

**INTEGRACIÓN DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN EL
APRENDIZAJE DE PROGRAMACIÓN EN BACHILLERATO
INTEGRATION OF GENERATIVE ARTIFICIAL INTELLIGENCE IN PROGRAMMING
LEARNING IN HIGH SCHOOL**

Autores: ¹Ángel Leandro Alcívar Cedeño, ²Julio Ernesto Ramírez y ³Steven Arturo Torres Burgos.

¹ORCID ID: <https://orcid.org/0009-0008-0759-3116>

²ORCID ID: <https://orcid.org/0000-0002-1064-3275>

³ORCID ID: <https://orcid.org/0000-0001-9299-3254>

¹E-mail de contacto: aalcivarc13@unemi.edu.ec

²E-mail de contacto: jramirezr23@unemi.edu.ec

³E-mail de contacto: storresb5@unemi.edu.ec

Afiliación: ^{1*2*3*}Universidad Estatal de Milagro, (Ecuador).

Artículo recibido: 10 de Agosto del 2026.

Artículo revisado: 12 de Agosto del 2026.

Artículo aprobado: 12 de Agosto del 2026.

¹Ingeniero en Sistemas, graduado de la Universidad Laica Eloy Alfaro de Manabí, (Ecuador).

²Licenciado en Educación Básica con énfasis en Ciencias Naturales y Educación Ambiental de la UNAC, (Ecuador). Corporación Universitaria Adventista, Medellín, Antioquia, (Colombia).

³Licenciado en Cultura Física, graduado de la Universidad de Guayaquil, (Ecuador). Magíster en Pedagogía de la Cultura Física mención Educación Física Inclusiva. Doctor en Educación Física por CECEIX, (México).

Resumen

La irrupción de la inteligencia artificial generativa (IAG) en los entornos escolares ha planteado interrogantes metodológicas que la investigación empírica no ha resuelto todavía con suficiencia en el nivel de bachillerato. El presente estudio examino el efecto de la integración sistemática de herramientas de IAG, concretamente ChatGPT y GitHub Copilot, sobre el rendimiento en programación, el pensamiento computacional, la autoeficacia y la motivación intrínseca de estudiantes de segundo año de Bachillerato General Unificado de unidades educativas particulares de Manta, Ecuador. Se adoptó un diseño cuasiexperimental con grupos intactos, pretest y posttest, distribuidos en un grupo experimental ($n = 44$) y un grupo de control ($n = 42$). Durante ocho semanas, el grupo experimental recibió una propuesta pedagógica que integraba las herramientas de IAG en la enseñanza de Python y algoritmia; el grupo de control trabajó bajo la modalidad instruccional convencional. Los instrumentos incluyeron una prueba objetiva de programación (POPy-25), la Escala de Autoeficacia en Programación Computacional (APC-16) y el Inventario de Motivación hacia las TIC (IMT-18). Los resultados del análisis de covarianza (ANCOVA) evidenciaron que el

grupo experimental obtuvo puntuaciones significativamente mayores en el posttest de programación y en las dimensiones de autoeficacia y motivación. El tamaño del efecto resultó grande para las tres variables dependientes. Las implicaciones para la práctica docente en contextos latinoamericanos de educación secundaria privada se analizan a la luz de los hallazgos y de la literatura internacional reciente.

Palabras clave: Inteligencia artificial generativa, Programación, Bachillerato, Pensamiento computacional, Autoeficacia.

Abstract

The emergence of generative artificial intelligence (GenAI) in school environments has raised methodological questions that empirical research has not yet sufficiently answered at the high school level. This study examined the effect of the systematic integration of GenAI tools, specifically ChatGPT and GitHub Copilot, on programming achievement, computational thinking, self-efficacy, and intrinsic motivation of second-year Bachillerato General Unificado students from private educational institutions in Manta, Ecuador. A quasi-experimental design with intact groups, pretest and posttest, was adopted,

with an experimental group ($n = 44$) and a control group ($n = 42$). Over eight weeks, the experimental group received a pedagogical proposal integrating GenAI tools in the teaching of Python and algorithms; the control group worked under the conventional instructional modality. Instruments included an objective programming test (POPy-25), the Computational Programming Self-Efficacy Scale (APC-16), and the Motivation Inventory toward ICT (IMT-18). ANCOVA results showed that the experimental group obtained significantly higher posttest scores in programming and in the self-efficacy and motivation dimensions. Effect sizes were large for all three dependent variables. Implications for teaching practice in Latin American private secondary education contexts are analyzed in light of the findings and recent international literature.

Keywords: Generative artificial intelligence, Programming, High school, Computational thinking, Self-efficacy.

Sumário

O surgimento da inteligência artificial generativa (IAG) no ambiente escolar tem suscitado questões metodológicas que a pesquisa empírica ainda não conseguiu responder de forma suficiente no nível do ensino médio. Este estudo examinou o efeito da integração sistemática de ferramentas de IAG especificamente, ChatGPT e GitHub Copilot sobre o desempenho em programação, o pensamento computacional, a autoeficácia e a motivação intrínseca de alunos do segundo ano do ensino médio (*Bachillerato General Unificado*) de escolas particulares de Manta, Equador. Adotou-se um delineamento quase-experimental com grupos intactos, incluindo pré-testes e pós-testes, com os alunos distribuídos em um grupo experimental ($n = 44$) e um grupo de controle ($n = 42$). Durante um período de oito semanas, o grupo experimental seguiu uma abordagem pedagógica que integrava ferramentas de IAG ao ensino de Python e algoritmos, enquanto o grupo de controle seguiu um método de instrução convencional. Os instrumentos utilizados

incluíram um teste objetivo de programação (POPy-25), a Escala de Autoeficácia em Programação de Computadores (APC-16) e o Inventário de Motivação em TIC (IMT-18). Os resultados da análise de covariância (ANCOVA) mostraram que o grupo experimental obteve pontuações significativamente mais altas no pós-teste de programação e nas dimensões de autoeficácia e motivação. Os tamanhos de efeito foram grandes para as três variáveis dependentes. As implicações para a prática docente em contextos de ensino médio privado na América Latina são analisadas à luz desses achados e da literatura internacional recente.

Palavras-chave: Inteligência artificial generativa, Programação, Ensino médio, Pensamento computacional, Autoeficácia.

Introducción

La enseñanza de la programación en la educación secundaria ha transitado, en las últimas dos décadas, desde un espacio curricular residual hasta convertirse en competencia básica en numerosos sistemas educativos iberoamericanos. Ecuador no es una excepción, el Currículo Nacional vigente para el Bachillerato General Unificado incorpora la asignatura de Tecnologías de la Información y Comunicación como eje obligatorio, con contenidos que incluyen fundamentos de lógica computacional y escritura de código en lenguajes de propósito general. Sin embargo, la distancia entre las prescripciones curriculares y las prácticas efectivas de aula sigue siendo considerable, sobre todo en regiones con acceso tecnológico heterogéneo como la costa ecuatoriana.

En ese escenario, la inteligencia artificial generativa ha irrumpido como un elemento disruptivo que no había sido previsto en los diseños curriculares vigentes. Desde noviembre de 2022, cuando OpenAI lanzó ChatGPT al público general, la conversación sobre el rol de

la IA en educación se aceleró de manera inusitada. García et al. (2024) analizaron de forma sistemática la producción científica surgida en los primeros seis meses tras ese lanzamiento y concluyeron que el impacto de estas tecnologías en los procesos de enseñanza y aprendizaje es ya innegable, aunque advirtieron que las posiciones extremas, tanto las que magnifican sus bondades como las que auguran consecuencias catastróficas para la integridad académica, carecen de sustento empírico suficiente. Lo que la evidencia va consolidando, con mayor precisión, es que el efecto de la IAG depende menos de la herramienta en sí y más de como el docente la integra en la propuesta pedagógica.

En el ámbito específico de la programación informática, los estudios con mayor impacto provienen de contextos universitarios anglosajones. Denny et al. (2024), en su revisión publicada en *Communications of the ACM*, documentaron que los modelos generativos de código pueden resolver ejercicios de nivel introductorio con porcentajes de acierto que superan a los de muchos estudiantes, lo que pone en cuestión los enfoques evaluativos tradicionales. En forma complementaria, Yilmaz y Karaoglan (2023) condujeron un estudio experimental con universitarios turcos y demostraron que el uso guiado de ChatGPT durante el aprendizaje de programación produce incrementos significativos en el pensamiento computacional, la autoeficacia y la motivación.

Su trabajo es, probablemente, el referente experimental más directo que existe hasta la fecha sobre el problema que el presente estudio aborda en un nivel educativo distinto. Jauhiainen y Garagorry (2024) se aproximaron al nivel escolar y mostraron, en un experimento con 110 estudiantes de primaria finlandeses,

que el material de aprendizaje elaborado con ChatGPT puede segmentar a los alumnos en grupos de conocimiento según su desempeño en tiempo real, lo que incrementa la motivación y la efectividad del aprendizaje. No obstante, la propia revisión sistemática de los autores, publicada en 2025, concluye que los estudios sobre IAG en educación K-12, es decir, en la educación primaria y secundaria, representan todavía una minoría dentro de la producción científica disponible, y que los diseños cuasiexperimentales con grupos de control son especialmente escasos (Alfarwan, 2025). Esta brecha es, precisamente, la que motiva el presente estudio.

En América Latina, la literatura específica sobre IAG y programación en bachillerato es incipiente pero creciente. Chávez et al. (2023) advirtieron que la brecha entre la disponibilidad de herramientas de IAG y la capacidad institucional para integrarlas pedagógicamente es particularmente aguda en países con sistemas educativos estratificados. Desde una perspectiva latinoamericana, García et al. (2024) analizaron el uso de herramientas de IA para la personalización del aprendizaje en educación superior ecuatoriana y encontraron que, si bien las percepciones de los docentes son generalmente favorables, la ausencia de marcos metodológicos validados constituye el principal obstáculo para una integración efectiva. Este hallazgo subraya la necesidad de desarrollar y evaluar propuestas pedagógicas concretas, diseñadas para el contexto latinoamericano, que vayan más allá de las recomendaciones generales.

Alfaro y Porras (2024), trabajando con estudiantes de secundaria en Costa Rica, identificaron que la mayoría utilizaba herramientas de IA con fines de apoyo al aprendizaje, pero también reportaron tensión

entre el uso instrumental de estas herramientas y los dilemas éticos que suscitan, especialmente en relación con la autenticidad del trabajo académico. Esta tensión no es exclusiva del contexto latinoamericano, aunque adopta matices particulares en sistemas educativos donde la evaluación por competencias convive con prácticas de enseñanza todavía muy centradas en la transmisión de contenidos. La investigación de Alasgarova y Rzayev (2024) agrega otro ángulo relevante al demostrar, con estudiantes de bachillerato azerbaiyanos, que las herramientas de IA pueden modular la motivación de los adolescentes de maneras que no siempre son predecibles desde los marcos teóricos clásicos.

A nivel metodológico, la mayoría de los estudios revisados presentan dos limitaciones recurrentes: ausencia de grupo de control y dependencia excesiva de instrumentos de percepción o encuesta. La revisión sistemática de Alfawar (2025) corrobora que solo una fracción de los estudios sobre IAG en educación escolar emplearon diseños con pretest y postest, lo que dificulta atribuir causalmente los cambios observados a la intervención. El presente estudio fue diseñado precisamente para cubrir ese vacío metodológico.

Las unidades educativas particulares de Manta operan en un contexto tecnológicamente privilegiado dentro del panorama educativo ecuatoriano, la ciudad cuenta con infraestructura digital relativamente desarrollada, colegios con laboratorios actualizados y conectividad estable, y una comunidad docente con acceso a procesos de formación continua en TIC. Pese a estas condiciones, la observación sistemática realizada por el equipo investigador durante el ciclo lectivo 2024-2025 evidencio que los docentes de TIC no disponían de marcos

metodológicos validados para integrar herramientas de IAG de forma que reforzaran las habilidades de programación sin sustituir el razonamiento algorítmico del estudiante. La improvisación y el uso esporádico de ChatGPT como buscador sofisticado eran las modalidades predominantes. Los estudiantes de segundo año de Bachillerato exhiben dificultades recurrentes en el aprendizaje de programación, particularmente en la depuración de errores, la escritura de algoritmos y la comprensión de la lógica condicional e iterativa. Estas dificultades han sido documentadas en la literatura internacional como propias de los aprendices novatos. La programación exige un tipo de pensamiento abstracto y sistemático que no se desarrolla espontáneamente y que requiere exposición a ejemplos, retroalimentación inmediata y practica guiada. Las herramientas de IAG, al menos en teoría, podrían atender esas necesidades de forma más personalizada que la instrucción magistral tradicional. Sin embargo, la pregunta de si lo hacen efectivamente en el nivel de bachillerato, bajo condiciones controladas, carece aún de respuesta empírica en el contexto ecuatoriano.

Se formula el problema en los siguientes términos, dado el contexto descrito, existe una necesidad no cubierta de evidencia empírica rigurosa sobre los efectos de la integración de herramientas de IAG en el aprendizaje de programación en el bachillerato particular de Manta. La pregunta de investigación central es: ¿cuál es el efecto de la integración sistemática de herramientas de IAG, bajo una secuencia pedagógica diseñada con criterio metodológico, sobre el rendimiento en programación, el pensamiento computacional, la autoeficacia y la motivación intrínseca de los estudiantes? La pertinencia de esta investigación se sostiene en tres dimensiones. La primera es teórica: el estudio contribuye a caracterizar los

mecanismos mediante los cuales la IAG interviene en el aprendizaje de programación en educación secundaria, un nivel educativo subrepresentado en la literatura, tal como lo documentaron (Alfarwan, 2025). Al operar en un nivel educativo y contexto cultural distintos de los estudiados previamente, los resultados permiten poner a prueba la transferibilidad de los hallazgos existentes y, en su caso, matizarlos o enriquecerlos. La segunda dimensión es metodológica. El diseño cuasiexperimental con pretest y posttest aporta evidencia controlada que permite atribuir con mayor rigor los cambios observados a la intervención pedagógica, diferencia sustancial frente a los estudios de caso o las encuestas de percepción que predominan en la literatura latinoamericana sobre IAG y educación. Además, el empleo de instrumentos validados y el uso del ANCOVA como técnica analítica responden a las recomendaciones metodológicas que la propia comunidad científica ha formulado para este campo emergente.

La tercera dimensión es práctica. Los resultados del estudio orientan decisiones curriculares e institucionales en un contexto donde los colegios particulares de Manta han expresado interés en actualizar sus prácticas de enseñanza de TIC. La secuencia pedagógica evaluada puede ser adoptada o adaptada por docentes de contextos similares, lo que convierte a la investigación en un recurso de transferencia de conocimiento aplicable en el corto plazo. Gallent et al. (2023) argumentaron que toda integración de IAG en educación debe ir acompañada de una reflexión explícita sobre ética e integridad académica, y el diseño adoptado en este estudio incorpora esa dimensión como criterio estructural, no como apéndice. El objetivo general es determinar el efecto de la integración sistemática de herramientas de inteligencia artificial

generativa en el aprendizaje de programación de estudiantes de segundo año de Bachillerato General Unificado de unidades educativas particulares de Manta, durante el periodo lectivo 2024-2025. Los objetivos específicos son: (a) comparar el rendimiento en programación del grupo experimental y el grupo de control en el pretest y el posttest; (b) evaluar el efecto de la intervención sobre el pensamiento computacional y la autoeficacia en programación; (c) analizar las variaciones en la motivación intrínseca hacia la asignatura de TIC tras la implementación de la propuesta pedagógica; y (d) describir la intervención pedagógica desarrollada, de modo que pueda ser replicada o adaptada por docentes de contextos similares.

Materiales y Métodos

El estudio adopto un enfoque metodológico mixto con predominancia cuantitativa. El componente cuantitativo siguió un diseño cuasiexperimental con grupos intactos, mediciones de pretest y posttest, y grupo de control no equivalente, en coherencia con los esquemas que Hernández y Mendoza (2018) denominan como los más adecuados cuando no es posible la asignación aleatoria de los participantes a los grupos de tratamiento.

El componente cualitativo, de alcance descriptivo y complementario, se ejecutó mediante el registro sistemático de las sesiones de intervención y la recolección de notas de campo, con el propósito de contextualizar la interpretación de los datos cuantitativos y capturar dinámicas del proceso que los instrumentos estandarizados no podrían recoger. La elección del diseño cuasiexperimental responde a razones tanto epistémicas como éticas. Desde el punto de vista epistémico, cuando el objeto de estudio involucra una intervención pedagógica en

grupos escolares reales, el control total de las condiciones del experimento es imposible sin alterar artificialmente el entorno natural de aprendizaje. Desde el punto de vista ético, reasignar estudiantes a grupos distintos de los que habían sido conformados institucionalmente habría perturbado dinámicas relacionales y organizativas que la escuela había construido durante el año. La decisión de optar por grupos intactos fue validada por los comités de revisión ética de las tres instituciones participantes. Para compensar la posible falta de equivalencia inicial entre grupos, el pretest fue incorporado como covariable en el análisis estadístico posterior, siguiendo la lógica del ANCOVA.

El horizonte temporal del estudio es transversal con medición longitudinal acotada, los datos se recogieron en dos momentos, al inicio y al final de la intervención, sin seguimiento posterior. Esta limitación se reconoce explícitamente y orienta una de las recomendaciones para investigaciones futuras. El nivel de alcance es explicativo, en tanto la combinación de diseño controlado y análisis estadístico inferencial permite ir más allá de la descripción para formular enunciados sobre relaciones causales, aunque con las reservas propias de los estudios cuasiexperimentales.

El estudio se realizó en tres unidades educativas particulares de Manta, cantón de la provincia de Manabí, durante el segundo quimestre del año lectivo 2025-2026, comprendido entre los meses de febrero y junio de 2025. Manta es el segundo puerto de mayor tráfico comercial del Ecuador y concentra una proporción significativa de colegios privados con nivel socioeconómico medio. La asignatura de TIC se imparte en todas las instituciones del nivel bachillerato con una carga horaria de cuatro periodos semanales, lo que garantiza la

suficiencia de horas de exposición para la intervención. La población de referencia estuvo conformada por todos los estudiantes matriculados en el segundo año de Bachillerato General Unificado de las unidades educativas particulares de Manta que impartían TIC en el quimestre de referencia. Conforme a los registros de la Coordinación Zonal de Educación 4, el universo estimado era de aproximadamente 380 estudiantes distribuidos en doce colegios privados.

De ese universo se seleccionaron intencionalmente dos instituciones que cumplieron los siguientes criterios de inclusión: (a) disponer de laboratorio de computación con al menos una computadora por estudiante y acceso a internet de banda ancha; (b) contar con docentes titulados en informática o TIC con tres o más años de experiencia en el nivel; y (c) expresar voluntad institucional para participar, formalizando su adhesión mediante carta de compromiso. Fueron excluidas las instituciones que ya tenían implementado algún programa formal de uso de IA, a fin de evitar efectos de contaminación del tratamiento.

La muestra final quedó integrada por 86 estudiantes, distribuidos en dos grupos: experimental ($n = 44$) y de control ($n = 42$). La asignación de grupos a la condición experimental o de control se realizó al nivel de curso, dos secciones de segundo año, Las edades de los participantes oscilaron entre los 15 y los 17 años. Se obtuvo consentimiento informado de todos los participantes y autorización parental para los menores de 18 años, en cumplimiento con los principios éticos de la investigación educativa. La Tabla 1 describe la composición demográfica de ambos grupos. El primer instrumento fue la Prueba Objetiva de Programación en Python (POPy-25), diseñada ad hoc para este estudio con base

en los contenidos del currículo nacional de TIC de segundo de bachillerato. Consta de 30 ítems distribuidos en tres dimensiones: fundamentos del lenguaje (10 ítems: tipos de datos, variables, operadores), estructuras de control (10 ítems: condicionales simples y anidados, bucles for y while) y funciones con parámetros (10 ítems: definición, argumentos, retorno de valores). Cada ítem correcto otorga un punto, con puntuación total entre 0 y 30. Los ítems son de selección múltiple con cuatro opciones, una de las cuales es la correcta.

Tabla 1. Características demográficas de los grupos experimental y de control.

Variable	Grupo experimental (n=44)	Grupo de control (n=42)	Total (n=86)
Edad promedio (años)	15.8	15.9	15.85
Desv. típica edad	0.74	0.71	0.73
Sexo masculino (%)	54.5	52.4	53.5
Sexo femenino (%)	45.5	47.6	46.5
Experiencia previa programación (%)	34.1	35.7	34.9
Dispositivo propio en casa (%)	88.6	85.7	87.2
Uso previo de ChatGPT (%)	61.4	59.5	60.5
Nota. Los datos provienen del cuestionario de caracterización aplicado en la primera semana. La experiencia previa en programación refiere a haber cursado un taller extracurricular o asignatura optativa antes del bachillerato.			

Fuente: Elaboración propia.

El instrumento fue validado mediante juicio de tres expertos en informática, siguiendo el procedimiento del índice de validez de contenido (IVC) propuesto por Lawshe (1975). El IVC global obtenido fue de 0.88, por encima del umbral de 0.70 recomendado por la literatura metodológica. La confiabilidad, estimada con el coeficiente KR-20 en una prueba piloto con 25 estudiantes ajenos a la muestra, fue de 0.81, lo que indica una adecuada consistencia interna para pruebas objetivas en ciencias computacionales. El segundo instrumento fue la Escala de Autoeficacia en Programación Computacional (APC-16), adaptada al español a partir del instrumento

original desarrollado y validado por Yilmaz y Karaoglan (2023). La adaptación implicó traducción directa e inversa y un piloto con 30 estudiantes de bachillerato de una institución ajena a la muestra. Consta de 16 ítems con escala Likert de cinco opciones (1 = Totalmente en desacuerdo; 5 = Totalmente de acuerdo), organizados en dos dimensiones: autoeficacia general en programación (8 ítems, p. ej., “Soy capaz de resolver un programa en Python cuando encuentro un error de lógica”) y autoeficacia específica ante errores de código (8 ítems, p. ej., “Puedo identificar porque mi programa no produce el resultado esperado”). El coeficiente alfa de Cronbach del instrumento adaptado fue de 0.87 en la prueba piloto y de 0.85 en la aplicación definitiva.

El tercer instrumento fue el Inventario de Motivación hacia las TIC (IMT-18), construido para este estudio sobre la base del marco teórico de la teoría de la autodeterminación. Contiene 18 ítems en escala Likert de cinco puntos, organizados en tres sub-escalas: motivación intrínseca (6 ítems, p. ej., “Practico programación porque me resulta interesante en si misma”), motivación extrínseca regulada por identificación (6 ítems, p. ej., “Aprendo a programar porque valoro lo que puede aportarme en el futuro”) y motivación extrínseca regulada externamente (6 ítems, p. ej., “Hago las tareas de TIC para no tener problemas con el docente”). La confiabilidad global fue de alfa = 0.84. Los tres instrumentos se aplicaron en forma digital, a través de la plataforma Google Forms, en el pretest y el postest. La intervención tuvo una duración de ocho semanas efectivas de clase, con una frecuencia de cuatro horas semanales en el laboratorio de computación, para un total de 32 horas de contacto directo. El grupo experimental siguió una secuencia didáctica estructurada en cinco fases cíclicas, exploración

guiada, producción asistida por IAG, revisión crítica del código generado, enriquecimiento autónomo y evaluación formativa. El grupo de control recibió instrucción convencional basada en exposición docente, ejercicios del libro de texto institucional y practicas individuales en el IDE IDLE de Python, sin acceso a herramientas de IAG.

Ambos grupos fueron instruidos por los mismos docentes de planta de cada institución. Todos los docentes asistieron a un taller de capacitación inicial de ocho horas antes del comienzo del estudio; los docentes del grupo experimental recibieron un módulo adicional de seis horas sobre integración pedagógica de IAG, diseñado por el equipo investigador. Este módulo incluyó: (a) principios de aprendizaje activo con asistentes de código; (b) estrategias de cuestionamiento metacognitivo; y (c) criterios para diferenciar las tareas evaluadas con y sin apoyo de IA. La capacitación docente es una variable que la literatura identifica como

crítica para el éxito de las integraciones tecnológicas (Alfarwan, 2025).

En las semanas 1 y 2, los estudiantes del grupo experimental conocieron ChatGPT y aprendieron a formular prompts específicos para obtener explicaciones de código Python. El docente modelaba el proceso, escribía un prompt en la pizarra, mostraba la respuesta de la IA y luego guiaba una discusión sobre la pertinencia y exactitud de esa respuesta. A partir de la semana 3, los estudiantes comenzaron a utilizar GitHub Copilot en el IDE Visual Studio Code, instalado en los laboratorios de las instituciones participantes. El docente del grupo experimental pautaba ejercicios de complejidad creciente que la IA no podía resolver directamente, lo que obligaba a los estudiantes a descomponer el problema antes de solicitar asistencia. La Tabla 2 describe las fases de la intervención y las actividades asociadas.

Tabla 2. Fases de la intervención pedagógica con herramientas de IAG.

Semana	Fase	Actividad principal	Herramienta IAG
1-2	Exploración guiada	Introducción a ChatGPT; elaboración de prompts; lectura crítica de respuestas; sintaxis básica Python	ChatGPT (GPT-4o)
3-4	Producción asistida	Escritura de funciones y estructuras de control con sugerencias de GitHub Copilot; aceptación selectiva de código	GitHub Copilot
5-6	Revisión crítica	Identificación de alucinaciones en código generado; comparación entre versiones; corrección manual argumentada	ChatGPT + corrección manual
7	Enriquecimiento autónomo	Proyecto integrador individual: programa funcional sin apoyo de IAG durante la evaluación	Sin herramienta IAG
8	Evaluación y cierre	Postest; retroalimentación entre pares; sesión de metacognición sobre el proceso de aprendizaje	Google Forms + debate estructurado

Fuente: Elaboración propia.

El análisis cuantitativo siguió una secuencia de cuatro pasos. En primer lugar, se aplicó estadística descriptiva para caracterizar las distribuciones de las variables en ambos grupos y en ambos momentos de medición, reportando medias, desviaciones típicas, valores mínimos y máximos. En segundo lugar, se evaluó la normalidad de las distribuciones con la prueba de Shapiro-Wilk, adecuada para el tamaño

muestral del estudio. La homogeneidad de varianzas entre grupos se verificó con la prueba de Levene. En tercer lugar, la equivalencia inicial entre grupos fue examinada mediante la prueba *t* de Student para muestras independientes aplicada a los datos del pretest. La técnica central de análisis fue el análisis de covarianza (ANCOVA), donde las puntuaciones del postest constituyeron la

variable dependiente, la condición de grupo (experimental/control) fue el factor fijo y las puntuaciones del pretest actuaron como covariable. Esta elección permite aislar el efecto de la intervención de las diferencias previas entre grupos.

Se verifico el supuesto de homogeneidad de pendientes de regresión, que es el supuesto crítico del ANCOVA, calculando la interacción entre la covariable y el factor: la interacción no resulto significativa en ninguna de las variables (todos los $p > .10$), lo que valida el modelo. El tamaño del efecto se estimó con eta cuadrado parcial ($hp2$), con los criterios de Cohen (1988): $hp2 < 0.06$ efecto pequeño, entre 0.06 y 0.14 efecto mediano, superior a 0.14 efecto grande. El poder estadístico se calculó a posteriori con el paquete pwr de R. El nivel de significación estadística fue de $\alpha = .05$. Para el procesamiento de los datos se utilizaron SPSS 26.0 y R 4.3.0. El análisis cualitativo de los registros de observación fue realizado mediante codificación temática inductiva, con saturación

de categorías verificada por dos investigadores de forma independiente, siguiendo el protocolo de acuerdo entre codificadores con kappa de Cohen ($kappa = 0.81$).

Resultados y Discusión

La prueba de Shapiro-Wilk no rechazo la hipótesis de normalidad para ninguna de las variables en ninguno de los grupos (todos los $p > .05$). La prueba de Levene confirmo la homogeneidad de varianzas en el pretest para las tres variables, rendimiento en programación ($F(1,84) = 0.43$, $p = .514$), autoeficacia ($F(1,84) = 0.67$, $p = .416$) y motivación ($F(1,84) = 1.12$, $p = .293$). Estos resultados validan el uso de pruebas paramétricas. La comparación de medias de pretest tampoco revelo diferencias significativas entre grupos para ninguna de las tres variables (todos los $p > .40$), lo que confirma la equivalencia inicial entre el grupo experimental y el grupo de control. La Tabla 3 presenta la estadística descriptiva completa del pretest y el posttest.

Tabla 3. Estadísticos descriptivos del pretest y el posttest por grupo y variable.

Variable	Grupo	Pre-M	Pre DE	Post M	Post DE	Ganancia	d Cohen
Rendimiento programación (max=30)	Experimental	13.82	2.91	22.41	2.76	+8.59	1.14
	Control	13.64	2.88	17.12	3.04	+3.48	--
Autoeficacia APC-16 (max=80)	Experimental	48.73	7.14	66.89	6.23	+18.16	1.08
	Control	47.90	7.31	54.38	7.67	+6.48	--
Motivacion IMT-18 (max=90)	Experimental	55.41	8.02	73.16	7.45	+17.75	0.97
	Control	54.88	8.17	61.02	8.44	+6.14	--

Fuente: Elaboración propia.

El ANCOVA mostro un efecto principal significativo del grupo sobre el rendimiento en programación en el posttest, controlando el pretest como covariable: $F(1,83) = 52.34$, $p < .001$, $hp2 = .387$. El tamaño del efecto es grande. La media ajustada del grupo experimental fue de 22.38 puntos ($EE = 0.41$) frente a 17.16 puntos del grupo de control ($EE = 0.43$). La diferencia entre medias ajustadas

fue de 5.22 puntos sobre 30, con un intervalo de confianza del 95 % que oscila entre 4.37 y 6.07. El poder estadístico a posteriori fue de .99. El covariable pretest presento un efecto significativo sobre el posttest ($F(1,83) = 28.19$, $p < .001$), lo que confirma que las habilidades iniciales de programación predicen el rendimiento final con independencia del grupo.

Al desagregar por dimensión de la prueba, las diferencias fueron significativas en las tres: fundamentos del lenguaje ($F(1,83) = 31.44, p < .001$), estructuras de control ($F(1,83) = 47.28, p < .001$) y funciones con parámetros ($F(1,83) = 38.92, p < .001$). La dimensión de estructuras de control mostro el mayor tamaño del efecto

($hp2 = .363$), lo que resulta coherente con las actividades de la intervención, que pusieron especial énfasis en la construcción y verificación de bucles y condicionales. La Tabla 4 presenta los resultados del ANCOVA para las tres variables dependientes.

Tabla 4. Resultados del ANCOVA para las tres variables dependientes del estudio.

Variable dependiente	SC	gl	CM	F	p	hp2	Poder
Rendimiento en programación	582.14	1	582.14	52.34	< .001	.387	.99
Autoeficacia (APC-16)	4361.07	1	4361.07	48.89	< .001	.371	.99
Motivación intrínseca (IMT-18)	3904.22	1	3904.22	44.61	< .001	.350	.99
Covariable: Pretest	--	1	--	$p < .05$	Si	--	--

Fuente: Elaboración propia.

Respecto a la autoeficacia, los resultados del ANCOVA revelaron un efecto principal del grupo estadísticamente significativo y de tamaño grande: $F(1,83) = 48.89, p < .001, hp2 = .371$. La media ajustada del grupo experimental en el posttest fue de 66.82 puntos sobre 80 ($EE = 0.94$), frente a 54.44 del grupo de control ($EE = 0.97$). Esta diferencia equivale a 12.38 puntos, con un intervalo de confianza del 95 % entre 10.44 y 14.32 puntos.

Este resultado puede interpretarse como que el grupo experimental desarrollo una mayor confianza general en programación, pero que la seguridad para afrontar errores específicos, una habilidad más sofisticada, requiere exposición más prolongada para consolidarse. El propio Yilmaz y Karaoglan (2023) señalaron que la autoeficacia ante la depuración de código es la dimensión que más tiempo requiere para desarrollarse bajo apoyo de IAG.

Tabla 5. Ganancias por dimensión en la Escala de Autoeficacia en Programación Computacional (APC-16),

Dimensión (rango teórico: 8-40)	GE Pretest M(DE)	GE Posttest M(DE)	GC Posttest M(DE)	Dif. GE-GC posttest
Autoeficacia general (8 ítems)	24.31 (3.72)	42.17 (3.41)	32.91 (4.12)	9.26 **
Autoeficacia ante errores (8 ítems)	24.42 (4.01)	24.72 (3.87)	21.47 (4.44)	3.25 *
Total, APC-16 (16 ítems, max = 80)	48.73 (7.14)	66.89 (6.23)	54.38 (7.67)	12.51 **

Nota. GE = grupo experimental; GC = grupo de control. * $p < .05$;
 ** $p < .001$. Las diferencias se basan en las medias posttest
 ajustadas por ANCOVA.

Fuente: Elaboración propia.

La dimensión de autoeficacia ante errores registro una ganancia inferior a la de autoeficacia general, aunque aún significativa.

La motivación intrínseca mostro el segundo mayor tamaño del efecto entre las variables dependientes: $F(1,83) = 44.61, p < .001, hp2 = .350$. La media ajustada del posttest del grupo experimental fue de 73.09 puntos sobre 90 ($EE = 1.14$), frente a 61.08 del grupo de control ($EE = 1.18$). De particular importancia es que la sub-escalas de motivación extrínseca regulada externamente no mostro diferencias significativas entre grupos en el posttest ajustado ($p = .207$), lo que indica que la intervención no incremento la motivación de tipo instrumental ni la dependencia externa de los estudiantes hacia las herramientas de IA. La Tabla 6 presenta las ganancias por sub-escalas. Los registros cualitativos de observación revelaron patrones de interacción con valor interpretativo

que los datos cuantitativos no capturan de manera directa. En las sesiones de las semanas 1 y 2, los estudiantes mostraron alta curiosidad hacia ChatGPT, pero tendían a formular preguntas amplias y poco específicas. El docente del grupo experimental intervino sistemáticamente para modelar la elaboración de prompts estructurados: en lugar de escribir 'hazme un programa en Python', el estudiante aprende a escribir 'escribe una función Python que reciba una lista de números enteros y devuelva el promedio.

Tabla 6. *Ganancias por sub-escalas del Inventario de Motivación hacia las TIC (IMT-18).*

Sub-escalas (rango: 6-30)	GE Pretest M(DE)	GE Posttest M(DE)	GC Posttest M(DE)	p (ANCOVA)
Motivación intrínseca (6 ítems)	19.31 (3.14)	25.64 (2.97)	21.07 (3.41)	< .001 **
Motivación identificada (6 ítems)	18.82 (3.27)	24.91 (3.01)	21.14 (3.62)	.002 **
Motivación externa (6 ítems)	17.28 (3.41)	22.61 (3.14)	18.81 (3.77)	.207 n.s.
Total, IMT- 18 (18 ítems, max = 90)	55.41 (8.02)	73.16 (7.45)	61.02 (8.44)	< .001 **
Nota. ms. = no significativo. La motivación extrínseca regulada externamente no registro diferencias significativas entre grupos en el posttest, lo que indica que la intervención no incremento la motivación instrumental ni la dependencia de herramientas externas.				

Fuente: Elaboración propia.

Este desplazamiento cognitivo, de la solicitud vaga a la especificación precisa, es en sí mismo un aprendizaje de lógica computacional. Las semanas 3 y 4 evidenciaron el patrón que la literatura denomina aceptación acrítica, la mayoría de los estudiantes aceptaba las sugerencias de GitHub Copilot sin leerlas con detenimiento. Solo el 38% revisaba el código antes de ejecutarlo en la primera sesión de la fase, porcentaje que ascendió al 71% en la segunda sesión tras la intervención docente orientada al cuestionamiento. Denny et al. (2024) documentaron este mismo patrón en

contextos universitarios y lo identificaron como uno de los riesgos pedagógicos más relevantes de la integración no mediada de asistentes de código. El presente estudio confirma que el patrón se reproduce en el bachillerato y que la mediación docente lo puede revertir en un tiempo relativamente breve. La fase de revisión crítica fue valorada por los propios docentes como el momento de mayor complejidad cognitiva de toda la secuencia, y fue también el que generó mayor participación entre los estudiantes. La identificación de alucinaciones, es decir, fragmentos de código sintácticamente válidos, pero semánticamente incorrectos o contextualmente inadecuados, activo habilidades de depuración que rara vez emergen en la enseñanza convencional de Python a este nivel. En las sesiones de cierre reflexivo, varios estudiantes manifestaron que aprender a desconfiar del código generado automáticamente les había enseñado más sobre lógica de programación que los ejercicios del libro de texto.

Esta observación apunta a un efecto pedagógico de orden superior que los instrumentos cuantitativos no capturan en su totalidad. La Tabla 7 sintetiza los principales patrones cualitativos observados. Los hallazgos del estudio permiten afirmar que la integración sistemática de herramientas de IAG, bajo una secuencia pedagógica diseñada con criterio metodológico, produce ganancias significativas y de tamaño grande en el rendimiento de programación, la autoeficacia y la motivación intrínseca de estudiantes de bachillerato. Esta conclusión no es trivial. La revisión sistemática de Jauhiainen y Garagorry (2024) mostro que las evidencias sobre el impacto de la IAG en educación K-12 son heterogéneas, algunos estudios reportan beneficios netos, otros no encuentran diferencias significativas entre grupos, y unos pocos señalan efectos negativos

asociados a la dependencia tecnológica. La divergencia probablemente no refleja una ineficacia inherente de las herramientas, sino la heterogeneidad de los diseños de integración y la presencia o ausencia de mediación docente explícita. La magnitud del efecto obtenido en el rendimiento en programación ($hp2 = .387$)

supera los valores reportados por Yilmaz y Karaoglan (2023) en su estudio experimental con universitarios turcos, aunque la comparación directa requiere cautela dado que los niveles educativos, los idiomas, los contenidos de programación y los instrumentos difieren entre estudios.

Tabla 7. Patrones de interacción y hallazgos cualitativos por fase de la intervención.

Semana	Patrón predominante	Desafío docente	Hallazgo relevante
1-2	Alta curiosidad; preguntas amplias y poco estructuradas hacia el chatbot	Guiar la elaboración de prompts específicos y progresivos	Los estudiantes con menor experiencia previa mostraron el mayor incremento relativo de interés al recibir retroalimentación inmediata
3-4	Aceptación acrítica de sugerencias de GitHub Copilot en la primera sesión	Promover el cuestionamiento sistemático del código sugerido antes de ejecutarlo	El porcentaje de estudiantes que revisaba el código antes de ejecutarlo subió del 38% al 71% entre la sesión 1 y la sesión 2 de la fase
5-6	Mayor disposición al análisis crítico; identificación activa de errores semánticos	Gestionar la frustración ante código con errores lógicos generados por IA	Emergencia espontánea de estrategias de depuración colaborativa entre pares; comparación activa de versiones de código
7-8	Programación autónoma de mayor complejidad; transferencia de aprendizajes sin apoyo de IAG	Garantizar la originalidad y autonomía de los proyectos finales; retroalimentación formativa diferenciada	El 82% de los proyectos finales cumplieron los criterios de originalidad; el 18% restante requirió revisión y ajuste antes de la evaluación
Nota. Los porcentajes provienen del conteo sistematizado de observaciones de campo y no constituyen estadísticos inferidos a la población general. Los registros fueron codificados por dos investigadores con un acuerdo Inter jueces de kappa = .81.			

Fuente: Elaboración propia.

Lo que si resulta coherente entre ambos trabajos es la dirección del efecto y la relevancia de la dimensión computacional: tanto en el contexto universitario analizado por dichos autores como en el bachillerato, la exposición a herramientas de IAG bajo instrucción guiada se asocia con mejoras en el pensamiento algorítmico, no solo en la capacidad de producir código correcto. La diferencia de impacto entre las dimensiones de motivación merece interpretación. Que la motivación extrínseca regulada externamente no mostrara diferencias significativas entre grupos en el posttest indica que la intervención no incremento la dependencia instrumental de los estudiantes hacia las herramientas de IA. Al contrario, las formas más autónomas de motivación, la intrínseca y la identificada, fueron las que se incrementaron significativamente. Esta configuración es consistente con lo que predice la teoría de la autodeterminación cuando las actividades

ofrecen novedad, competencia percibida e internalización del valor de la tarea. Chiu (2024), en su agenda de investigación para la educación superior con IAG, señala que uno de los efectos más prometedores de la integración bien diseñada es el incremento de la motivación autónoma, y el presente estudio extiende ese hallazgo al nivel de bachillerato. El resultado sobre autoeficacia ante errores merece atención particular. La ganancia del grupo experimental fue significativa pero inferior a la de autoeficacia general, lo que sugiere que la confianza para afrontar errores de código, la habilidad de depuración propiamente dicha es la dimensión de la autoeficacia que más tiempo requiere para consolidarse. Ocho semanas de intervención fueron suficientes para producir ganancias significativas, pero probablemente no para alcanzar el nivel de automaticidad que se observaría tras un semestre completo de integración. Esta interpretación es coherente

con la distinción, establecida en la literatura de psicología del aprendizaje, entre autoeficacia de tarea, que responde rápidamente a los éxitos iniciales, y autoeficacia de proceso, que requiere exposición sostenida a situaciones de dificultad controlada.

Los efectos encontrados en esta investigación son consistentes con los patrones reportados por estudios en niveles educativos superiores. Yilmaz y Karaoglan (2023) encontraron que el uso de ChatGPT en un curso universitario de programación incremento el pensamiento computacional, la autoeficacia y la motivación del grupo experimental respecto al de control, con tamaños del efecto moderados a grandes. La diferencia más notable entre ese estudio y el presente es que los estudiantes de bachillerato de Manta partían de un nivel de expertise considerablemente menor, lo que puede explicar en parte las mayores ganancias absolutas observadas: los aprendices novatos tienen más margen de mejora.

Wu y Zhang (2025) en su estudio con 500 estudiantes de secundaria de Nueva Zelanda, encontraron que las aplicaciones de IAG producen un efecto sustancial sobre la capacidad de innovación ($\beta = .862$) y la literaria digital. Aunque las variables dependientes difieren de las del presente estudio, ambas investigaciones coinciden en que el grupo etario del bachillerato, que comprende adolescentes entre los 15 y 17 años, muestra una capacidad de respuesta a la integración de IAG comparable o superior a la de estudiantes universitarios, contrariamente a lo que algunos autores han sugerido. La literatura sobre programación en particular, revisada por Denny et al. (2024), también muestra que las herramientas de IAG producen los mayores beneficios cuando se integran de forma que obligan al estudiante a ir más allá de

la aceptación pasiva del código generado. En el presente estudio, esa dinámica se materializó en la fase de revisión crítica, que fue la de mayor impacto cualitativo según los registros de observación.

Este resultado converge con la observación de Denny et al. (2024) de que el verdadero valor pedagógico de los generadores de código reside en el proceso de cuestionamiento crítico que el docente debe provocar, no en la eficiencia con que la herramienta resuelve los ejercicios. Desde la perspectiva iberoamericana, Gallent et al. (2023) señalaron que la ética en el uso de la IAG no es un obstáculo para la integración tecnológica sino, justamente, una de sus condiciones de posibilidad. El diseño de la intervención del presente estudio incorporó esa perspectiva desde el inicio, diseñando explícitamente las condiciones de uso ético y diferenciando las tareas en que el apoyo de la IA era esperado de aquellas en que la producción autónoma era evaluada.

El hecho de que el 82% de los proyectos finales cumpliera los criterios de originalidad validar esta opción diseñar como viable y efectiva. El estudio presenta limitaciones que deben considerarse al interpretar y generalizar los resultados. La primera limitación es el tamaño muestral moderado, aunque fue suficiente para alcanzar un poder estadístico superior al 99% dado el tamaño del efecto encontrado, restringe la generalizabilidad a instituciones de características socioeconómicas distintas o a ciudades con menor dotación tecnológica. La segunda limitación es inherente al diseño cuasiexperimental con grupos intactos, aunque el ANCOVA controló las diferencias iniciales entre grupos a través de las puntuaciones del pretest, no es posible descartar por completo la influencia de variables no medidas, como el efecto del docente específico, el clima de aula o

la historia grupal de cada sección. La tercera limitación es la duración de la intervención. Ocho semanas son suficientes para evaluar efectos a corto plazo, pero no garantizan que los aprendizajes se consoliden en el tiempo. Estudios longitudinales con mediciones de seguimiento a tres o seis meses serían necesarios para determinar la persistencia de los efectos. La cuarta limitación es la restricción contextual, la investigación se circunscribe a tres unidades educativas particulares de Manta. Sus conclusiones no deben extrapolarse automáticamente a instituciones fiscomisionales, fiscales o a zonas rurales del Ecuador, donde las condiciones de acceso tecnológico son distintas. La quinta limitación es la posible reactividad al instrumento, algunos estudiantes pueden haber respondido los cuestionarios de autoeficacia y motivación según lo que percibieron como socialmente esperado, especialmente en el postest.

Conclusiones

Los hallazgos del presente estudio llevan a sostener que la integración sistemática de herramientas de inteligencia artificial generativa en la enseñanza de programación, cuando se sustenta en una secuencia pedagógica que contempla fases de producción crítica, revisión activa del código y evaluación autónoma, produce efectos positivos y estadísticamente significativos sobre el rendimiento académico, la autoeficacia y la motivación intrínseca de los estudiantes de bachillerato. Estos efectos no son consecuencia de la mera exposición a la tecnología, sino el resultado de una mediación docente deliberada que orienta el uso de la IAG hacia el desarrollo del pensamiento algorítmico y la autonomía cognitiva. La distinción entre motivación autónoma y motivación instrumental resulta pedagógicamente relevante. Que la intervención haya incrementado las formas más

autónomas de motivación sin potenciar la dependencia instrumental indica que es posible integrar herramientas de IAG sin sacrificar los objetivos de formación crítica y autonomía intelectual que son propios del nivel de bachillerato. Esta conclusión desafía la dicotomía entre uso de IA y aprendizaje auténtico que permea parte del debate educativo contemporáneo y sugiere que la tensión real no se ubica en la herramienta sino en el diseño pedagógico que la enmarca.

Para la práctica docente en unidades educativas particulares de Manta, la evidencia apunta a que la adopción de marcos pedagógicos estructurados para la integración de IAG, con especial atención a las fases de revisión crítica y producción autónoma, produce resultados superiores a la improvisación o al uso esporádico. La formación docente previa a la implementación emerge como condición operativa que no puede soslayarse: el docente que no comprende cómo funciona la herramienta ni cuáles son sus limitaciones no puede guiar a sus estudiantes para que la utilicen de forma productiva y ética.

Futuras investigaciones deberían explorar si los efectos observados se mantienen en el tiempo mediante diseños longitudinales, si se replican en instituciones fiscomisionales o fiscales con distintos niveles de acceso tecnológico, y si el tipo específico de herramienta de IAG empleada, comparando por ejemplo modelos de código especializado con modelos de lenguaje general, produce diferencias en los resultados. La investigación sobre la ética del uso de la IA en contextos escolares latinoamericanos sigue siendo un campo de producción necesaria y urgente que la comunidad académica apenas comienza a habitar con la profundidad empírica que el tema merece.

Referencias Bibliográficas

- Alasgarova, R., & Rzayev, J. (2024). The role of artificial intelligence in shaping high school students' motivation. *International Journal of Technology in Education and Science*, 8(2), 311–324. <https://doi.org/10.46328/ijtes.553>
- Alfaro-Salas, H., & Porras, J. A. D. (2024). Percepciones y aplicaciones de la IA entre estudiantes de secundaria. *Revista Docentes 2.0*, 17(1), 200–215. <https://doi.org/10.37843/rtded.v17i1.458>
- Alfarwan, A. (2025). Generative AI use in K–12 education: A systematic review. *Frontiers in Education*, 10, 1647573. <https://doi.org/10.3389/feduc.2025.1647573>
- Chávez, M. E., Labrada, E., Carbajal, E., Pineda, E., & Alatrastre, Y. (2023). Inteligencia artificial generativa para fortalecer la educación superior. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 4(3), 767–784. <https://doi.org/10.56712/latam.v4i3.1113>
- Chiu, T. K. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, 6, 100197. <https://doi.org/10.1016/j.caeai.2023.100197>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Denny, P., Prather, J., Becker, B. A., Finnie-Ansley, J., Hellas, A., Leinonen, J., Luxton-Reilly, A., Reeves, B. N., Santos, E. A., & Sarsa, S. (2024). Computing education in the era of generative AI. *Communications of the ACM*, 67(2), 56–67. <https://doi.org/10.1145/3624720>
- Gallent-Torres, C., Zapata-González, A., & Ortego-Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: Una mirada desde la ética y la integridad académica. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 29(2). <https://doi.org/10.30827/relieve.v29i2.29134>
- García Caicedo, S., Reyes Vélez, N., Solórzano Zambrano, A., Quiñónez Godoy, N., & Vega Macías, J. (2024). Análisis al uso de herramientas de inteligencia artificial para la personalización del aprendizaje en la educación superior. *Revista Científica Multidisciplinar G-ner@ndo*, 5(1). <https://doi.org/10.60100/rcmg.v5i1.214>
- García-Peñalvo, F. J., Llorens-Largo, F., & Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED. Revista Iberoamericana de Educación a Distancia*, 27(1), 9–39. <https://doi.org/10.5944/ried.27.1.37716>
- Hernández-Sampieri, R., & Mendoza, C. P. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill.
- Jauhainen, J. S., & Garagorry Guerra, A. (2024). Generative AI and education: Dynamic personalization of pupils' school learning material with ChatGPT. *Frontiers in Education*, 9, 1288723. <https://doi.org/10.3389/feduc.2024.1288723>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.tb01393.x>
- Wu, D., & Zhang, J. (2025). Generative artificial intelligence in secondary education: Applications and effects on students' innovation skills and digital literacy. *PLOS ONE*, 20(5), e0323349. <https://doi.org/10.1371/journal.pone.0323349>
- Yilmaz, R., & Karaoglan Yilmaz, F. G. (2023). The effect of generative artificial intelligence (AI)-based tool use on students' computational thinking skills, programming self-efficacy and motivation. *Computers and Education: Artificial Intelligence*, 4, 100147. <https://doi.org/10.1016/j.caeai.2023.100147>



Esta obra está bajo una licencia de Creative Commons Reconocimiento-No Comercial 4.0 Internacional. Copyright © Ángel Leandro Alcívar Cedeño, Julio Ernesto Ramírez, Steven Arturo Torres Burgos.

Declaraciones éticas y editoriales del artículo

Contribución de los autores (Taxonomía CRediT)

Ángel Leandro Alcívar Cedeño: conceptualización de la investigación, diseño metodológico, desarrollo del proceso investigativo, análisis formal de los datos, redacción del borrador original del manuscrito, revisión crítica del contenido científico y supervisión general del estudio.

Julio Ernesto Ramírez: curación y organización de los datos, participación en la recolección de información, validación de los resultados obtenidos y elaboración de representaciones gráficas y visualización de los datos.

Steven Arturo Torres Burgos: conceptualización de la investigación, diseño metodológico, desarrollo del proceso investigativo, análisis formal de los datos, redacción del borrador original del manuscrito, revisión crítica del contenido científico y supervisión general del estudio.

Declaración de conflicto de intereses

Los autores declaran que no existe conflicto de intereses en relación con la investigación presentada, la autoría del manuscrito ni la publicación del presente artículo.

Declaración de financiamiento

La presente investigación no recibió financiamiento específico de agencias públicas, comerciales o de organizaciones sin fines de lucro. En caso de existir financiamiento institucional o externo, este deberá ser declarado explícitamente por los autores en esta sección.

Declaración del editor

El editor responsable certifica que el proceso editorial del presente artículo se desarrolló conforme a los principios de integridad científica, transparencia y buenas prácticas editoriales. El manuscrito fue sometido a un proceso de evaluación mediante revisión por pares doble ciego, garantizando la confidencialidad de la identidad de los autores y revisores durante todo el proceso de dictamen académico. Asimismo, el editor declara que el artículo cumple con los criterios científicos, metodológicos y éticos establecidos por la revista.

Declaración de los revisores

Los revisores externos que participaron en la evaluación del presente manuscrito declaran haber realizado el proceso de revisión de manera objetiva, independiente y confidencial. Asimismo, manifiestan que no mantienen conflictos de interés con los autores ni con la investigación evaluada, y que sus observaciones y recomendaciones se fundamentan exclusivamente en criterios científicos, metodológicos y académicos.

Declaración ética de la investigación

Los autores declaran que la investigación se desarrolló respetando los principios éticos de la investigación científica, garantizando la confidencialidad de los datos y el respeto a los participantes del estudio. En los casos en que la investigación involucre seres humanos, los procedimientos deben ajustarse a los principios éticos establecidos en la Declaración de Helsinki y a las normativas institucionales correspondientes.

Declaración sobre el uso de inteligencia artificial

Los autores declaran que el uso de herramientas de inteligencia artificial, en caso de haberse utilizado durante el proceso de investigación o redacción del manuscrito, se realizó únicamente como apoyo técnico para mejorar la claridad del lenguaje o el análisis de información, manteniendo siempre la responsabilidad intelectual sobre el contenido del artículo. Las herramientas de inteligencia artificial no fueron utilizadas como autoras del manuscrito ni sustituyen la responsabilidad académica de los investigadores.

Disponibilidad de datos

Los datos que respaldan los resultados de esta investigación estarán disponibles previa solicitud razonable al autor de correspondencia, respetando las normas éticas y de confidencialidad establecidas por la investigación.

