes convexas (óptica), movimiento parabólico (dinámica) y disminución del punto de congelación (termodinámica). El objetivo central es analizar cómo elementos de juego cuidadosamente diseñados —basados en marcos como Octalysis, Six Facets Design y teorías motivacionales como *flow* y autodeterminación— influyen en la experiencia de aprendizaje y la percepción del estudiantado.

El trabajo se estructura en dos componentes principales: (1) el proceso de diseño del juego, que selecciona y adapta elementos colaborativos y motivacionales para su integración en actividades relacionadas con Física, y (2) una evaluación empírica centrada en las reacciones de los participantes durante y después del uso del sistema. A diferencia de los estudios experimentales estrictos, la investigación adopta un enfoque empírico exploratorio para documentar el comportamiento del estudiantado dentro del entorno IVR y valorar la pertinencia pedagógica del diseño.

La fase de evaluación involucró a 22 participantes universitarios (12 mujeres y 10 hombres, edad promedio 22.7 años), quienes interactuaron en parejas dentro del entorno virtual. Tras una sesión introductoria, los estudiantes completaron actividades del juego, respondieron cuestionarios posteriores y resolvieron pequeñas evaluaciones integradas dentro del entorno IVR. Este diseño permitió observar tanto la participación colaborativa como la respuesta a los retos y recompensas que caracterizan al *serious game*.

Los resultados indican que los elementos inmersivos y colaborativos favorecen la motivación y la implicación de los estudiantes durante la actividad. La combinación de retos, mecánicas narrativas, avatares, enemigos y misiones, junto con los experimentos de Física incorporados, genera una



experiencia atractiva que, según los cuestionarios, incrementa la intención de continuar aprendiendo dentro del entorno virtual. Aunque el estudio no reporta mejoras de aprendizaje conceptual medidas de manera formal, sí ofrece evidencia de que el diseño colaborativo del juego facilita que los participantes se mantengan activos, comunicativos y orientados al objetivo.

Entre las limitaciones señaladas, los autores destacan el tamaño reducido de la muestra, la brevedad de las sesiones y el hecho de que el estudio se realizó en condiciones controladas, lo cual limita la generalización a cursos reales de laboratorio. Asimismo, al centrarse principalmente en la motivación, el estudio no evalúa con rigor el impacto del IVR en la comprensión conceptual, por lo que se requieren investigaciones futuras con diseños más comparativos y evaluaciones de aprendizaje más detalladas.

En conjunto, el trabajo de Sun et al. aporta evidencia preliminar del potencial de los entornos VR colaborativos para aumentar la motivación en actividades de Física, y subraya la importancia de integrar principios de diseño de juegos y dinámicas cooperativas en propuestas educativas inmersivas. Su contribución radica tanto en el marco de diseño ofrecido como en la documentación empírica inicial que valida su pertinencia en contextos de enseñanza de la Física.

#### **A5 – Mapping Physical Moroccan Sciences Students’ Perceptions of AI: A Case Study on Generative AI (ChatGPT) (Tafhi et al., 2025)**

El estudio de Tafhi y colaboradores analiza las percepciones que estudiantes marroquíes de Física tienen sobre la IAG, con énfasis en el uso y la valoración de herramientas generativas como ChatGPT. La investigación parte de un interés por comprender cómo se posiciona el estudiantado frente a la creciente presencia de la IAG en entornos educativos y cuáles son las expectativas, reservas y posibilidades que asocian a su integración en la enseñanza de las ciencias.

El enfoque metodológico es cuantitativo, descriptivo y de carácter exploratorio. Los autores elaboraron un cuestionario estructurado, validado por cinco docentes expertos, con reactivos orientados a captar conocimientos previos sobre IAG, experiencias de uso y percepciones respecto a la posibilidad de que sistemas generativos sustituyan funciones humanas, incluido al profesorado de Física. El instrumento también se diseñó en colaboración con estudiantes, lo que permitió ajustar la redacción a sus marcos de comprensión y reducir sesgos asociados a interpretaciones ambiguas.



La muestra fue amplia: 300 respuestas válidas de un total de 350 participantes, pertenecientes a diferentes niveles educativos dentro de la formación en ciencias físicas (L1, L2, L3, M1 y M2), además de estudiantes de bachillerato especializados en Física. La diversidad del grupo permitió obtener una visión transversal de cómo se entiende y valora la IA en distintos momentos del recorrido formativo. El análisis se realizó con herramientas básicas en Excel, centrado en frecuencias y porcentajes.

Los resultados revelan un panorama mixto. Una mayoría significativa declara estar familiarizada con la IAG, aunque gran parte del estudiantado nunca había interactuado directamente con ChatGPT. Pese a ello, la percepción general sugiere que la IA es vista como una herramienta con potencial para complementar el aprendizaje, pero no como un sustituto del profesorado: más del 70 % considera que ChatGPT no puede reemplazar la labor docente en Física. Asimismo, una proporción considerable de participantes expresa dudas sobre la capacidad de la IAG para desempeñar roles humanos en ámbitos más amplios. Estas tendencias reflejan una combinación de curiosidad, cautela y conciencia de las limitaciones actuales de las herramientas generativas.

Entre las limitaciones del estudio, los autores señalan que la investigación se basa únicamente en autoinformes, sin triangulación con datos de desempeño o análisis cualitativos. Además, al centrarse en percepciones generales sobre la IAG y no en experiencias concretas de uso en tareas de Física, el alcance de los hallazgos es principalmente actitudinal. Aun así, el estudio ofrece una aproximación valiosa a la manera en que los futuros profesionales de ciencias físicas interpretan el papel emergente de la IAG en la educación, destacando la necesidad de estrategias formativas que integren estas herramientas de manera crítica y contextualizada.

#### **A6 – Development of an AI-Supported Chatbot as an Interactive Learning Platform in STEM Education (Fekets, 2025)**

El trabajo de Fekets presenta el diseño y la evaluación inicial de un chatbot educativo apoyado en modelos generativos, integrado en una plataforma orientada a estudiantes de áreas STEM. Aunque la propuesta no se centra exclusivamente en Física, la asignatura forma parte del currículo al que va dirigida la herramienta, por lo que el estudio ofrece información relevante para comprender cómo podrían incorporarse sistemas conversacionales en actividades de aprendizaje científico.



El artículo describe el proceso de desarrollo del chatbot, que combina técnicas de PLN, generación de lenguaje y análisis de intenciones para ofrecer respuestas personalizadas a dudas conceptuales y procedimentales. La plataforma permite que el estudiante interactúe por medio de consultas abiertas, reciba explicaciones, acceda a ejemplos y explore contenido suplementario. El objetivo es facilitar un entorno de apoyo accesible que complemente el trabajo en el aula y fomente el aprendizaje autónomo. La evaluación realizada es de naturaleza exploratoria y se centra en aspectos de uso, usabilidad y experiencia percibida. Participaron 54 estudiantes de formación STEM, quienes utilizaron la plataforma y respondieron cuestionarios orientados a valorar la utilidad del chatbot, la claridad de las respuestas generadas y la facilidad de interacción. El análisis se apoyó en indicadores descriptivos que permitieron identificar patrones de satisfacción y áreas susceptibles de mejora.

Los resultados señalan que los estudiantes valoran positivamente la inmediatez de las respuestas y la posibilidad de aclarar dudas sin depender del docente. La mayoría percibe que el chatbot contribuye a reforzar explicaciones y a generar un ambiente de aprendizaje más flexible. Sin embargo, el estudio también muestra que las respuestas del modelo no siempre son suficientemente específicas y que, en ocasiones, requieren interpretación o verificación adicional. Esto evidencia la necesidad de ajustar los mecanismos de retroalimentación y de mejorar la alineación entre las preguntas del usuario y las respuestas generadas.

Entre las limitaciones reconocidas, se menciona que la evaluación se basó únicamente en percepciones, sin mediciones directas de aprendizaje o desempeño. Además, el uso de un único grupo limita la posibilidad de establecer comparaciones o generalizar los resultados a otros contextos educativos. Aun así, el estudio aporta una aproximación clara y útil sobre la factibilidad de integrar chatbots basados en IAG en entornos STEM, destacando su potencial para apoyar el aprendizaje autónomo y complementar las explicaciones del docente.

## **Discusión**

La discusión de los hallazgos se fundamenta en un corpus reducido de seis estudios empíricos, por lo que las interpretaciones deben entenderse como tendencias iniciales más que como conclusiones consolidadas del campo. Aun así, los resultados permiten identificar patrones relevantes sobre el uso de



la IAG en la enseñanza de la Física, así como las condiciones bajo las cuales estas tecnologías pueden favorecer el aprendizaje y la participación estudiantil.

En primer lugar, los estudios analizados muestran que la IAG puede contribuir a transformar ciertos aspectos de la enseñanza de la Física, especialmente en tareas que requieren explicación paso a paso, retroalimentación inmediata o la generación de múltiples representaciones del mismo concepto. En varios casos se observaron indicios de mejora en la comprensión conceptual, la motivación o la autonomía del alumnado; sin embargo, estos efectos varían según el diseño, el contexto y el tipo de interacción con el modelo generativo, lo que impide establecer conclusiones generalizables.

Otro patrón relevante se relaciona con las formas de mediación docente. La evidencia disponible apunta a una posible efectividad cuando la IAG se integra como herramienta de apoyo guiado —por ejemplo, para estructurar experimentos, resolver problemas de manera incremental o generar explicaciones alternativas— más que como sustituto de la instrucción. En este sentido, la IAG aparece como una estrategia pedagógica emergente con potencial para complementar prácticas tradicionales y promover un aprendizaje más activo, siempre que su uso se acompañe de orientaciones claras y objetivos didácticos definidos.

Este panorama coincide con hallazgos recientes en educación científica que destacan tanto el potencial de la IA para apoyar procesos de comprensión como los desafíos asociados a su precisión, sesgos y necesidad de supervisión disciplinar (Erümit & Sarialioğlu, 2025). Estas investigaciones externas refuerzan la importancia de introducir la IAG en ambientes controlados, con una guía docente explícita y con criterios sólidos para validar la pertinencia y calidad de las respuestas generadas por los modelos. Asimismo, los estudios ponen de relieve la necesidad de avanzar hacia diseños metodológicos más robustos, que incluyan comparaciones entre grupos, mediciones longitudinales o análisis que permitan distinguir entre mejoras atribuibles a la IAG y aquellas derivadas de factores contextuales. Los enfoques exploratorios predominan en el corpus, lo que resulta comprensible dado el carácter emergente del campo, pero también limita la capacidad de comprender cómo y por qué la IAG influye en los procesos de aprendizaje en Física.

Finalmente, los hallazgos destacan la importancia de considerar aspectos éticos, pedagógicos y de alfabetización digital. El uso de la IAG exige que el alumnado desarrolle criterios para evaluar la



pertinencia y calidad de las respuestas generadas, mientras que el profesorado requiere estrategias para integrar estas tecnologías sin generar dependencia o reducir la autonomía cognitiva de los estudiantes. En conjunto, los estudios analizados permiten identificar patrones preliminares sobre los beneficios y desafíos de la IAG en la enseñanza de la Física, aunque sin pretender generalizar los efectos observados.

## CONCLUSIONES

La revisión sistemática permitió identificar tendencias iniciales sobre el uso de la IAG en la enseñanza de la Física, basadas en un corpus reducido de seis estudios empíricos. Aunque la evidencia disponible es todavía incipiente, los resultados sugieren que la IAG puede desempeñar un papel relevante como herramienta de apoyo para la explicación de conceptos, la generación de representaciones alternativas y la retroalimentación inmediata, siempre que su integración se diseñe de forma pedagógicamente fundamentada.

Los estudios analizados muestran indicios de mejora en aspectos como la comprensión conceptual, la autonomía y la motivación estudiantil, aunque estos efectos varían según el tipo de tarea, el nivel educativo y la forma de mediación docente. En consecuencia, los hallazgos no permiten establecer generalizaciones amplias, pero sí delinear escenarios prometedores para la incorporación de la IAG en actividades experimentales, resolución de problemas y entornos de práctica guiada.

Asimismo, la revisión destaca desafíos persistentes: la necesidad de validar la precisión de las respuestas generadas, evitar sesgos o simplificaciones excesivas, fortalecer la alfabetización digital del alumnado y asegurar un rol docente claro en la supervisión del proceso. Estas consideraciones son especialmente relevantes ante la rápida evolución de los modelos generativos y su creciente disponibilidad en contextos educativos.

En conjunto, esta revisión contribuye a clarificar el estado actual del campo y a sintetizar la evidencia emergente sobre los usos pedagógicos de la IAG en Física. Los resultados invitan a profundizar en líneas de investigación orientadas a: (a) desarrollar diseños experimentales más robustos, (b) explorar efectos sostenidos a través de estudios longitudinales, (c) analizar la interacción entre IAG y distintos perfiles de estudiantes, y (d) evaluar el impacto de estas herramientas en competencias científicas específicas más allá de la comprensión conceptual. Avanzar en estas direcciones permitirá consolidar un marco más preciso para la integración efectiva y responsable de la IAG en la enseñanza de la Física.



## Limitaciones del estudio

Esta revisión presenta algunas limitaciones que deben considerarse al interpretar los resultados. En primer lugar, aunque el corpus final integra estudios que abordan distintos niveles educativos, tecnologías y contextos de aplicación, la muestra sigue siendo reducida y heterogénea. Esta diversidad, si bien permite observar un panorama amplio, dificulta establecer relaciones comparativas sólidas entre los enfoques analizados.

En segundo lugar, la calidad metodológica de los estudios incluidos es variable. Una parte importante del corpus se basa en diseños exploratorios, muestras pequeñas o evaluaciones centradas en percepciones, lo que restringe la posibilidad de generalizar los hallazgos. La ausencia de estudios longitudinales o diseños experimentales consistentes limita, además, la capacidad de valorar el impacto sostenido de la IAG en la comprensión conceptual de la Física.

Asimismo, la revisión se circunscribe a trabajos publicados entre 2018 y 2025 y a un conjunto específico de bases de datos. Es posible que existan investigaciones recientes o en otros repositorios que no fueron capturadas por los criterios de búsqueda aplicados. Por último, aunque el análisis se realizó con procedimientos sistemáticos y criterios explícitos de inclusión y exclusión, la interpretación de los resultados depende inevitablemente del juicio académico, lo cual introduce un componente subjetivo que no puede eliminarse por completo. Estas restricciones metodológicas y de cobertura obligan a interpretar los hallazgos con cautela, sin pretender generalizaciones amplias, aunque ello no disminuye la relevancia de las evidencias identificadas ni la necesidad de continuar investigando este ámbito emergente. Aun con estas limitaciones, la revisión ofrece una síntesis sólida y actualizada del estado de la investigación sobre IAG en la enseñanza de la Física, y constituye un punto de partida para orientar futuras líneas de trabajo en este campo emergente.

Asimismo, esta revisión presenta algunas restricciones adicionales. En primer lugar, el uso exclusivo de literatura en inglés introduce un posible sesgo de idioma, ya que investigaciones pertinentes publicadas en otros contextos lingüísticos pudieron quedar fuera del corpus final. En segundo lugar, la exclusión deliberada de literatura gris y preprints —incluyendo repositorios como arXiv, ResearchGate o informes institucionales— restringe el acceso a trabajos emergentes en un campo que avanza con rapidez, lo cual podría limitar la detección de tendencias muy recientes. Finalmente, el reducido número de estudios



empíricos disponibles obliga a interpretar los hallazgos con cautela, sin pretender generalizaciones amplias, aunque ello no disminuye la relevancia de las evidencias identificadas ni la necesidad de continuar investigando este ámbito.

### **Líneas de investigación futura**

Los estudios revisados muestran avances importantes, pero también evidencian vacíos que deben atenderse para consolidar el campo. Una primera línea de investigación consiste en desarrollar estudios con diseños experimentales y cuasi-experimentales más robustos, que permitan comparar la efectividad de diferentes tipos de intervenciones basadas en IAG y evaluar su impacto en la comprensión conceptual, el razonamiento científico y la resolución de problemas en Física.

Otra vía prioritaria es la realización de investigaciones longitudinales que analicen efectos sostenidos de la interacción con modelos generativos, particularmente en cursos que integran actividades de laboratorio, simulación computacional o entornos inmersivos. Este enfoque permitiría comprender cómo evoluciona el desempeño del estudiantado a lo largo del tiempo y qué condiciones favorecen aprendizajes duraderos.

También resulta necesario estudiar con mayor profundidad los mecanismos cognitivos involucrados en la interacción con la IAG. El uso de técnicas como el seguimiento ocular, el análisis de trazas digitales o la evaluación de patrones discursivos puede ayudar a identificar cómo estas herramientas influyen en la atención, la construcción de explicaciones o la toma de decisiones durante la resolución de tareas de Física.

Asimismo, se requieren investigaciones que aborden la integración pedagógica de la IAG desde marcos teóricos explícitos. Explorar cómo se articulan estas herramientas con enfoques como el aprendizaje por indagación, el modelado científico o el aprendizaje colaborativo permitiría diseñar intervenciones más coherentes y ajustadas a las necesidades formativas de distintas poblaciones estudiantiles.

Por último, es importante avanzar hacia estudios que consideren aspectos éticos, epistemológicos y sociotécnicos relacionados con el uso de modelos generativos en educación científica. Analizar cómo se enseñan prácticas de verificación, manejo de incertidumbre y detección de errores en la interacción con IA contribuirá a formar estudiantes más críticos y capaces de utilizar estas herramientas de forma responsable.



## **Implicaciones para la práctica docente**

Los estudios analizados muestran que la IAG puede desempeñar un papel valioso como recurso de acompañamiento en la enseñanza de la Física, siempre que su integración se realice de manera crítica y guiada. Una primera implicación para la práctica es la necesidad de incorporar estas herramientas como apoyo al razonamiento y no como sustitutos de la explicación docente. La IA puede facilitar el acceso a ejemplos, retroalimentación inmediata o estrategias de modelación, pero requiere que los estudiantes aprendan a validar, contrastar y reformular la información que reciben.

Asimismo, la evidencia señala que la incorporación de sistemas conversacionales, asistentes de laboratorio o entornos inmersivos debe acompañarse de orientaciones claras sobre cómo formular preguntas, interpretar sugerencias y reconocer posibles imprecisiones en las respuestas generadas por la IA. Estas habilidades metacognitivas resultan esenciales para evitar una dependencia acrítica y para fortalecer la autonomía en el aprendizaje científico.

Otra implicación relevante es la necesidad de diseñar actividades que integren la IA dentro de secuencias didácticas coherentes. En programación computacional, por ejemplo, los modelos generativos pueden utilizarse para explorar alternativas de código o identificar errores, mientras que en laboratorios presenciales pueden apoyar la comprensión de procedimientos y decisiones experimentales. En ambos casos, su uso es más efectivo cuando se articula con metas de aprendizaje explícitas y con espacios de discusión que permitan interpretar los resultados.

Finalmente, la diversidad de tecnologías analizadas sugiere que la formación docente debe incluir oportunidades para conocer y evaluar críticamente distintas aplicaciones de la IAG. Esto implica no solo familiarizarse con sus usos potenciales, sino también comprender sus limitaciones, los sesgos que pueden introducir y las condiciones que favorecen un aprendizaje significativo. Integrar estas herramientas de manera responsable puede contribuir a crear experiencias educativas más inclusivas, motivadoras y adaptadas a las necesidades de los estudiantes.



## REFERENCIAS BIBLIOGRÁFICAS

- Archambault, É., Campbell, D., Gingras, Y., & Larivière, V. (2009). Comparing bibliometric statistics obtained from the Web of Science and Scopus. *Journal of the American Society for Information Science and Technology*, 60(7), 1320-1326. <https://doi.org/10.1002/asi.21062>
- Aria, M., & Cuccurullo, C. (2017). bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959-975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Avila, K. E., Steinert, S., Ruzika, S., Kuhn, J., & Küchemann, S. (2024). Using ChatGPT for teaching physics. *The Physics Teacher*, 62(6), 536-537. <https://doi.org/10.1119/5.0227132>
- Bitzenbauer, P. (2023). ChatGPT in physics education: A pilot study on easy-to-implement activities. *Contemporary Educational Technology*, 15(3), ep430. <https://doi.org/10.30935/cedtech/13176>
- Booth, A., Papaioannou, D., & Sutton, A. (2016). *Systematic Approaches to a Successful Literature Review* (Second Edition).
- Coban, A., Dzsotjan, D., Küchemann, S., Durst, J., Kuhn, J., & Hoyer, C. (2025). AI support meets AR visualization for Alice and Bob: Personalized learning based on individual ChatGPT feedback in an AR quantum cryptography experiment for physics lab courses. *EPJ Quantum Technology*, 12(1), 15. <https://doi.org/10.1140/epjqt/s40507-025-00310-z>
- Donthu, N., Kumar, S., Mukherjee, D., Pandey, N., & Lim, W. M. (2021). How to conduct a bibliometric analysis: An overview and guidelines. *Journal of Business Research*, 133, 285-296. <https://doi.org/10.1016/j.jbusres.2021.04.070>
- Erümit, A. K., & Sarıalioğlu, R. Ö. (2025). Artificial intelligence in science and chemistry education: A systematic review. *Discover Education*, 4(1), 178. <https://doi.org/10.1007/s44217-025-00622-3>
- Fekets, G. (2025). DEVELOPMENT OF AN ARTIFICIAL INTELLIGENCE-SUPPORTED CHATBOT AS AN INTERACTIVE LEARNING PLATFORM IN STEM EDUCATION: EXPLORING USABILITY AND STUDENT EXPERIENCE. *Journal of Baltic Science Education*, 24(5), 878-893. <https://doi.org/10.33225/jbse/25.24.878>
- Gast, D. L., & Ledford, J. R. (2024). *Single Case Research Methodology: Applications in Special Education and Behavioral Sciences* (Fourth Edition). <https://www.routledge.com/Single-Case->



Research-Methodology-Applications-in-Special-Education-and-Behavioral-Sciences/Ledford-Gast/p/book/9781032265810

- Gough, D., Oliver, S., Thomas, J., & Thomas, J. (Eds.). (2017). *An introduction to systematic reviews* (2nd edition). SAGE.
- Harzing, A.-W., & Alakangas, S. (2016). Google Scholar, Scopus and the Web of Science: A longitudinal and cross-disciplinary comparison. *Scientometrics*, *106*(2), 787-804. <https://doi.org/10.1007/s11192-015-1798-9>
- Holmes, N. G., Olsen, J., Thomas, J. L., & Wieman, C. E. (2017). Value added or misattributed? A multi-institution study on the educational benefit of labs for reinforcing physics content. *Physical Review Physics Education Research*, *13*(1), 010129. <https://doi.org/10.1103/PhysRevPhysEducRes.13.010129>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Latif, E., Parasuraman, R., & Zhai, X. (2024). PhysicsAssistant: An LLM-Powered Interactive Learning Robot for Physics Lab Investigations. *IEEE Int. Workshop Robot Human Commun., RO-MAN*, 864-871. <https://doi.org/10.1109/RO-MAN60168.2024.10731312>
- Lee, G., Yun, M., Zhai, X., & Crippen, K. (2025). Artificial Intelligence in Science Education Research: Current States and Challenges. *Journal of Science Education and Technology*. <https://doi.org/10.1007/s10956-025-10239-8>
- Liang, Y., Zou, D., Xie, H., & Wang, F. L. (2023). Exploring the potential of using ChatGPT in physics education. *Smart Learning Environments*, *10*(1), 52. <https://doi.org/10.1186/s40561-023-00273-7>
- Lo, C. K. (2023). What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature. *Education Sciences*, *13*(4). <https://doi.org/10.3390/educsci13040410>



- Luckin, R., & Holmes, W. (2016). *Intelligence Unleashed: An argument for AI in Education*: Pearson.  
[https://www.researchgate.net/publication/299561597\\_Intelligence\\_unleashed\\_An\\_argument\\_for\\_AI\\_in\\_education](https://www.researchgate.net/publication/299561597_Intelligence_unleashed_An_argument_for_AI_in_education)
- Martín-Martín, A., Orduna-Malea, E., Thelwall, M., & Delgado López-Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of Informetrics*, 12(4), 1160-1177. <https://doi.org/10.1016/j.joi.2018.09.002>
- Meltzer, D. E., & Thornton, R. K. (2012). Resource Letter ALIP-1: Active-Learning Instruction in Physics. *American Journal of Physics*, 80(6), 478-496. <https://doi.org/10.1119/1.3678299>
- Nishiwaki, Y. (2025). Application of generative AI-based programming in physics education for non-English speaking students. *European Journal of Physics*, 46(5), 055703. <https://doi.org/10.1088/1361-6404/adfc16>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). *The PRISMA 2020 statement: An updated guideline for reporting systematic reviews*. <https://doi.org/10.1136/bmj.n71>
- Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education*, 147, 103778. <https://doi.org/10.1016/j.compedu.2019.103778>
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research*, 104, 333-339. <https://doi.org/10.1016/j.jbusres.2019.07.039>
- Sun, L., Lee, B. G., Chieng, D., & Yang, S. (2024). Exploring Collaborative Immersive Virtual Reality Serious Games for Enhancing Learning Motivation in Physics Education. *2024 IEEE 48th Annual Computers, Software, and Applications Conference (COMPSAC)*, 115-120. <https://doi.org/10.1109/COMPSAC61105.2024.00026>
- Tafhi, O., Khattabi, K. E., Lamrabet, M., Ouzennou, F., & Chikhaoui, A. (2025). Mapping Physical Moroccan Sciences Student's Perceptions of AI: A Case Study on Generative AI (ChatGPT). En



- International Journal of Information and Education Technology* (Vol. 15, Número 1, pp. 175-186). <https://doi.org/10.18178/ijiet.2025.15.1.2230>
- Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8(1), 45. <https://doi.org/10.1186/1471-2288-8-45>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. <https://doi.org/10.1186/s40561-023-00237-x>
- UNESCO. (2021). *AI and education: Guidance for policy*. <https://unesdoc.unesco.org/>
- UNESCO. (2023). *Guidance for generative AI in education and research*. <https://unesdoc.unesco.org/>
- VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring Systems, and Other Tutoring Systems. *Educational Psychologist*, 46(4), 197-221. <https://doi.org/10.1080/00461520.2011.611369>
- Wang, S., Wang, F., Zhu, Z., Wang, J., Tran, T., & Du, Z. (2024). Artificial intelligence in education: A systematic literature review. *Expert Systems with Applications*, 252, 124167. <https://doi.org/10.1016/j.eswa.2024.124167>
- Wieman, C. E. (2014). Large-scale comparison of science teaching methods sends clear message. *Proceedings of the National Academy of Sciences*, 111(23), 8319-8320. <https://doi.org/10.1073/pnas.1407304111>
- Xiao, Y., & Watson, M. (2019). Guidance on Conducting a Systematic Literature Review. *Journal of Planning Education and Research*, 39(1), 93-112. <https://doi.org/10.1177/0739456X17723971>
- Xu, W., & Ouyang, F. (2022). The application of AI technologies in STEM education: A systematic review from 2011 to 2021. *International Journal of STEM Education*, 9(1), 59. <https://doi.org/10.1186/s40594-022-00377-5>
- Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39. <https://doi.org/10.1186/s41239-019-0171-0>



Zhai, X. (2024). Transforming Teachers' Roles and Agencies in the Era of Generative AI: Perceptions, Acceptance, Knowledge, and Practices. *Journal of Science Education and Technology*.  
<https://doi.org/10.1007/s10956-024-10174-0>

Zupic, I., & Čater, T. (2015). Bibliometric Methods in Management and Organization. *Organizational Research Methods*, 18(3), 429-472. <https://doi.org/10.1177/1094428114562629>

