

Artículo Científico

Aprendizaje personalizado con IA generativa para enseñar la conservación de la energía mecánica según niveles cognitivos.

Personalized learning with generative AI used to teach mechanical energy conservation across cognitive levels.

Guerrero-Zambrano, Marcos¹; Cevallos-Torres, Lorenzo²; Rumbaut-Rangel, Dayron³.

¹ Universidad Bolivariana del Ecuador (UBE); Universidad Estatal de Milagro; Ecuador, Durán; <https://orcid.org/0000-0002-5617-6836>; mfguerreroz_a@ube.edu.ec; mguerreroz@unemi.edu.ec

² Universidad Bolivariana del Ecuador (UBE); Ecuador, Durán; <https://orcid.org/0000-0002-7211-2891>; ljcevallost@ube.edu.ec

³ Universidad Bolivariana del Ecuador (UBE); Ecuador, Durán; <https://orcid.org/0009-0001-9087-0979>; drumbautr@ube.edu.ec

¹ Autor Correspondencia



<https://doi.org/10.63618/omd/isj/v4/n3/326>

Cita: Guerrero-Zambrano, M., Cevallos-Torres, L., & Rumbaut-Rangel, D. (2026). Aprendizaje personalizado con IA generativa para enseñar la conservación de la energía mecánica según niveles cognitivos. *Innova Science Journal*, 4(3), 33-51. <https://doi.org/10.63618/omd/isj/v4/n3/326>

Recibido: 26/12/2025

Aceptado: 12/05/2026

Publicado: 31/07/2026



Copyright: © 2026 por los autores. Este artículo es un artículo de acceso abierto distribuido bajo los términos y condiciones de la Licencia Creative Commons, Atribución-NoComercial 4.0 Internacional. (CC BY-NC).

(<https://creativecommons.org/licenses/by-nc/4.0/>)

Resumen: La investigación explora el potencial de la inteligencia artificial generativa para personalizar el aprendizaje de la ley de conservación de la energía mecánica a futuros profesores de Física de una universidad pública ecuatoriana, considerando las limitaciones de infraestructura y los retos para incorporar pedagógicamente tecnologías emergentes. Método: Se empleó un enfoque mixto explicativa secuencial con predominio cuantitativo y un diseño preexperimental de grupo único (n = 9). Para la recolección de datos, se empleó una prueba de opción múltiple alineado con los niveles cognitivos, tanto para el pretest como para el posttest; una encuesta de Likert que incluyó preguntas abiertas sobre la experiencia con un tutor generativo usando Gemini. Resultados: Los hallazgos mostraron un aumento en el promedio de aciertos que va desde el 52,8 % hasta el 88,0 % donde se observaron avances notables en la comprensión, aplicación y creación, así como un progreso hacia el nivel máximo cognitivo. Además, la percepción estudiantil fue positiva en cuanto a utilidad, motivación e impacto en el aprendizaje, con una carga cognitiva moderada. Conclusiones: La tutoría socrática mediada por IA, con supervisión docente, favorece aprendizajes más profundos y un uso crítico de la tecnología, aunque se requieren muestras mayores y diseños comparativos para fortalecer la validez externa de los hallazgos.

Palabras clave: Inteligencia artificial generativa; Aprendizaje personalizado; Tutoría socrática; Conservación de la energía mecánica; Formación de profesores de Física.

Abstract: This study explores the potential of generative artificial intelligence to personalize the learning of the law of conservation of mechanical energy among pre-service Physics teachers at an Ecuadorian public university, taking into account infrastructure limitations and the challenges of pedagogically integrating emerging technologies. Method: A quantitatively driven sequential explanatory mixed-methods approach was employed, using a single-group pre-experimental design (n = 9). Data were collected through a multiple-choice test aligned with cognitive levels, administered as both a pretest and a posttest. In addition, a Likert-scale survey was applied, which included open-ended questions regarding participants' experience with a generative tutor based on Gemini. Results: The findings showed an increase in the average percentage of correct answers, rising from 52.8% to 88.0%. Notable gains were observed in understanding, application, and creation, along with progress toward the highest cognitive level. Conclusions: AI-mediated Socratic tutoring, under teacher supervision, fosters deeper learning and a more critical use of technology; however, larger samples and comparative designs are required to strengthen the external validity of the findings.

Keywords: Generative Artificial Intelligence; Personalized Learning; Socratic Tutoring; Mechanical Energy Conservation; Physics Teacher Education.

1. Introducción

En los últimos años, la Inteligencia Artificial Generativa (IAG) se ha perfilado como una tecnología potencialmente capaz de transformar el futuro de la educación, personalizando el aprendizaje y mejorando los resultados de aprendizaje (Ogunleye et al., 2024; Prather et al., 2025). En educación, este potencial es especialmente relevante cuando se trabaja con contenidos que requieren múltiples representaciones, explicación causal y transferencia a nuevas situaciones.

En la enseñanza de la Física se siguen presentando dificultades en la comprensión de conceptos relacionados con energía, el trabajo y muy en especial, la Ley de Conservación de la Energía Mecánica (LCEM). La literatura informa concepciones alternativas recurrentes, por ejemplo, la idea de que la energía se pierde o que solo está presente cuando hay movimiento observable, lo cual produce aprendizajes fragmentados y explicaciones inestables (Guerra-Reyes et al., 2024; Halilovic et al., 2021; Kotsis & Panagou, 2023). Estas dificultades no son únicamente conceptuales; sino que comprometen la interpretación de representaciones, la elección del sistema físico y la justificación de los resultados.

Desde el punto de vista ético como pedagógico, la Inteligencia Artificial (IA) educativa no es neutra. Su valor para la formación está sujeto a la construcción de la interacción, al rol del docente y al tipo de andamiaje ofrecido al estudiante (Holmes et al., 2022). En esta línea, estudios recientes indican que los tutores basados en modelos de lenguaje pueden apoyar procesos de explicación, práctica guiada y retroalimentación inmediata, cuando se articulan con principios pedagógicos sólidos y con supervisión docente (Kestin et al., 2025; Memon & Kwan, 2025; Tu et al., 2025).

En América Latina y, en particular, en Ecuador, esta discusión adquiere una importancia especial. En la literatura se identifican desconexiones entre los textos de ciencias, las prácticas pedagógicas en el aula y la incorporación de tecnologías emergentes, así como también limitaciones de infraestructura, acceso y formación docente para integrar la IA de forma equitativa y pedagógicamente pertinente (Ade Nandang Mustafa, 2025; Barrera & Azeez, 2024; Castillo Villegas et al., 2025; Valdivieso & González, 2025). En este escenario, todavía se cuenta con escasa evidencia empírica situada que articule tutoría con la IAG, personalización explícita por niveles cognitivos y en un contenido exigente como la LCEM.

A partir de la brecha antes mencionada, el presente estudio se plantea como un piloto orientado a explorar los cambios observados en el aprendizaje de la LCEM y la percepción asociada al uso de un tutor basado en IAG, estructurado según la Taxonomía de Bloom, en futuros docentes de Física. En particular, el propósito general de la investigación es analizar el impacto de una intervención de aprendizaje personalizado con IAG, diseñada por niveles cognitivos de la taxonomía de Bloom, sobre el desarrollo del conocimiento conceptual y aplicado de la LCEM en futuros docentes de Física.

En coherencia con este propósito, la investigación se orienta por las siguientes preguntas: ¿qué nivel de conocimiento inicial por niveles de la Taxonomía de Bloom presentan los futuros docentes sobre la LCEM?, ¿cómo se estructura la adaptación del tutor con IAG en los distintos niveles?, ¿qué percepción tienen los futuros

docentes de Física sobre el uso de la IAG como apoyo al aprendizaje personalizado de la LCEM?, y ¿qué cambios se observan en el rendimiento académico y en el perfil cognitivo tras la intervención en aula?. Responder estas preguntas permitirá aportar evidencia empírica situada sobre el potencial de la IAG para apoyar procesos de personalización del aprendizaje en Física en contextos ecuatorianos, así como orientar el diseño de guías, rúbricas y recomendaciones pedagógicas transferibles a otros contenidos y niveles educativos.

El aporte del estudio se ubica en tres planos que se complementan entre sí. En términos pedagógicos, evidencia una secuencia de interacción basada en preguntas y verificación para un contenido de gran densidad conceptual (Fakour & Imani, 2025; Tong et al., 2023). Desde el punto de vista metodológico, organiza la evaluación por niveles cognitivos con el fin de describir con mayor resolución el perfil de logro del grupo (Almatrafi & Johri, 2025). En el ámbito educativo, revaloriza la opinión del futuro profesor acerca de la utilidad y las limitaciones de la IAG como ayuda en la clase (García Ramos & Wilson-Kennedy, 2024; Kalniņa et al., 2024).

2. Materiales y Métodos

2.1. Diseño, enfoque, tipo y nivel de investigación

La investigación se realizó mediante un enfoque mixto secuencial explicativo con predominio cuantitativo. En una primera fase se estudiaron los cambios observados en el dominio conceptual y cognitivo aplicado de la LCEM, a través de mediciones pretest y posttest, y en una segunda fase se incorporaron respuestas abiertas y trazas de interacción con el tutor para complementar la lectura de los resultados numéricos. La integración de ambos componentes se hizo a nivel interpretativo, contrastando los patrones de desempeño con la percepción reportada por los participantes (Araya-Muñoz et al., 2019; Moore et al., 2015).

En la parte cuantitativa, se empleó un diseño preexperimental de grupo único con medición pre-post, organizado en una estrategia didáctica con cuatro momentos, denominada DEAPE, que significa Diagnóstico, Explicación, Aprendizaje Individualizado con IAG y Evaluación, lo cual permite el seguimiento intra-sujeto del desempeño en tareas alineadas a niveles cognitivos progresivos y comparables. Debido al carácter de estudio piloto con un único grupo y $n = 9$ participantes, los resultados se ofrecen como evidencia preliminar de cambios observados, pero no como prueba concluyente (Araya-Muñoz et al., 2019; Hernández-Sampieri & Mendoza, 2023; Krathwohl, 2002).

Por su finalidad, el tipo de la investigación es aplicado y evaluativo, dado que tiene como objetivo evaluar una estrategia de personalización del aprendizaje con IAG en un contenido disciplinar concreto como es LCEM y la generación de evidencia valiosa para la optimización de la práctica docente, conforme a las propuestas relativas a la evaluación de intervenciones y recursos educativos en contextos de enseñanza superior (Castillo Villegas et al., 2025). Respecto a su nivel, se propone un enfoque preliminar y piloto, dirigido a la refinación y perfeccionamiento de secuencias, instrumentos y criterios de implementación (Araya-Muñoz et al., 2019; Castillo Villegas et al., 2025).

2.2. Participantes y muestra

La muestra estuvo conformada por 9 estudiantes matriculados en la asignatura Mecánica y Laboratorio de la carrera Pedagogía de las Ciencias Experimentales, en una institución de educación superior pública de la ciudad de Milagro, Ecuador. Son futuros profesores de Física en una etapa inicial de formación y para quienes la LCEM es un contenido relevante para su futura práctica docente en el Bachillerato General Unificado.

La asignatura integra elementos teóricos y experimentales de mecánica, tales como el movimiento, las fuerzas, el trabajo, la energía cinética, la energía potencial gravitacional, la energía potencial elástica, las fuerzas conservativas y las no conservativas, energía degradada, sistema, entorno, frontera y universo. La intervención se llevó a cabo en dicho entorno educativo, donde la LCEM desempeña un papel fundamental como eje de conexión entre la comprensión conceptual y la resolución de problemas.

Los participantes tenían a su disposición Chromebooks con acceso a internet y cuentas institucionales de Google, lo que permitió utilizar el tutor generativo a través de un prompt en Gemini en condiciones homogéneas de acceso durante las sesiones presenciales. La muestra fue no probabilística por conveniencia, característica propia del trabajo con grupos naturales de clase en contextos educativos (Hernández-Sampieri & Mendoza, 2023). Se incluyeron estudiantes que estuviesen matriculados en el curso, aceptasen participar mediante consentimiento informado y tuviesen acceso a internet y dispositivo asignado por la institución.

Se utilizó un único grupo de intervención, sin grupo control, en concordancia con el carácter piloto del estudio y con la intención de documentar la implementación en un contexto real de aula universitaria (Kestin et al., 2025).

2.3. Contexto educativo

La intervención se realizó en una institución universitaria pública ecuatoriana de la ciudad de Milagro, con 9 estudiantes en formación para futuros docentes de Física. La asignatura de Mecánica y su Laboratorio incluye clases teóricas, resolución de problemas y prácticas experimentales, utilizando tanto recursos tradicionales como digitales.

En este espacio, la LCEM no solo fue tratada como contenido conceptual, sino que también fue empleada como eje central para organizar el programa de Física, sirviendo para explicar fenómenos relacionados con movimiento y oscilaciones como los sistemas masa-resorte y péndulos. Por ello, resulta pertinente evaluar el desempeño conceptual y cognitivo aplicado en este tema, a fin de valorar la preparación del futuro docente. En Ecuador se dan retos ya descritos en la literatura, como la diversidad metodológica, las dificultades para incorporar tecnologías emergentes y la persistencia de concepciones alternativas en Física (Guerra-Reyes et al., 2024; Valdivieso & González, 2025), debido a estas condiciones se aplicó un tutor de IAG usando Gemini, acoplado con la Taxonomía de Bloom y mediado por la supervisión docente.

2.4. Estrategia didáctica DEAPE

La duración total de la estrategia didáctica a la que los autores la han denominado DEAPE tuvo una duración de 9 horas, desarrollándose en cuatro fases: Diagnóstico, Explicación, Aprendizaje Personalizado con IAG y Evaluación.

2.4.1. Fase de diagnóstico

La fase de diagnóstico duró una hora y se aplicó una prueba inicial orientada a diagnosticar el nivel cognitivo y el dominio del contenido sobre la LCEM.

El instrumento estuvo formado por 12 ítems de opción múltiple de tipo conceptual y/o numérico, con 4 opciones de respuesta y una respuesta correcta. Los reactivos se alinearon explícitamente con los 6 niveles de la Taxonomía de Bloom, con dos ítems por nivel cada uno, e incorporaron situaciones contextualizadas en problemas típicos de LCEM.

Con el fin de garantizar la validez de contenido, los ítems fueron sometidos en primera instancia al Índice de Validez de Contenido (IVC) mediante el juicio de tres expertos en pedagogía de la física, quienes juzgaron la pertinencia, claridad y representatividad de cada ítem. A partir de ello, se modificaron los ítems de acuerdo a la evaluación del IVC, y nuevamente con los cambios realizados se evaluaron los ítems con el IVC, obteniendo la máxima nota por parte de los expertos.

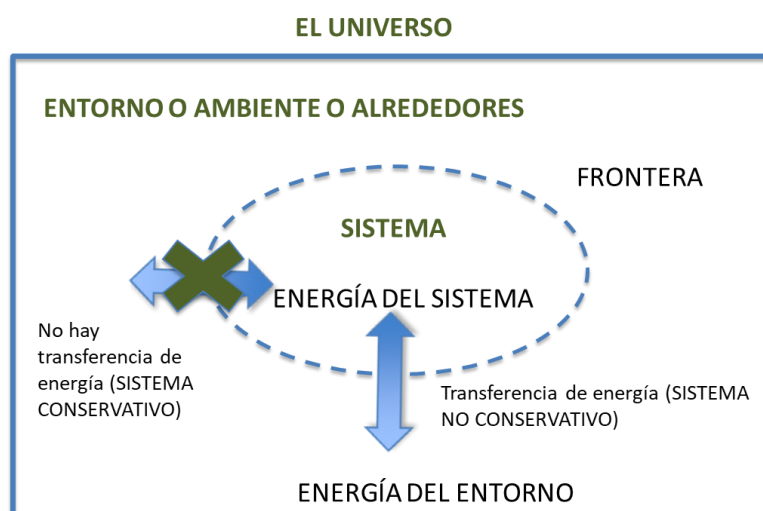
2.4.2. Fase de Explicación

Tras el diagnóstico, el docente desarrolló una fase de explicación de 2 horas, en los que se abordó los temas de energía cinética, energía potencial gravitacional y elástica, fuerzas conservativas y no conservativas, trabajo, energía degradada, sistema, entorno, frontera y universo.

La explicación se apoyó en la Figura 1, utilizada como organizador visual de la relación entre energía y transferencia de energía (trabajo y calor).

Figura 1.

Diagrama utilizado por el docente para la explicación de la LCEM



Nota. Elaboración propia del diagrama conceptual utilizado durante la fase de explicación para representar la relación entre sistema, entorno, frontera y transferencia de energía.

Esta figura permitió hacer explícita la relación entre las fuerzas conservativas, no conservativas y la LCEM, además de contrastar casos con y sin fricción y ejemplificar sistemas físicos de interés para la futura práctica docente, como sistemas de masa-

resorte, péndulos, juego de parque de diversiones, cuerpos en planos inclinados, entre otras situaciones.

En esta fase se mezclaron miniexposiciones con preguntas conceptuales y numéricas de control sobre la LCEM, como preparación para la fase de personalización con IAG.

2.4.3. Fase de aprendizaje individualizado con IAG

En la tercera fase, de 3 horas de duración, los estudiantes interactuaron con un tutor basado en Gemini, configurado para actuar como tutor experto en LCEM y en la Taxonomía de Bloom. La interacción se organizó en una secuencia de preguntas socráticas distribuidas en los niveles recordar, comprender, aplicar, analizar, evaluar y crear.

En cada nivel cognitivo de la Taxonomía de Bloom, el tutor generaba una primera pregunta principal de tipo conceptual o numérica. Si la respuesta era correcta y completa, procedía al siguiente nivel; si era incorrecta o incompleta, generaba preguntas socráticas intermedias, una por una, hasta reconstruir el razonamiento necesario para volver a plantear la pregunta principal.

De esta manera, el tutor ofrecía problemas alineados con los niveles cognitivos de la taxonomía de Bloom, ajustaba la complejidad de manera implícita dependiendo de la calidad de la respuesta que daba el estudiante y promovía la metacognición a través de justificación y verificación constante. Además, se cuidó que el lenguaje fuese claro y apropiado para el perfil del Bachillerato General Unificado (BGU) del Ecuador.

El docente observó la interacción entre el estudiante y el tutor, además revisó las conversaciones de muestra y orientó el uso crítico de la herramienta, en especial en la comprobación de los resultados, su contraste con los apuntes de la clase y la detección de posibles errores de la IAG. También se recogieron número de intentos, secuencia de niveles alcanzados y patrones de error.

El prompt diseñado, fue probado previamente con estudiantes de la misma carrera para ajustar su claridad operativa. En esta versión del manuscrito se sintetiza la lógica funcional del mismo: tutoría socrática, avance por niveles cognitivos según la taxonomía de Bloom, retroalimentación de las respuestas dadas por los estudiantes sin entrega directa de la misma y finalmente el avance al siguiente nivel según la calidad de la respuesta del estudiante.

2.4.4. Fase de evaluación

Al finalizar la intervención, se realizó un posttest equivalente al pretest. El posttest preservó correspondencia conceptual, numérica y de nivel cognitivo con la versión original del pretest, pero empleó enunciados y contextos diferentes para reducir efectos de memorización y práctica (Siantuba et al., 2023).

También se aplicó una encuesta de percepción con escala Likert sobre el uso de la IAG para aprender la LCEM, la misma consistió en 24 afirmaciones agrupadas en 6 dimensiones, entre ellas: utilidad percibida, motivación e interés, carga cognitiva, aceptabilidad ética y confianza, impacto en el aprendizaje de la LCEM y facilidad de uso y personalización. Incluyó, además, 3 preguntas abiertas sobre la utilidad de la experiencia, las dificultades encontradas y las sugerencias de mejora.

Los datos cuantitativos fueron analizados mediante estadísticos descriptivos y pruebas de comparación de medidas repetidas. Los puntajes por nivel cognitivo se analizaron mediante la prueba de Wilcoxon de rangos con signo y, previa verificación de la normalidad de las diferencias mediante la prueba de Shapiro-Wilk, se utilizó una *t* de muestras pareadas para el puntaje total. El cambio por nivel se representó además en un gráfico de bosque con las diferencias porcentuales post-pre e intervalos de confianza del 95%, calculados sobre 18 respuestas por nivel, o sea 9 estudiantes por cada 2 ítems. Las respuestas abiertas y los fragmentos de interacción con el tutor aportaron datos cualitativos que se agruparon temáticamente según utilidades percibidas, dificultades y sugerencias de mejora.

2.5. Consideraciones éticas

El estudio se desarrolló según los principios de ética relacionados a la investigación educativa en educación superior, con la aprobación correspondiente de la institución.

La participación de los estudiantes fue voluntaria y se realizó mediante consentimiento informado, donde se explicitaron los objetivos, los procedimientos, los posibles beneficios y el derecho a retirarse sin consecuencias académicas.

Los datos de los estudiantes se codificaron con identificadores alfanuméricos para mantener la confidencialidad de los mismos y los resultados se presentaron de forma anónima. Además, la información analizada se trabajó de forma independiente a los registros de calificaciones del curso y se dejó en claro que, con la participación de forma correcta, se otorgaría una bonificación de 3 puntos extras para la asignatura.

2.6. Validez interna y limitaciones

Debido al diseño de un solo grupo preexperimental, la investigación está expuesta a las amenazas tradicionales a la validez interna, como historia, maduración, efectos de prueba y fluctuaciones mínimas en la aplicación de los instrumentos. En el periodo de intervención, los estudiantes tuvieron la posibilidad de acceder a otros recursos externos, lo cual puede contribuir como una influencia sobre el aprendizaje. También es posible que parte de la mejoría observada se deba al avance natural en el contenido de la asignatura y con el formato de evaluación. De igual manera, la repetición de tareas similares como el pretest y el posttest puede propiciar efectos de mejoría aun cuando se hayan utilizado formas paralelas. Por último, las pequeñas diferencias de horario, de ambiente de aula y de explicaciones operativas en el momento de la aplicación pueden introducir pequeñas variaciones de instrumentación.

Para disminuir esas amenazas se utilizaron versiones paralelas del test, condiciones similares de aplicación, instrucciones estandarizadas y un control uniforme del uso del tutor durante las sesiones presenciales.

Respecto de la validez externa, los resultados deben leerse como evidencia situacional y preliminar debido al tamaño reducido de la muestra, el trabajo con una sola asignatura, la implementación en una única institución y el uso de una sola plataforma de IAG, ya que la muestra fue pequeña ($n = 9$) y perteneció a un solo curso, la intervención se desarrolló en una sola asignatura y en una única institución de educación superior; además, se utilizó una única plataforma de IAG, Gemini, con una configuración

específica del tutor y el periodo de intervención fue breve y se concentró en 9 horas dentro del semestre académico.

Por lo tanto, los hallazgos se ofrecen como resultados preliminares que pueden orientar estudios posteriores con muestras más amplias y diseños comparativos mejor controlados.

2.7. Prueba de entrada y salida basada en dominios de contenido y niveles cognitivos en la LCEM: Diseño y medición

Con el fin de evaluar los niveles cognitivos relacionados al aprendizaje de la LCEM se elaboró una prueba de pretest y postest de opción múltiple basada en la Taxonomía de Bloom revisada por expertos en la asignatura (Anderson & Krathwohl, 2001; Krathwohl, 2002). La prueba estuvo conformada por 12 ítems, con 2 preguntas por cada nivel cognitivo de la Taxonomía de Bloom: recordar, comprender, aplicar, analizar, evaluar y crear.

Los dominios de contenido considerados fueron: C1: Conceptos básicos de energía mecánica y LCEM; C2: Transformación y conservación de la energía en contextos sin rozamiento; y C3: Razonamiento sobre rozamiento, degradación de energía y toma de decisiones sobre procedimientos o propuestas de solución, tal como se muestra en la tabla 1. De acuerdo con esta estructura, la prueba estuvo compuesta por preguntas conceptuales, numéricas y de juicio sobre opciones de diseño y explicación, con el fin de aproximarse a cada uno de los niveles cognitivos, sin dejar de lado el formato de respuesta cerrada.

Cada ítem se elaboró con cuatro opciones de respuesta, donde una de ellas es correcta y las otras 3 opciones son erróneas y que fueron diseñadas en base a concepciones alternativas y errores conceptuales frecuentes sobre la LCEM. La organización en niveles permite describir el nivel de pensamiento que evidencia el estudiante en cada tramo del instrumento, más allá de la puntuación global.

Los reactivos fueron calificados de forma dicotómica, otorgando el valor de 1 a la respuesta correcta y el valor de 0 a la respuesta incorrecta. Los ítems fueron agrupados por pares para conformar un puntaje por nivel cognitivo posterior con rango de 0 a 2. Se consideró que cada nivel correspondía a un rango de logro: 0 de 2, bajo logro; 1 de 2, logro medio; y 2 de 2, alto logro o dominio del nivel. Esta manera de agrupar los datos permite reconocer las fortalezas y los vacíos por nivel, y permite relacionar el análisis de aprendizaje con la estructura cognitiva del instrumento. Además, el puntaje por nivel, se determinó con el nivel máximo cognitivo consolidado de cada estudiante mediante la siguiente regla: se considera que un nivel cognitivo está consolidado cuando el estudiante obtiene 2 de 2 en ese nivel y mantiene al menos 1 de 2 en todos los niveles inferiores. La combinación de dicha codificación dicotómica, por niveles y de criterio de dominio permitió construir perfiles cognitivos y describir la diferencia observada entre el pretest y el postest (Stanny, 2016).

Tabla 1.

Clasificación de la prueba por ítems, dominios de contenido y niveles cognitivos

Item	Dominio de contenido.	Nivel de taxonomía de Bloom.	Tipo de ítem.
1	C1. Conceptos básicos de energía mecánica y ley de conservación de la energía mecánica, definición de energía mecánica y reconocimiento de sus componentes	Recordar	Conceptual de definición, identifica y describe qué es la energía mecánica en un contexto físico
2	C1. Conceptos básicos de energía mecánica, energía potencial gravitatoria, relación entre masa, gravedad y altura, reconocimiento de variables y estructura de la expresión	Recordar	Conceptual de fórmula, reconoce la expresión de energía potencial gravitatoria y el significado de cada variable
3	C2. Transformación de energía en un sistema sin rozamiento durante caída libre, explicación del cambio entre energía potencial gravitatoria y energía cinética	Comprender	Conceptual de comprensión, explica con palabras el cambio de forma de energía en caída libre sin rozamiento
4	C2. Transformación de energía en movimiento vertical sin rozamiento, interpretación del punto más alto, relación entre velocidad instantánea y energía potencial gravitatoria máxima	Comprender	Conceptual de interpretación, analiza qué ocurre con la energía y la velocidad en el punto más alto del movimiento
5	C2. Aplicación de la ley de conservación de la energía mecánica en caída libre sin rozamiento, cálculo de rapidez o energía cinética a partir de una altura inicial.	Aplicar	N Numérico de aplicación, resuelve un cálculo usando conservación de energía para obtener rapidez o energía cinética al caer desde una altura dada
6	C2. Aplicación de la ley de conservación de la energía mecánica en lanzamiento vertical sin rozamiento, determinación de altura máxima a partir de la velocidad inicial.	Aplicar	N Numérico de aplicación, calcula altura máxima usando conservación de energía y despeje de la variable buscada
7	C2. Análisis comparativo de dos trayectorias alternativas sin rozamiento, consistencia de la energía mecánica a lo largo del recorrido, justificación física de diferencias o equivalencias.	Analizar	Conceptual de comparación, contrasta trayectoria A y trayectoria B y justifica con criterios de conservación de energía
8	C2. Interpretación de una tabla con valores de energía potencial gravitatoria y energía cinética, verificación de constancia de energía mecánica y detección de incoherencias.	Analizar	Mixto con datos, interpreta información tabular, calcula, verifica consistencia y argumenta conclusiones
9	C3. Efectos del rozamiento, distinción entre conservación de energía mecánica y conservación de energía total, juicio sobre validez de afirmaciones en contextos con pérdidas.	Evaluar	Conceptual de juicio, valora una afirmación sobre conservación de energía considerando rozamiento y transferencia de energía.
10	C3. Evaluación crítica de un procedimiento o explicación de solución, detección de errores conceptuales sobre energía mecánica, energía total y referencia, argumentación correctiva.	Evaluar	Conceptual de análisis de solución, revisa el razonamiento de un estudiante, identifica fallas y propone corrección justificada
11	C3. Diseño experimental para estudiar conservación de energía mecánica, planteamiento de variables, controles, mediciones, instrumentos, registro de datos y criterios de validez.	Crear	Conceptual de planificación, selecciona la mejor propuesta experimental entre alternativas cerradas.
12	C3. Diseño de una actividad didáctica sobre LCEM incorporando rozamiento, secuencia de trabajo, recursos, consignas, evidencias y evaluación.	Crear	Conceptual de diseño didáctico, selecciona la alternativa más pertinente para estructurar una actividad sobre LCEM.

Nota. Los ítems 11 y 12 mantienen formato de opción múltiple y se orientan a la selección de alternativas de planificación o diseño, no a la producción abierta de una propuesta.

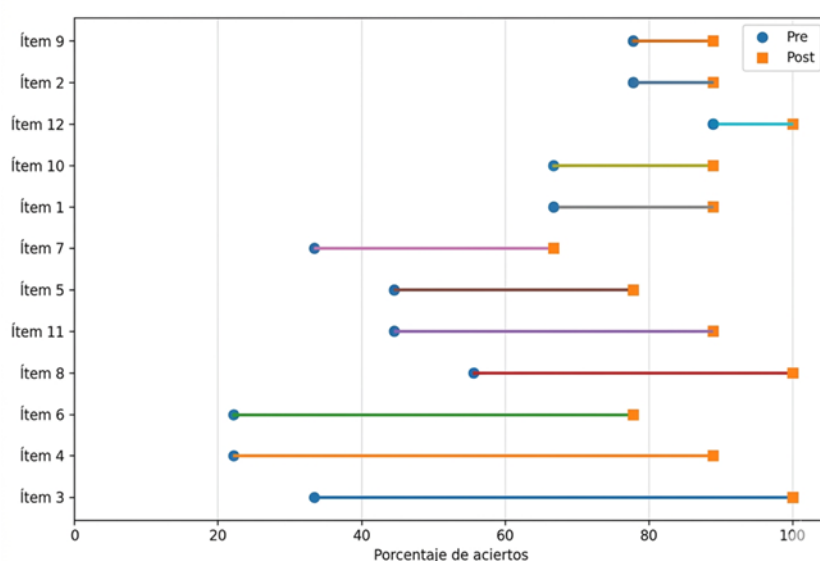
3. Resultados

En la Figura 2 se muestra el porcentaje de aciertos por ítem en el pretest y en el postest para los 12 reactivos del instrumento. Los 12 ítems muestran incrementos entre ambas mediciones y no se aprecian retrocesos. De promedio, el grupo aumentó su número de aciertos desde el 52,8 % en el pretest al 88,0 % en el postest, con una ganancia media de 35,2 puntos porcentuales. Los mayores contrastes se agruparon en los ítems 3 y 4, ambos relacionados con el nivel Comprender. El ítem 3 pasó de un 33,3 % a un 100 %

y el ítem 4 de un 22,2 % a un 88,9 %. Resalta, asimismo, el ítem 6, del nivel Aplicar, que pasó de 22,2 % a 77,8 %. El segundo grupo está conformado por los ítems 8, 11, 5, 7 y 12. El ítem 8 pasó de 55,6 % a 100 %; el ítem 11, de 44,4 % a 88,9 %; el ítem 5, de 44,4 % a 77,8 %; el ítem 7, de 33,3 % a 66,7 %; y el ítem 12, de 88,9 % a 100 %. Las diferencias más pequeñas son para los ítems que ya salían con porcentajes altos en el pretest. En Recordar, los ítems 1 y 2 aumentaron de 66,7 % y 77,8 % a 88,9 %. En Evaluar, los ítems 9 y 10 subieron de 77,8 % y 66,7 % a 88,9 %, respectivamente. Al final de la medición, tres ítems alcanzaron un 100 % de aciertos y seis se situaron en 88,9 %.

Figura 2.

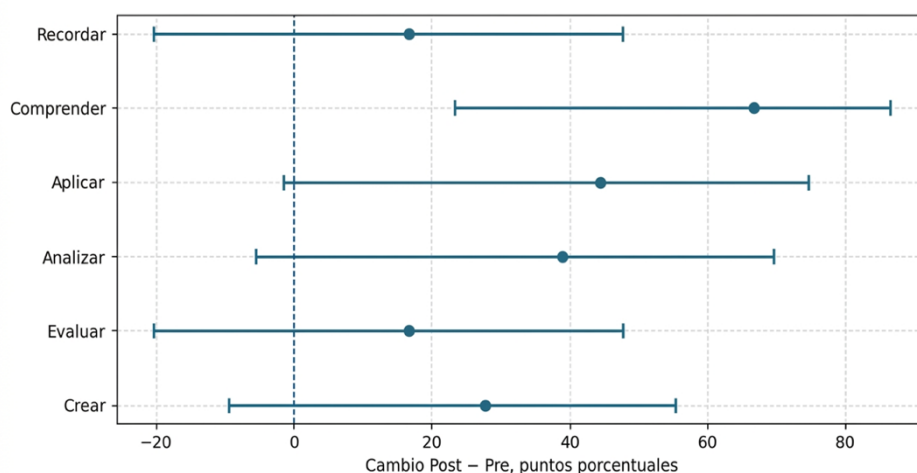
Porcentaje de aciertos por ítem en pretest y postest, ordenado por magnitud de mejora



Nota. Elaboración propia con base en los porcentajes de aciertos obtenidos en el pretest y el postest para los 12 ítems del instrumento (n = 9).

La Figura 3 resume el cambio porcentual post-pre por nivel cognitivo, incorporando intervalos de confianza del 95 %. Todas las estimaciones puntuales están por encima de cero. El cambio puntual más elevado lo aporta entender, con 66,67 puntos porcentuales. Los niveles a aplicar y analizar presentan diferencias de 44,44 y 38,89 puntos porcentuales, respectivamente. El nivel crear presenta una diferencia de 27,78 puntos porcentuales, mientras que Recordar y Evaluar tienen una diferencia de 16,67 puntos porcentuales cada uno. Al entenderlo, el intervalo de confianza permanece por encima de cero. En los restantes niveles, los intervalos son más amplios e incluyen o rozan el cero.

Figura 3.

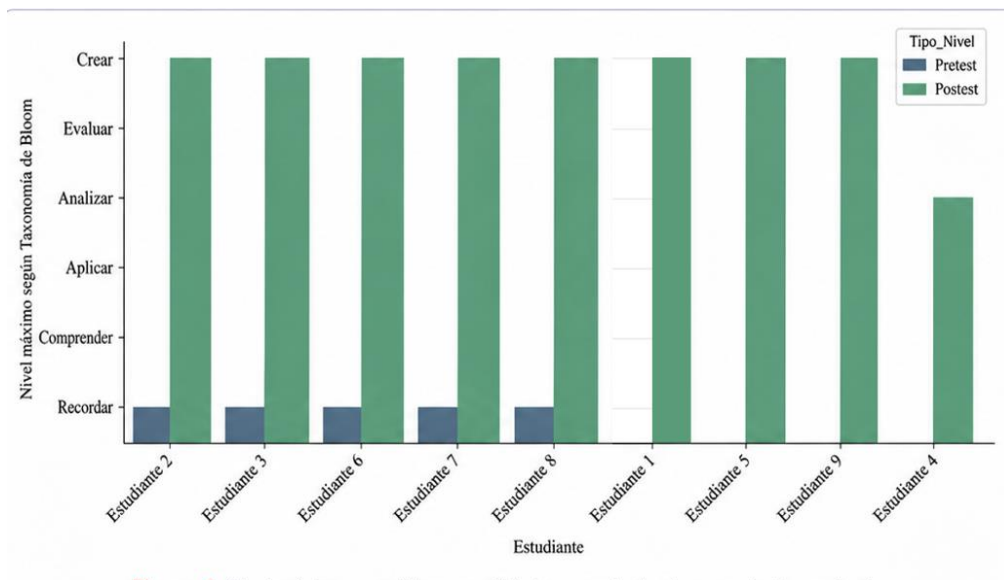
Cambio porcentual post-pre por nivel cognitivo con intervalos de confianza del 95

IC 95% por Newcombe-Wilson para diferencia de proporciones. Base: 18 respuestas por nivel

Nota. Elaboración propia con base en las diferencias porcentuales posttest-pretest por nivel cognitivo; los intervalos de confianza del 95 % se calcularon sobre 18 respuestas por nivel.%.

En la Figura 4 se ilustra el nivel máximo cognitivo alcanzado por cada estudiante en el pretest y en el posttest. En el pretest, 5 estudiantes alcanzaron a recordar como nivel máximo y 4 no cumplieron el umbral mínimo para consolidar un nivel. En el posttest, ocho estudiantes alcanzaron como nivel máximo crear y uno llegó a analizar.

Figura 4.

Nivel máximo cognitivo consolidado por estudiante en pretest y posttest

Nota. Elaboración propia con base en el nivel máximo cognitivo consolidado por estudiante en el pretest y el posttest (n = 9).

3.1. Análisis por estudiante y por nivel cognitivo

A cada estudiante se le calculó un puntaje por nivel cognitivo en el pretest y en el posttest. Como cada nivel se operacionalizó con dos ítems, el puntaje por nivel varió de 0 a 2. Debido al carácter discreto de estos puntajes y al escaso tamaño muestral, se empleó la prueba de rangos con signo de Wilcoxon para muestras apareadas para las comparaciones por nivel. Para cada nivel, la hipótesis nula estableció que la mediana de las diferencias posttest-pretest era igual a 0; la hipótesis alternativa estableció que la mediana fue distinta de 0. Para el RIC, que significa rango intercuartílico, se reporta la correlación biserial por rangos, r biserial, como medida del tamaño del efecto en la prueba de Wilcoxon.

Tabla 2.

Comparación de los rangos obtenidos en los niveles cognitivos entre el pretest y posttest (n = 9)

Nivel	Pretest Mdn y RIC	Posttest Mdn y RIC	Mejoran; Igual; Retroceden	Prueba	p valor	Tamaño de efecto
Recordar	2 ; 1 a 2	2 ; 2 a 2	2 ; 7 ; 0	Wilcoxon	0,158	r biserial 1,00
Comprender	0 ; 0 a 1	2 ; 2 a 2	7 ; 2 ; 0	Wilcoxon	0,011	r biserial 1,00
Aplicar	1 ; 0 a 1	2 ; 1 a 2	7 ; 2 ; 0	Wilcoxon	0,01	r biserial 1,00
Analizar	1 ; 0 a 2	2 ; 1 a 2	5 ; 3 ; 1	Wilcoxon	0,076	r biserial 0,81
Evaluar	2 ; 1 a 2	2 ; 2 a 2	4 ; 3 ; 2	Wilcoxon	0,353	r biserial 0,43
Crear	1 ; 1 a 2	2 ; 2 a 2	4 ; 5 ; 0	Wilcoxon	0,047	r biserial 1,00

Nota. Mdn = Mediana; RIC = Rango Intercuartílico. La columna "Dirección del cambio" representa el conteo de estudiantes que mejoran, se mantienen igual o retroceden, respectivamente. r biserial = coeficiente de correlación rango-biserial. $p < 0,050$ indica una diferencia estadísticamente significativa.

En la Tabla 2 se observa que comprender pasó de una mediana de 0 a 2, con 7 estudiantes que mejoraron y ningún retroceso ($p = 0,011$). Aplicar pasó de 1 a 2, también con siete mejoras y ningún retroceso ($p = 0,010$). Crear pasó de 1 a 2, con cuatro mejoras y ningún retroceso ($p = 0,047$). Analizar pasó de 1 a 2, con cinco mejoras, tres igualdades y un retroceso ($p = 0,076$). Recordar y Evaluar mantuvieron una mediana de 2, con predominio de igualdades.

Para el puntaje total, compuesto por 12 ítems, se evaluó la normalidad de las diferencias entre posttest y pretest mediante Shapiro-Wilk. El valor $p = 0,687$ indicó que no había evidencia para rechazar la normalidad de las diferencias, por lo que se aplicó una t pareada.

La hipótesis nula para el puntaje total estableció que la media de las diferencias posttest-pretest era igual a 0; la hipótesis alternativa planteó que era distinta de 0.

Tabla 3.

Análisis de diferencia de medias (Prueba t de Student) para muestras relacionadas (n = 9)

n	Pre media	Pre DE	Post media	Post DE	Δ media	IC95% Δ	Shapiro p (Δ)	t	p	dz
9	6.33	0.87	10.56	1.01	4.22	3.08 a 5.36	0.687	8.55	< 0.001	2.85

Nota. M = Media; DE = Desviación estándar; IC 95% = Intervalo de confianza del 95%; Shapiro-Wilk = valor-p de la prueba de normalidad de Shapiro-Wilk aplicada a las diferencias de los puntajes. dz = Tamaño del efecto de Cohen para muestras pareadas. Un valor de $p < 0,05$ se considera estadísticamente significativo.

La Tabla 3 muestra que la media pasó de $6,33 \pm 0,87$ en el pretest a $10,56 \pm 1,01$ en el posttest. La diferencia media fue de 4,22 ítems, con un intervalo de confianza del 95 % entre 3,08 y 5,36. La prueba t pareada arrojó $t = 8,55$ y $p < 0,001$; el tamaño de efecto reportado fue $dz = 2,85$.

3.2. Percepción de futuros docentes sobre la IAG y personalización del aprendizaje

La percepción sobre el uso de la IAG como apoyo al aprendizaje personalizado de la LCEM se estimó con una encuesta aplicada a los 9 participantes, conformada por 24 afirmaciones en escala Likert de 1 a 5 y tres preguntas abiertas. Para cada dimensión se calcularon mediana, RIC, media, desviación estándar (D) y porcentaje de respuestas altas que corresponden a 4 o 5 en la escala Likert.

Tabla 4.

Resumen descriptivo de los resultados por dimensiones del instrumento (n = 9)

Dimensión.	Ítems.	Mediana; RIC	Media $\pm$ DE	Respuestas 4 o 5, porcentaje.
Utilidad percibida.	4	5,00; (4,75 a 5,00)	$4,83 \pm 0,28$	100
Motivación e interés.	4	5,00; (4,75 a 5,00)	$4,83 \pm 0,22$	100
Impacto en el aprendizaje de la LCEM.	4	4,75; (4,50 a 5,00)	$4,64 \pm 0,42$	88,9
Facilidad de uso y personalización	4	4,25; (4,25 a 5,00)	$4,47 \pm 0,54$	77,8
Aceptabilidad ética y confianza	4	4,00; (3,75 a 4,50)	$4,06 \pm 0,37$	66,7
Carga cognitiva	4	2,25; (1,75 a 2,75)	$2,39 \pm 0,99$	11,1

Nota. Mdn = Mediana; RIC = Rango Intercuartílico; M = Media; DE = Desviación Estándar. La columna "Respuestas Altas (%)" representa la frecuencia relativa acumulada de las valoraciones 4 ("De acuerdo") y 5 ("Totalmente de acuerdo") en la escala Likert utilizada.

Las dimensiones con valores centrales más altos fueron utilidad percibida y motivación e interés, ambas con mediana 5,00 y 100 % de respuestas altas. El impacto percibido en el aprendizaje de la LCEM alcanzó una mediana de 4,75 y la facilidad de uso y personalización, una mediana de 4,25. La aceptabilidad ética y confianza registró una mediana de 4,00. La carga cognitiva mostró una mediana de 2,25.

En la personalización del aprendizaje, los 2 ítems referidos al ajuste del nivel de explicación registraron mediana 5 y 100 % de respuestas 4 o 5. En carga cognitiva se observó mayor dispersión. Las respuestas abiertas se organizaron en 3 temas: utilidad

del andamiaje socrático, dificultades puntuales durante la interacción y sugerencias de mejora centradas en dosificar preguntas, usar mini preguntas, incorporar autoevaluación al cierre y extender la estrategia a otras asignaturas.

4. Discusión

Los resultados de este estudio piloto demuestran que la intervención de aprendizaje personalizado con IAG, estructurada por niveles cognitivos de la taxonomía de Bloom, está asociada con mejoras sustantivas en el desempeño sobre la LCEM, en futuros docentes de Física (Guerra-Reyes et al., 2024; Halilović et al., 2021). De promedio el grupo pasó de un 52,8% de aciertos en el pretest a un 88,0% en el postest, con incrementos especialmente importantes en los niveles comprender, aplicar y crear, lo que sugiere un tránsito desde un dominio principalmente reproductivo hacia formas de comprensión más profundas y capacidades de uso aplicado y diseño didáctico (Anderson & Krathwohl, 2001; Krathwohl, 2002). Estos resultados coinciden con investigaciones anteriores que sugieren que la IAG puede favorecer la personalización del aprendizaje y elevar el desempeño académico cuando se incorpora propuestas pedagógicas estructuradas y con objetivos claros (Ayeni et al., 2024; Ogunleye et al., 2024).

Desde el punto de vista del entorno educativo, la secuencia DEAPE contribuyó a la construcción de un contexto híbrido que articula exposición docente, resolución de problemas, laboratorio y tutoría generativa en un aula universitaria pública ecuatoriana (Castillo Villegas et al., 2025). Este ambiente combinó recursos tradicionales con acceso homogéneo a un tutor generativo a través de dispositivos institucionales, lo que permitió reducir barreras de acceso y mantener condiciones relativamente equitativas entre los estudiantes (Valdivieso & González, 2025). La literatura sobre personalización del aprendizaje señala que la efectividad de las tecnologías inteligentes no depende sólo de los algoritmos, sino también de la forma en que se insertan en ambientes educativos que facilitan la participación activa, la autorregulación y la interacción significativa entre estudiantes, docente y contenidos (Barrios-Pérez & Amaya-Melgar, 2025; Rivera Viñan et al., 2025).

La parte de tutoría socrática implementada con IAG parece haber sido una pieza clave para la configuración de este ambiente de aprendizaje. El tutor generativo planteó secuencias de preguntas graduadas según el nivel cognitivo, ajustando la complejidad en función de las respuestas de cada estudiante, evitando la entrega directa de soluciones y fomentando la explicación argumentada (Fakour & Imani, 2025). Esta dinámica se alinea con investigaciones que evidencian el potencial de tutores basados en modelos de lenguaje para sostener procesos de explicación, práctica guiada y retroalimentación inmediata, siempre y cuando se diseñen como sistemas de andamiaje y no de sustitución del docente (Kestin et al., 2025; Memon & Kwan, 2025; Tu et al., 2025). Los participantes reportaron altos valores de utilidad percibida y motivación, lo que indica que, en este ambiente, la interacción con el tutor se percibió como un apoyo relevante para enfrentar un contenido de alta densidad conceptual como la LCEM (Siantuba et al., 2023).

El contexto latinoamericano y, en particular, la realidad ecuatoriana, brindan importantes matices a la interpretación de estos resultados. Diversos trabajos han señalado limitaciones de infraestructura, heterogeneidad metodológica y dificultades para

incorporar tecnologías emergentes en entornos reales de aula en la región (Ade Nandang Mustafa, 2025; Barrera & Azeez, 2024). En este estudio la experiencia se desarrolló en una institución pública con recursos acotados pero gestionados de manera organizada, lo que evidencia que es posible diseñar ambientes de aprendizaje personalizados con IAG en condiciones de recursos moderados cuando se planifica cuidadosamente el acceso y se acompaña al profesorado en el uso pedagógico de la tecnología (Castillo Villegas et al., 2025; Valdivieso & González, 2025). Sin embargo, el establecimiento de este tipo de entornos requiere políticas institucionales más generales, financiamiento sostenido y programas de formación docente que atiendan tanto dimensiones técnicas como éticas (García Ramos & Wilson-Kennedy, 2024; Kalniņa et al., 2024).

Los resultados de percepción de la encuesta ayudan a comprender cómo se vivió el ambiente de aprendizaje de los estudiantes. Las dimensiones de utilidad percibida y motivación e interés alcanzaron medianas máximas y un 100% de respuestas altas, mientras que la dimensión de carga cognitiva se situó en valores moderados (Siantuba et al., 2023). Estos resultados reflejan un equilibrio razonable entre el desafío y el apoyo, coherente con investigaciones anteriores, en la que los entornos adaptativos mediante IAG pueden aumentar el compromiso y la autorregulación al ajustar el nivel de dificultad a la zona de desarrollo del estudiante (Barrios-Pérez & Amaya-Melgar, 2025; Rivera Viñan et al., 2025). En cambio, las respuestas abiertas, resaltan la utilidad del andamiaje, recomiendan dosificar la cantidad de interrogantes e incorporar momentos explícitos de autoevaluación, para seguir afinando el diseño para que la personalización se traduzca en oportunidades de reflexión, y no de sobrecarga cognitiva (Stanny, 2016).

Al mismo tiempo, los resultados deben enfocarse dentro de las limitaciones del diseño metodológico. Se empleó un diseño preexperimental de un solo grupo, con una muestra pequeña ($n = 9$), por lo que las amenazas a la validez interna, como historia, maduración, práctica con el formato de prueba y otras influencias externas no pueden descartarse por completo (Hernández-Sampieri & Mendoza, 2023). A pesar de emplearse formas paralelas del instrumento, condiciones similares de aplicación e instrucciones estandarizadas, parte de la mejoría observada podría estar relacionada con el progreso natural en la materia y la repetición de tareas conceptualmente similares (Ghasemi & Zahediasl, 2012; Kim, 2015). La gran amplitud de algunos intervalos de confianza por nivel cognitivo y los elevadísimos tamaños del efecto refuerzan la recomendación de replicar el estudio con muestras mayores y, de ser posible, diseños cuasiexperimentales o experimentales que incorporen grupos de comparación (Araya-Muñoz et al., 2019).

También cabe destacar la dimensión ética del ambiente educativo mediado por IAG. Diferentes contribuciones ya han puesto sobre la mesa riesgos asociados a la opacidad de los modelos, a la gestión de datos y a la reproducción de sesgos en contextos educativos, sobre todo cuando los sistemas empiezan a asumir funciones tutoriales (Holmes et al., 2022; Ogunleye et al., 2024). En esta investigación, la participación continua del docente, la revisión de las interacciones y la invitación a comparar las respuestas de la IAG con los apuntes y criterios de la asignatura contribuyeron a enmarcar la experiencia dentro de una perspectiva crítica y responsable (García Ramos & Wilson-Kennedy, 2024; Kalniņa et al., 2024). Sin embargo, la posible expansión de estas herramientas exige reforzar la formación ética y digital de los futuros docentes,

para que puedan incorporar la IAG en sus aulas de forma reflexiva, transparente y orientada a la equidad (Memon & Kwan, 2025; Valdivieso & González, 2025).

Por último, esta investigación aporta elementos a la discusión en torno al diseño de ambientes educativos que integran IAG y evaluaciones por niveles cognitivos. El diseño de una prueba pretest y postes alineada con la taxonomía de Bloom, la determinación de niveles máximos consolidados y la combinación de análisis cuantitativos y cualitativos permitieron caracterizar con mayor resolución el perfil de logro de cada estudiante, antes y después de la intervención (Almatrafi & Johri, 2025; Anderson & Krathwohl, 2001). Esta postura se alinea con aquellas que subrayan la necesidad de conectar la personalización del aprendizaje con marcos curriculares claros, criterios de evaluación transparentes y espacios que permitan la participación activa del estudiantado (Barrios-Pérez & Amaya-Melgar, 2025; Rivera Viñan et al., 2025). En contextos como Ecuador, donde se plantean estrategias para mejorar la enseñanza de la Física y reducir brechas de aprendizaje, experiencias piloto como esta, pueden funcionar como laboratorio para ajustar los diseños didácticos, instrumentos e infraestructuras hacia entornos de aprendizaje más inclusivos y personalizados (Barrera & Azeez, 2024; Rodríguez Castillo et al., 2025).

5. Conclusiones

La primera conclusión es que una intervención de aprendizaje personalizado con IAG, organizada por niveles cognitivos, y situada en un ambiente universitario público, se asocia con mejoras importantes en el desempeño conceptual y aplicado sobre la LCEM de futuros docentes de Física. El aumento en el porcentaje de aciertos, las diferencias encontradas en comprender, aplicar y crear y el cambio en el nivel máximo cognitivo alcanzado desde recordar hacia niveles superiores sugieren un avance hacia formas de aprendizaje más profundas y transferibles.

En segundo lugar, los participantes evaluaron positivamente la estrategia educativa dada por la secuencia DEAPE, que integra explicación docente, actividades prácticas, laboratorio y tutoría generativa, en términos de utilidad, motivación e impacto percibido en su aprendizaje. La combinación de recursos tradicionales y de las tecnologías de IAG, en condiciones relativamente equitativas de acceso, permitió crear un entorno de trabajo exigente pero manejable, donde el diálogo con el tutor socrático y el acompañamiento continuo del docente, parecen haber favorecido la reflexión.

En tercer lugar, esta investigación revela que es posible llevar a cabo experiencias de personalización con IAG usando Gemini en el contexto universitario latinoamericano con recursos limitados, siempre que se planifiquen con cuidado las condiciones de acceso, se ajuste el diseño instruccional al contexto y se conserve la supervisión continua del docente durante el uso de las herramientas. Sin embargo, para que estas experiencias se generalicen, se necesitan políticas institucionales que fortalezcan la infraestructura, la formación digital y los marcos éticos para el uso de IA en la educación.

En cuarto lugar, debido a las características del diseño preexperimental, con un solo grupo y tamaño muestral reducido, los resultados deben ser tomados como evidencias preliminares. Para llegar a conclusiones más robustas, se requieren estudios con muestras más amplias, la comparación con un grupo de control y un análisis más

profundo de las interacciones y percepciones de los estudiantes, consideran con conjuntamente métodos cuantitativos y cualitativos.

Finalmente, es importante que los futuros docentes se formen en el dominio conceptual de la Física, así como en el uso pedagógico, crítico y ético de herramientas de IAG como Gemini. En un escenario donde estas herramientas tenderán a integrarse en el aula, es fundamental que los docentes en formación aprendan a diseñar secuencias didácticas apoyadas en IAG, a interpretar los datos que se generan, y a garantizar que la personalización del aprendizaje se oriente a la equidad y el desarrollo del pensamiento crítico.

Referencias Bibliográficas

- Ade Nandang Mustafa. (2025). Reimagining education in the age of artificial intelligence: Challenges and strategies for developing nations. *International Journal of Science and Research Archive*, 16(2), 291–298. <https://doi.org/10.30574/ijrsra.2025.16.2.2029>
- Aljuraiban, G. S., Gibson, R., Chan, D. S. M., Van Horn, L., & Chan, Q. (2025). Leveraging generative AI for course learning outcome categorization using Bloom's taxonomy. *Computers and Education: Artificial Intelligence*, 8, 100404. <https://doi.org/10.1016/j.caeai.2025.100404>
- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Araya-Muñoz, I., Fonseca-Argüello, M., Majano-Benavides, J., & Ugalde-Villalobos, M. E. (2019). Evaluation of the design and didactic development of three blended learning subjects: A pilot plan of the Faculty of Social Sciences at the Universidad Nacional, Costa Rica. *Revista Electrónica Educare*, 23(1), 1–16. <https://doi.org/10.15359/ree.23-1.6>
- Ayeni, O. O., Hamad, N. M. A., Chisom, O. N., Osawaru, B., & Adewusi, O. E. (2024). AI in education: A review of personalized learning and educational technology. *GSC Advanced Research and Reviews*, 18(2), 261–271. <https://doi.org/10.30574/gscarr.2024.18.2.0062>
- Barrera, A. G., & Azeez, G. K. (2024). Artificial intelligence in higher education: Lessons from Chile and Mexico. *International Higher Education*, (121), 24–25.
- Barrios-Pérez, E. G., & Amaya-Melgar, S. (2025). Aprendizaje personalizado y entornos personales de aprendizaje en la era de la inteligencia artificial: Apuntes teórico-conceptuales. *Emerging Trends in Education*, 8(15), 105–121. <https://doi.org/10.19136/etie.v8n15.6341>
- Castillo Villegas, K. G., Villagómez Samaniego, E. P., & Chuquin Cuatis, M. R. M. (2025). Educación basada en competencias: Diseño, implementación y evaluación en programas de educación superior. *Revista InveCom*, 5(3), e050342. <https://doi.org/10.5281/zenodo.14270813>
- Fakour, H., & Imani, M. (2025). Socratic wisdom in the age of AI: A comparative study of ChatGPT and human tutors in enhancing critical thinking skills. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1528603>

- Garcia Ramos, J., & Wilson-Kennedy, Z. (2024). Promoting equity and addressing concerns in teaching and learning with artificial intelligence. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1487882>
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Guerra-Reyes, F., Guerra-Dávila, E., Naranjo-Toro, M., Basantes-Andrade, A., & Guevara-Betancourt, S. (2024). Misconceptions in the learning of natural sciences: A systematic review. *Education Sciences*, 14(5), 497. <https://doi.org/10.3390/educsci14050497>
- Halilović, A., Mešić, V., Hasović, E., & Dervić, D. (2021). Students' difficulties in applying the law of conservation of mechanical energy: Results of a survey research. *The European Educational Researcher*, 4(2), 171–192. <https://doi.org/10.31757/euer.423>
- Hernández-Sampieri, R., & Mendoza, C. (2023). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta* (7.ª ed.). McGraw-Hill.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, 32(3), 504–526. <https://doi.org/10.1007/s40593-021-00239-1>
- Kalniņa, D., Nīmanīte, D., & Baranova, S. (2024). Artificial intelligence for higher education: Benefits and challenges for pre-service teachers. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1501819>
- Kestin, G., Miller, K., Klaes, A., Milbourne, T., & Ponti, G. (2025). AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. *Scientific Reports*, 15(1), 17458. <https://doi.org/10.1038/s41598-025-97652-6>
- Kim, T. K. (2015). T test as a parametric statistic. *Korean Journal of Anesthesiology*, 68(6), 540–546. <https://doi.org/10.4097/kjae.2015.68.6.540>
- Kotsis, K. T., & Panagou, D. (2023). The determination of the learning curve on the concept of energy using the alternatives ideas. *Contemporary Mathematics and Science Education*, 4(1), ep23011. <https://doi.org/10.30935/conmaths/13022>
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice*, 41(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2
- Memon, T. D., & Kwan, P. (2025). A collaborative model for integrating teacher and GenAI into future education. *TechTrends*, 69(6), 1192–1206. <https://doi.org/10.1007/s11528-025-01105-w>
- Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, 350, h1258. <https://doi.org/10.1136/bmj.h1258>

- Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). A systematic review of generative AI for teaching and learning practice. *Education Sciences*, 14(6), 636. <https://doi.org/10.3390/educsci14060636>
- Prather, J., Leinonen, J., Kiesler, N., Gorson Benario, J., Lau, S., MacNeil, S., Norouzi, N., Opel, S., Pettit, V., Porter, L., Reeves, B. N., Savelka, J., Smith, D. H., Strickroth, S., & Zingaro, D. (2025). Beyond the hype: A comprehensive review of current trends in generative AI research, teaching practices, and tools. *2024 Working Group Reports on Innovation and Technology in Computer Science Education*, 300–338. <https://doi.org/10.1145/3689187.3709614>
- Rivera Viñan, L. M., Romero Ramón, C. A., Montero Ortiz, M. A., Erique Campoverde, L. J., & Erique Campoverde, J. A. (2025). Estrategias de aprendizaje personalizado: Adaptando la enseñanza en el aula a las diversas necesidades y estilos de aprendizaje de los estudiantes. *Arandu UTIC*, 12(3), 3460–3474. <https://doi.org/10.69639/arandu.v12i3.1568>
- Rodríguez Castillo, N. E., Posada González, J., Amores Campaña, F. A., Yacchirema Jiménez, M. G., Yacchirema Jiménez, S. A., & Balseca Guzmán, C. E. (2025). University Education 5.0: Artificial intelligence and emerging technologies for innovation in higher education. *Seminars in Medical Writing and Education*, 4, 433. <https://doi.org/10.56294/mw2025433>
- Siantuba, J., Nkhata, L., & De Jong, T. (2023). The impact of an online inquiry-based learning environment addressing misconceptions on students' performance. *Smart Learning Environments*, 10(1), 22. <https://doi.org/10.1186/s40561-023-00236-y>
- Stanny, C. J. (2016). Reevaluating Bloom's taxonomy: What measurable verbs can and cannot say about student learning. *Education Sciences*, 6(4), 37. <https://doi.org/10.3390/educsci6040037>
- Tong, D., Liu, J., Sun, Y., Liu, Q., Zhang, X., Pan, S., & Bao, L. (2023). Assessment of student knowledge integration in learning work and mechanical energy. *Physical Review Physics Education Research*, 19(1), 010127. <https://doi.org/10.1103/PhysRevPhysEducRes.19.010127>
- Tu, Y., Chen, J., & Huang, C. (2025). Empowering personalized learning with generative artificial intelligence: Mechanisms, challenges and pathways. *Frontiers of Digital Education*, 2(2), 19. <https://doi.org/10.1007/s44366-025-0056-9>
- Valdivieso, T., & González, O. (2025). Generative AI tools in Salvadoran higher education: Balancing equity, ethics, and knowledge management in the Global South. *Education Sciences*, 15(2), 214. <https://doi.org/10.3390/educsci15020214>

CONFLICTO DE INTERESES

“Los autores declaran no tener ningún conflicto de intereses”.