

**Recibido:** 2026-05-20

**Aceptado:** 2026-06-03

**Publicado:** 2026-06-17

**Alfabetización en inteligencia artificial generativa: del uso  
instrumental a la autorregulación en un curso intensivo de cálculo**

**Generative Artificial Intelligence Literacy: From Instrumental Use to  
Self-Regulation in an Intensive Calculus Course**

**Autor**

**Melvin Ramírez Bogantes <sup>1</sup>**

Escuela de Matemática

[mramirez@itcr.ac.cr](mailto:mramirez@itcr.ac.cr)

<https://orcid.org/0000-0001-5516-0085>

**Instituto Tecnológico de Costa Rica**

Cartago – Costa Rica

## Resumen

La incorporación de la inteligencia artificial generativa (IAGen) en la educación superior plantea un dilema: puede potenciar el aprendizaje o convertirse en un atajo que simula competencia sin desarrollarla. Esta investigación tuvo como objetivo analizar en qué medida un proceso estructurado de alfabetización en IAGen facilita la transición desde un uso instrumental de estas herramientas (como generadoras de respuestas) hacia un uso reflexivo y autorregulado (como tutor que fortalece la autonomía), en un curso intensivo de verano de Cálculo Diferencial e Integral. Se empleó un diseño mixto que combinó la comparación del desempeño de los estudiantes antes y después de la intervención (pruebas resueltas con y sin asistencia de IA,  $n = 24$ ) con entrevistas a estudiantes de distintos niveles de rendimiento. La intervención se organizó en cinco fases: reconocer los errores de la IA, aprender a formular buenas preguntas, auditar críticamente sus respuestas, adaptar el apoyo a cada nivel y acordar un compromiso ético de uso. Antes de la alfabetización, los estudiantes lograban buenos resultados con ayuda de la IA, pero su rendimiento caía de forma notable al resolver problemas por su cuenta. Tras la intervención esa brecha se redujo, el desempeño autónomo mejoró de manera significativa (un efecto moderado y educativamente relevante) y la proporción de estudiantes en riesgo de reprobación disminuyó a la mitad; las entrevistas confirmaron un uso más consciente y estratégico de la herramienta. Se concluye que el factor decisivo no es la herramienta en sí, sino la pedagogía que la acompaña: una alfabetización deliberada permite que la IAGen funcione como un genuino apoyo al aprendizaje y no como un sustituto del pensamiento.

**Palabras clave:** alfabetización en inteligencia artificial generativa; aprendizaje autorregulado; autonomía académica; educación matemática; cálculo diferencial e integral.

## Abstract

The integration of generative artificial intelligence (GenAI) into higher education poses a dilemma: it can enhance learning or become a shortcut that simulates competence without building it. This study aimed to analyze the extent to which a structured GenAI literacy process facilitates the transition from an instrumental use of these tools (as answer generators) toward a reflective and self-regulated use (as a tutor that strengthens autonomy), within an intensive summer Differential and Integral Calculus course. A mixed-methods design was used, combining a comparison of student performance before and after the intervention (assessments completed with and without AI assistance,  $n = 24$ ) with interviews of students at different achievement levels. The intervention was organized into five phases: recognizing AI errors, learning to formulate effective prompts, critically auditing AI responses, adapting support to each level, and agreeing on an ethical use commitment. Before the literacy process, students achieved strong results with AI assistance, but their performance dropped sharply when solving problems on their own. After the intervention this gap narrowed, autonomous performance improved significantly (a moderate, educationally meaningful effect), and the proportion of students at risk of failing was cut in half; interviews confirmed a more conscious and strategic use of the tool. The study concludes that the decisive factor is not the tool itself but the pedagogy surrounding it: deliberate literacy allows GenAI to function as genuine support for learning rather than a substitute for thinking.

**Keywords:** generative artificial intelligence literacy; self-regulated learning; academic autonomy; mathematics education; differential and integral calculus.

## Introducción

La irrupción de los modelos de lenguaje de gran escala en el ámbito educativo constituye uno de los fenómenos más disruptivos del último lustro. Desde la publicación masiva de ChatGPT a finales de 2022, las universidades enfrentan un dilema pedagógico inédito: integrar o prohibir herramientas que pueden generar soluciones matemáticas paso a paso, explicaciones conceptuales personalizadas y retroalimentación instantánea (Kasneji et al., 2023). En el campo específico de la educación matemática, la inteligencia artificial generativa (IAGen) ofrece oportunidades sin precedentes para la tutoría personalizada, el andamiaje cognitivo y la práctica deliberada; sin embargo, también plantea riesgos significativos de dependencia tecnológica, atrofia del pensamiento crítico y plagio académico (Cotton, Cotton y Shipway, 2024). Este escenario se agudiza en los cursos intensivos de verano, donde la compresión temporal del contenido amplifica tanto las posibilidades como los peligros del uso de IAGen.

La literatura reciente documenta un crecimiento exponencial en la adopción de herramientas de IAGen por parte de estudiantes universitarios. Chan y Hu (2023) reportan que más del 80 % de los estudiantes encuestados en una universidad de Hong Kong habían utilizado ChatGPT al menos una vez para tareas académicas, mientras que Lim, Gunasekara, Pallant, Rajendran y Pham (2023) encuentran patrones similares en contextos del sudeste asiático. En el ámbito de las STEM, Wardat, Tashtoush, AlAli y Saleh (2024) evidencian que la IAGen puede mejorar la comprensión conceptual en matemáticas cuando se integra con estrategias pedagógicas explícitas, hallazgo consistente con lo reportado por Dahlkemper, Lahme y Klein (2023) en física universitaria. No obstante, diversos estudios advierten que el uso no guiado de estas herramientas conduce a lo que Mollick y Mollick (2023) denominan "ilusión de competencia": los estudiantes producen respuestas correctas sin comprensión genuina, lo que se traduce en un desempeño deficiente cuando la asistencia tecnológica se retira. Baidoo-Anu y Ansah (2023) señalan que esta problemática es particularmente aguda en disciplinas que requieren razonamiento procedimental y pensamiento abstracto, como el cálculo diferencial e integral. En respuesta, investigadores como Long y Magerko (2020) y Ng, Leung, Chu y Qiao (2021) proponen marcos de alfabetización en IA que trascienden la mera competencia técnica e incorporan dimensiones críticas, éticas y metacognitivas. Más recientemente, Sabzalieva y Valentini (2023), en el marco de la UNESCO, enfatizan

la necesidad de integrar la alfabetización en IA dentro de los currículos universitarios como condición para un uso pedagógicamente responsable.

A pesar de estos avances, persiste una brecha significativa en la literatura: la mayoría de los estudios sobre IAGen en educación superior se centran en percepciones y actitudes (Grassini, 2023), pero son escasos los que documentan intervenciones sistemáticas de alfabetización y miden su impacto en el desempeño académico real de los estudiantes, especialmente en contextos de cursos intensivos de matemáticas universitarias. La presente investigación aborda esta laguna mediante el diseño, implementación y evaluación de un proceso de alfabetización en IAGen estructurado en cinco fases: deconstrucción del oráculo, prompting pedagógico, auditoría de la IA, andamiaje diferenciado y pacto ético, en un curso de verano de Cálculo Diferencial e Integral. El objetivo general es analizar en qué medida dicho proceso de alfabetización facilita la transición del uso instrumental de la IAGen (como generador de respuestas) hacia un empleo reflexivo y autorregulado (como tutor personalizado que potencia la autonomía). Se plantea la hipótesis de que los estudiantes alfabetizados mostrarán una menor brecha entre su rendimiento asistido por IA y su desempeño autónomo en pruebas presenciales, así como una mayor capacidad de autorregulación metacognitiva en el uso de la herramienta.

### **Antecedentes y marco conceptual**

Para delimitar el objeto de estudio, esta investigación comprende la alfabetización en inteligencia artificial generativa como un conjunto de competencias que integra comprensión funcional, pensamiento crítico, ética académica y autorregulación, en consonancia con marcos de alfabetización en IA que enfatizan dimensiones críticas y metacognitivas (Long y Magerko, 2020; Ng, Leung, Chu y Qiao, 2021) y con los paradigmas de la inteligencia artificial en educación descritos por Ouyang y Jiao (2021). Desde una perspectiva socioconstructivista, el aprendizaje se concibe como mediado por herramientas culturales y por la interacción, lo cual enlaza con la idea de complementariedad entre docente humano y modelos de lenguaje (Jeon y Lee, 2023) y con enfoques de diseño instruccional que comparan rasgos del conductismo, cognitivismo y constructivismo (Ertmer y Newby, 2013). A nivel pedagógico, la integración se

interpreta como una orquestación donde convergen conocimiento disciplinar, pedagógico y tecnológico (Mishra y Koehler, 2006).

En educación superior, el campo reporta un crecimiento acelerado y heterogéneo: revisiones sistemáticas y del estado del arte advierten tanto promesas como vacíos metodológicos, especialmente en la evaluación del aprendizaje (Zawacki-Richter, Marín, Bond y Gouverneur, 2019; Crompton y Burke, 2023; Lo, 2023). Síntesis recientes discuten oportunidades y riesgos para investigación, práctica y políticas institucionales (Dwivedi et al., 2023; Meyer et al., 2023), y se han reportado evidencias experimentales de efectos en productividad en tareas académicas, con variabilidad según el tipo de actividad y el grado de guía pedagógica (Noy y Zhang, 2023). En este marco, trabajos de panorama general señalan beneficios potenciales, limitaciones y desafíos de implementación, incluyendo reacciones institucionales de disrupción y alarma (Deng y Lin, 2023; Memarian y Doleck, 2023; García-Peñalvo, 2023; Farrokhnia, Banihashem, Noroozi y Wals, 2023).

De forma específica, la literatura subraya la tensión entre apoyo al aprendizaje y riesgos de integridad académica: se discuten desafíos de plagio, detección y reconfiguración de la evaluación en presencia de modelos de lenguaje (Perkins, 2023; Khalil y Er, 2023; Rudolph, Tan y Tan, 2023; Yan et al., 2024). En paralelo, se propone orientar el uso mediante prácticas de prompting y evaluación crítica, así como mediante andamiajes que favorezcan la verificación y la comprensión (Eager y Brunton, 2023; Dahlkemper, Lahme y Klein, 2023). Asimismo, la adopción se asocia con variables motivacionales y de aceptación tecnológica (Lai, Cheung y Chan, 2023) y con condiciones de interacción en entornos en línea que afectan la relación estudiante-docente.

Este debate no se limita al ámbito educativo, en otros dominios profesionales, particularmente en el campo de la salud, revisiones sistemáticas documentan aplicaciones emergentes de la inteligencia artificial generativa en educación médica, investigación y práctica clínica, al tiempo que advierten sobre riesgos asociados a la precisión de la información, sesgos algorítmicos y desafíos éticos en su implementación (Sallam, 2023). Este escenario refuerza la necesidad de procesos de alfabetización en IA que permitan a los futuros profesionales utilizar estas herramientas de forma crítica, responsable y reflexiva.

## **Materiales y métodos**

Se adoptó un enfoque metodológico mixto con un diseño pre-post de un solo grupo, complementado con un componente cualitativo de entrevistas semiestructuradas. Esta elección se fundamentó en la naturaleza multidimensional del fenómeno estudiado, que requiere tanto la medición cuantitativa del rendimiento académico como la comprensión profunda de las experiencias subjetivas de los estudiantes con la IAGen (Creswell y Creswell, 2018). El estudio se inscribe en lo que Hernández-Sampieri y Mendoza (2018) denominan diseños mixtos de triangulación concurrente, donde los datos cuantitativos y cualitativos se recolectan de manera simultánea y se integran en la fase de interpretación.

### **Contexto y participantes**

La investigación se realizó durante el curso intensivo de verano de Cálculo Diferencial e Integral, impartido en el Instituto Tecnológico de Costa Rica durante el periodo intensivo de verano 2025–2026 (diciembre de 2025 a febrero de 2026). El curso condensó el contenido de un semestre regular (16 semanas) en seis semanas de instrucción intensiva, con cuatro sesiones semanales de tres-cuatro horas diarias. Participaron 37 estudiantes en la fase inicial (Actividad A) y 33 en la fase posterior (Actividad B), de los cuales 30 pudieron ser emparejados para el análisis comparativo. Para la escala Likert se obtuvo una muestra de 24 respondientes (14 hombres, 9 mujeres, 1 otro/prefiere no decirlo), con una edad media de 20.14 años ( $DE = 1.46$ , rango 18-24). Respecto a la experiencia previa con IA en cursos universitarios, 11 estudiantes reportaron uso frecuente, 10 algunas veces, 2 rara vez y 1 nunca. Los participantes provenían de carreras de ingeniería y ciencias, y fueron seleccionados mediante muestreo por conveniencia dado que conformaban el grupo natural del curso. El profesor del curso fungió simultáneamente como investigador principal, lo que facilitó la implementación ecológica de la intervención dentro del contexto natural de enseñanza.

### **Intervención: proceso de alfabetización en cinco fases**

La intervención se estructuró como un proceso de alfabetización en IAGen integrado al desarrollo regular del curso. Se proporcionó a los estudiantes acceso a un asistente virtual (GPT de ChatGPT) diseñado específicamente para el curso, así como a herramientas LLM de acceso gratuito (ChatGPT, Gemini, DeepSeek, Claude, entre otras). El proceso comprendió cinco fases secuenciales, diseñadas a partir de los marcos teóricos de

alfabetización en IA de Long y Magerko (2020) y Ng, Leung, Chu y Qiao (2021), adaptados al contexto del cálculo universitario.

**La Fase 1**, denominada "*Deconstrucción del oráculo*", tuvo como objetivo demostrar que la IA comete errores y que el estudiante debe mantener su criterio. Se diseñaron actividades donde los estudiantes solicitaban a la IA la resolución de problemas con errores sutiles (por ejemplo, el cálculo del límite de  $\sin(x)/x$  mediante L'Hôpital, que incurre en razonamiento circular, o la evaluación directa de integrales impropias sin verificar continuidad). El debate guiado posterior permitió a los estudiantes constatar que la IA puede acertar en el resultado, pero fallar en el método.

**La Fase 2**, "*Prompting pedagógico*", dotó a los estudiantes de plantillas de prompts organizadas en cuatro modos: tutor explicativo (para conceptos nuevos, con analogías y lenguaje accesible), andamiaje o scaffolding (solo pistas, no respuestas completas), socrático (descubrimiento guiado mediante preguntas) y verificador (contraste del trabajo propio con la evaluación de la IA).

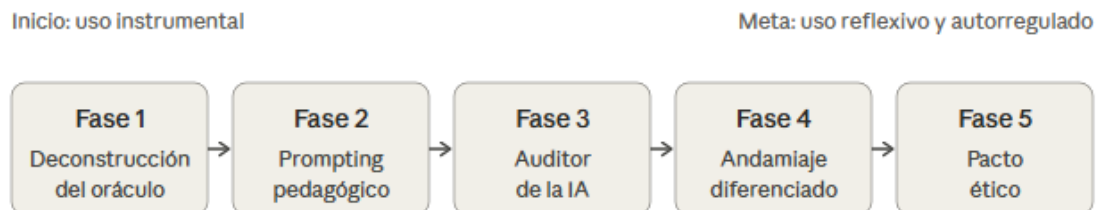
**La Fase 3**, "*Auditor de la IA*", convirtió la verificación de respuestas en una actividad gamificada donde los estudiantes "cazaban" errores deliberados o espontáneos de la IA en problemas de optimización, integración y análisis de funciones.

**La Fase 4**, "*Andamiaje diferenciado*", evidenció que un mismo contenido puede abordarse en niveles de profundidad distintos según el perfil de rendimiento del estudiante (definidos en la sección de instrumentos: X de rendimiento bajo, Z de rendimiento medio y W de rendimiento alto). Por ejemplo, el Teorema Fundamental del Cálculo se exploró con analogías concretas y cotidianas para los estudiantes de perfil Z (rendimiento medio, promedio entre 60 y 70), mientras que con los de perfil W (rendimiento alto, promedio superior a 70) se profundizó en su conexión con la integral de Riemann y en contraejemplos que delimitan sus hipótesis.

Finalmente, la **Fase 5**, "*Pacto ético*", culminó con la redacción de un contrato personal en el que cada estudiante definió sus propios límites de uso de la IA, comprometiéndose a intentar los problemas de manera autónoma antes de recurrir a la herramienta y a contrastar siempre las respuestas con fuentes confiables.

La Figura 1 sintetiza la secuencia de las cinco fases y su orientación general, que progresa desde un uso instrumental de la IA hacia un uso reflexivo y autorregulado. Cada fase se describe en detalle a continuación.

**Figura 1.** Proceso de alfabetización en IAGen en cinco fases.



**Fuente:** Elaboración propia.

### Instrumentos de recolección de datos

Se utilizaron tres instrumentos complementarios. Primero, las Actividades A y B, cada una compuesta por tres reactivos de tarea domiciliaria mediada por IAGen y tres reactivos semejantes en prueba presencial sin IAGen (escala 0-100). La Actividad A se aplicó antes de la alfabetización y la Actividad B después de esta. Segundo, una escala tipo Likert de 18 ítems (1-5) que evaluó cuatro dimensiones: utilidad para el aprendizaje, motivación e integración en clase, uso *Estudiante X* (rendimiento bajo, nota menor a 60 en el primer parcial), *Estudiante Z* (rendimiento medio, nota entre 60 y 70 en el primer parcial) y *Estudiante W* (rendimiento alto, nota mayor a 70 en el primer parcial). Las entrevistas fueron grabadas, transcritas y analizadas mediante codificación temática (Braun y Clarke, 2006).

### Análisis de datos

Para el análisis cuantitativo se calcularon estadísticos descriptivos (media, desviación estándar, mediana, cuartiles y porcentajes) por reactivo y globalmente. La comparación entre el rendimiento en tarea (con IA) y prueba (sin IA) dentro de cada actividad se realizó mediante pruebas t de Student pareadas, verificadas con la prueba no paramétrica de Wilcoxon de rangos con signo. Se reportaron tamaños de efecto mediante Cohen d para medidas repetidas (valores de referencia: 0.2 pequeño, 0.5 moderado, 0.8 grande) e intervalos de confianza del 95 %. La comparación pre-post se realizó tanto descriptivamente como mediante emparejamiento por nombre de los estudiantes presentes en ambas actividades (n = 30). La correlación tarea-prueba se evaluó mediante

el coeficiente de Pearson. Para la escala Likert se realizó un análisis descriptivo por ítems y por dimensiones, apropiado para el tamaño muestral ( $n = 24$ ). El análisis cualitativo siguió el procedimiento de análisis temático de seis fases de Braun y Clarke (2006): familiarización con los datos, generación de códigos iniciales, búsqueda de temas, revisión de temas, definición y denominación de temas, y producción del informe.

## Resultados

Los resultados se presentaron en dos secciones: hallazgos cuantitativos (rendimiento académico y percepción) y hallazgos cualitativos (experiencias de los estudiantes). Se utilizaron tres tablas para sintetizar la evidencia.

### Resultados cuantitativos

La Tabla 1 presentó los resultados comparativos por reactivo y globales de las Actividades A (pre-alfabetización) y B (post-alfabetización). En la Actividad A, se observó una brecha estadísticamente significativa en dos de los tres reactivos: A1 mostró una caída de  $-24.80$  puntos

**Tabla 1.** Comparación tarea (con IA) vs. prueba (sin IA) por reactivo y global, actividades A y B.

Indicador	Tarea (M)	Prueba (M)	Diferencia	p	dz	r
A1 (n=35)	82.29	57.49	-24.80	.001	-0.59	—
A2 (n=35)	79.43	49.89	-29.54	.001	-0.60	—
A3 (n=35)	85.71	85.26	-0.46	.953	-0.01	—
Global A (n=35)	82.48	64.21	-18.27	.003	-0.53	0.24
B1 (n=32)	91.09	78.28	-12.81	.038	-0.38	—
B2 (n=32)	92.81	75.83	-16.98	<.001	-0.65	—
B3 (n=32)	93.63	80.17	-13.45	.015	-0.45	—
Global B (n=32)	92.51	78.09	-14.42	.001	-0.60	0.355

**Fuente:** Elaboración propia

En la Actividad B (post-alfabetización), la brecha tarea-prueba persistió, pero con magnitudes menores en los reactivos B1 ( $-12.81$  puntos;  $p = .038$ ;  $dz = -0.38$ ) y B3 ( $-13.45$  puntos;  $p = .015$ ;  $dz = -0.45$ ), mientras que B2 mostró una caída mayor ( $-16.98$  puntos;  $p < .001$ ;  $dz = -0.65$ ). Globalmente, el promedio de la prueba presencial alcanzó

78.09 (frente a 64.21 en la Actividad A) y la proporción de estudiantes con puntaje inferior a 60 disminuyó del 42.9 % al 21.9 %. La correlación tarea-prueba aumentó a  $r = .355$ , indicando una mejor alineación entre el trabajo asistido y el desempeño autónomo.

La Tabla 2 sintetizó la comparación directa entre las Actividades A y B. En los 30 estudiantes emparejados por nombre que completaron ambas actividades, el promedio de la prueba presencial mejoró en +10.81 puntos ( $p = .005$ ;  $dz = 0.54$ , IC 95 %: [+3.40, +18.22]). La proporción de estudiantes con promedio de prueba inferior a 60 se redujo del 33.3 % al 20.0 %. Aunque las actividades evaluaron contenidos distintos de cálculo, se observó el mismo sentido de cambio en el promedio de la prueba presencial y en la proporción de promedios inferiores a 60.

**Tabla 2.** Comparación global de indicadores de desempeño autónomo entre actividades A y B.

Indicador	Actividad A	Actividad B
Prueba presencial (M)	64.21	78.09
% Prueba < 60	42.9 %	21.9 %
Correlación tarea-prueba ( $r$ )	.24	.355
Mejora emparejada ( $n=30$ )	—	+10.81 ( $p=.005$ ; $dz=0.54$ )

*Fuente:* Elaboración propia.

Respecto a la escala Likert ( $n = 24$ ), la Tabla 3 presentó los resultados por dimensión. La dimensión de uso crítico y metacognición obtuvo las puntuaciones más altas: el ítem "Es importante comprender lo que responde la IA" alcanzó una media de 4.92 con el 100 % de acuerdo (puntuaciones 4-5), mientras que "Siempre contrasté respuestas con mis conocimientos o clases" obtuvo una media de 4.67 con el 95.8 % de acuerdo. La dimensión de proyección futura también fue muy valorada (media 4.59), con el 95.8 % de los estudiantes indicando que la IA será útil en cursos futuros y que desean aprender a usarla mejor. La dimensión de utilidad para el aprendizaje mostró resultados altos (media 4.28), destacando que el 87.5 % valoró positivamente la ayuda de la IA para comprender conceptos e identificar errores en su razonamiento. La dimensión de motivación e integración fue la más heterogénea (media 3.81), con solo el 54.2 % concordando en que la IA motiva su interés por aprender matemáticas. Se registró que el 29.2 % de los estudiantes admitió haber aceptado respuestas de la IA sin cuestionarlas en alguna ocasión (media 2.92 en el ítem negativo correspondiente).

**Tabla 3.** Resultados de la escala Likert por dimensión (n = 24).

Dimensión	Media	% Acuerdo (4-5)
Uso crítico / metacognición	4.39	95.8 %
Proyección futura	4.59	95.8 %
Utilidad para el aprendizaje	4.28	87.5 %
Motivación e integración	3.81	54.2 %
Índice global	4.28	—

**Fuente:** Elaboración propia.

### Resultados cualitativos

El análisis temático de las entrevistas a los tres perfiles de rendimiento reveló cuatro temas principales que se describieron a continuación.

**Tema 1: La IA como andamio cognitivo personalizado.** Los tres estudiantes coincidieron en que la IAGen funcionó como un tutor personalizado capaz de adaptarse a su nivel. El Estudiante Z (rendimiento medio) expresó:

*"Mi problema no era no saber la teoría, era cómo empezar. Ahí le pregunté a la IA: 'Tengo que derivar esto. No me des la respuesta, pero dime cuál es el primer paso lógico'".*

Esta estrategia de solicitar orientación parcial en lugar de respuestas completas fue un patrón recurrente en los estudiantes alfabetizados. El Estudiante X (rendimiento bajo) relató cómo la IA le ayudó con el Teorema del Valor Medio:

*"Le pedí: 'Explícame el Teorema del Valor Medio como si tuviera 15 años, con un ejemplo de la vida real'. Me puso un ejemplo de un viaje en coche y la velocidad media, y de repente todo cobró sentido".*

Por su parte, el Estudiante W (rendimiento alto) utilizó la IA no para sobrevivir, sino para profundizar:

*"Le pregunté: 'Explícame la intuición geométrica profunda detrás del Teorema del Valor Medio, ....la IA actuó como un amplificador de mi curiosidad".*

**Tema 2: La transición del uso instrumental al uso reflexivo.** Las entrevistas revelaron una evolución en el modo de interacción con la IAGen a lo largo del curso. El Estudiante Z describió un episodio crítico de transición:

*"Una vez, le pedí que me resolviera un problema de optimización completo. Copié la respuesta y la entregué en el curso. En la siguiente clase, el profe puso un problema semejante a ése y yo no entendía de dónde salían los números. Me sentí vacío, .... desde ese día, hice el pacto conmigo mismo de que la IA sería mi tutor...".*

El Estudiante X señaló:

*"sin esa guía, probablemente habría caído en usarla solo para resolver ejercicios y ya".*

**Tema 3: El desarrollo del pensamiento crítico frente a la IA.** La fase de auditoría y la formación en detección de errores generaron una actitud escéptica constructiva. El Estudiante W reportó haber detectado un error de la IA en al menos una ocasión:

*"Le pedí a la IA que calculara una integral definida, al final me dio un resultado numérico. Por un momento pensé que estaba bien, porque la algebraica daba eso. Pero luego en clases se analizó el ejercicio y resulta que la función presentaba una discontinuidad en el intervalo de integración. Ahí entendí que la IA había aplicado una receta sin pensar. Ese día dejé de verla como un oráculo".*

El Estudiante X añadió:

*"A veces me daba respuestas muy genéricas o incluso errores", y comentó que la formación en verificación "Ahora, cuando voy a usarla, pienso: '¿Esto me está haciendo pensar o me está evitando pensar?' Si es lo segundo, borro el prompt y lo vuelvo a redactar".*

**Tema 4: Proyección de la competencia más allá del curso.** Los tres estudiantes percibieron la experiencia como transferible a otros ámbitos. El Estudiante Z afirmó:

*"Aprendí a usarla para mi bien".*

El Estudiante W elaboró:

*"No solo sé usar la herramienta, sino que entiendo sus límites y su potencial. Saber cómo plantearle un problema complejo a una IA, cómo criticar su respuesta para refinarla".*

El Estudiante X concluyó con una metáfora:

*"El asunto no está en la IA, sino en cómo la usan para construir su propio aprendizaje".*

## Discusión de resultados

Los resultados de esta investigación proporcionan evidencia convergente de que un proceso estructurado de alfabetización en IAGen facilita la transición del uso instrumental al uso reflexivo de estas herramientas en el contexto del cálculo universitario. La hipótesis central se ve respaldada por la reducción de la brecha tarea-prueba, el aumento de la correlación entre rendimiento asistido y autónomo, y los testimonios cualitativos que documentan procesos de autorregulación metacognitiva.

El hallazgo más relevante es la mejora significativa del desempeño autónomo tras la alfabetización. La mejora emparejada de +10.81 puntos ( $dz = 0.54$ ) indica un efecto moderado y educativamente significativo, consistente con los hallazgos de Wardat, Tashtoush, AlAli y Saleh (2024), quienes reportaron mejoras similares al integrar IAGen con estrategias pedagógicas explícitas en matemáticas. La reducción del porcentaje de estudiantes con rendimiento inferior a 60 (de 42.9 % a 21.9 %) tiene implicaciones prácticas relevantes: en un curso intensivo de verano, donde las tasas de reprobación suelen ser elevadas, una reducción a la mitad del grupo en riesgo resulta pedagógicamente significativa. Estos resultados contrastan con las advertencias de Mollick y Mollick (2023) sobre la "ilusión de competencia" generada por la IAGen: la fase pre-alfabetización de esta investigación confirma exactamente ese fenómeno (correlación débil  $r = .24$  entre tarea asistida y prueba presencial), pero la fase post-alfabetización muestra que es posible mitigarlo mediante una intervención pedagógica deliberada.

El incremento de la correlación tarea-prueba de  $r = .24$  a  $r = .355$  sugiere que, tras la alfabetización, el trabajo realizado con asistencia de IA se alinea mejor con el aprendizaje real del estudiante. Este hallazgo es coherente con la teoría del andamiaje cognitivo de Vygotsky (1978), actualizada al contexto de la IAGen: *cuando la herramienta se utiliza como andamio (proporcionando pistas y orientación) en lugar de como sustituto (proporcionando respuestas completas), el estudiante transita efectivamente por su zona de desarrollo próximo*. Las entrevistas corroboran este mecanismo: el Estudiante Z describió cómo pasó de solicitar soluciones completas a pedir únicamente el primer paso lógico, transformando así la interacción con la IA de una relación de dependencia a una de colaboración cognitiva.

Los resultados de la escala Likert revelan un perfil de uso predominantemente crítico y metacognitivo, con medias superiores a 4.3 en los ítems de contrastación y comprensión. Este hallazgo se alinea con lo reportado por Chan y Hu (2023), quienes encontraron que los estudiantes que reciben formación explícita en el uso de IAGen desarrollan actitudes más reflexivas. No obstante, la persistencia de un subgrupo (29.2 %) que admite aceptación acrítica ocasional de respuestas de la IA constituye una señal de alerta pedagógica que resuena con las preocupaciones de Cotton, Cotton y Shipway (2024) sobre la vulnerabilidad de los estudiantes ante respuestas convincentes pero potencialmente erróneas. Este resultado indica que la alfabetización, aunque efectiva, debe ser sostenida y reforzada a lo largo del curso, no limitarse a una intervención puntual.

La dimensión cualitativa enriquece significativamente la comprensión de los mecanismos subyacentes, los estudiantes de rendimiento alto utilizan la IAGen para ampliar horizontes conceptuales, mientras que los de rendimiento medio y bajo la emplean para acceder a explicaciones adaptadas a su nivel. Esta diferenciación es consistente con el concepto de andamiaje diferenciado propuesto por Kasneci et al. (2023) y con los hallazgos de Dai et al. (2023), quienes documentaron que la personalización mediada por IAGen beneficia particularmente a los estudiantes con mayores dificultades. El episodio relatado por el Estudiante Z: *copiar una respuesta, sentirse "como un fraude" y redefinir su relación con la herramienta*, constituye un ejemplo elocuente de lo que Zimmerman (2000) denominó fases de autorregulación del aprendizaje: la experiencia de fracaso (fase de autorreflexión) conduce a un reajuste de estrategias (fase de planificación) que se manifiesta en un uso más deliberado de la herramienta (fase de ejecución).

Es necesario señalar las limitaciones del estudio. Las Actividades A y B evaluaron contenidos distintos del curso de cálculo, por lo que la comparación directa debe interpretarse como evidencia de patrones consistentes y no como una equivalencia de dificultad. La ausencia de grupo control impide aislar completamente el efecto de la alfabetización del efecto de maduración y experiencia acumulada en el curso. El efecto techo observado en las tareas con IA (muchos puntajes de 100) limita la variabilidad y puede influir en las diferencias detectadas. Adicionalmente, el rol dual del profesor como investigador puede introducir sesgos de expectativa, aunque también garantiza la implementación fiel de la intervención en su contexto ecológico. El tamaño muestral moderado ( $n = 30$  para el análisis emparejado,  $n = 24$  para la escala Likert) limita la generalización de los resultados, si bien la triangulación con datos cualitativos proporciona robustez interpretativa.

Futuras investigaciones deberían incorporar diseños cuasi-experimentales con grupo control, evaluar la transferencia a largo plazo de las competencias de alfabetización en IA, explorar la efectividad diferencial de cada una de las cinco fases del proceso y replicar la intervención en otros cursos de matemáticas universitarias y en contextos culturales diversos. Asimismo, sería valioso investigar el impacto de la alfabetización en IAGen sobre variables motivacionales y de autoeficacia matemática, dada la heterogeneidad observada en la dimensión de motivación de la escala Likert.

## **Conclusiones**

Esta investigación aporta evidencia empírica de que un proceso de alfabetización en IAGen, estructurado en cinco fases progresivas, facilita la transición del uso instrumental al uso reflexivo de estas herramientas en un curso universitario intensivo de Cálculo Diferencial e Integral. Los hallazgos cuantitativos demuestran que la alfabetización se asocia con una mejora significativa del desempeño autónomo en pruebas presenciales sin IA (+10.81 puntos;  $d_z = 0.54$ ), una reducción a la mitad de la proporción de estudiantes en riesgo de reprobación y un fortalecimiento de la relación entre el trabajo asistido y el aprendizaje efectivo.

Los hallazgos cualitativos revelan que la alfabetización cataliza procesos de autorregulación metacognitiva: los estudiantes transitan desde la copia acrítica de respuestas hacia la formulación estratégica de prompts que solicitan orientación parcial, verificación de razonamientos propios y exploración conceptual diferenciada según su nivel de dominio. La convergencia entre los datos

Las implicaciones para la práctica educativa son claras: la mera provisión de herramientas de IAGen a los estudiantes, sin un proceso explícito de alfabetización, puede generar una ilusión de competencia que se desmorona cuando la asistencia tecnológica se retira. Por el contrario, una intervención que integre la deconstrucción de la falibilidad de la IA, la enseñanza explícita de estrategias de prompting, la práctica de auditoría crítica, la diferenciación por perfiles de aprendizaje y la promoción de un compromiso ético personal configura las condiciones para que la IAGen funcione como un genuino tutor personalizado y no como un atajo cognitivo. En un momento en que las universidades debaten entre prohibir o adoptar la IAGen, este estudio sugiere que la clave no reside en la herramienta, sino en la pedagogía que la acompaña.

### **Agradecimientos**

El autor expresa su sincero agradecimiento a los estudiantes del curso intensivo de verano de Cálculo Diferencial e Integral 2025, por su participación voluntaria, entusiasta y comprometida en este estudio. Su disposición para compartir sus experiencias de aprendizaje con la inteligencia artificial fue fundamental para el desarrollo de esta investigación. Se extiende un reconocimiento especial a la Red Académica MatIA, cuyo acompañamiento intelectual, espacios de discusión y análisis crítico enriquecieron significativamente el planteamiento teórico y metodológico del trabajo.

### **Referencias Bibliográficas**

Baidoo-Anu, D., y Ansah, L. O. (2023). *Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning*. *Journal of AI*, 7(1), 52-62. Recuperado de <https://dergipark.org.tr/en/pub/jai/issue/77844/1337500>

- Braun, V., y Clarke, V. (2006). *Using thematic analysis in psychology. Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Chan, C. K. Y., y Hu, W. (2023). *Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. International Journal of Educational Technology in* , 20(1), 43. <https://link.springer.com/article/10.1186/s41239-023-00411-8>
- Cotton, D. R. E., Cotton, P. A., y Shipway, J. R. (2024). *Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International*, 61(2), 228-239. <https://www.tandfonline.com/doi/full/10.1080/14703297.2023.2190148>
- Creswell, J. W., y Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Thousand Oaks, CA: SAGE Publications.
- Crompton, H., y Burke, D. (2023). *Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education*, 20(1), 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Dahlkemper, M. N., Lahme, S. Z., y Klein, P. (2023). *How do physics students evaluate artificial intelligence responses on comprehension questions? A study on the impact of AI literacy. Physical Review Physics Education Research*, 19(1), 010142. Recuperado de <https://journals.aps.org/prper/abstract/10.1103/PhysRevPhysEducRes.19.010142>
- Dai, Y., Liu, A., Dai, D., Cao, Y., Li, S., y Zhang, C. (2023). *Reconceptualizing ChatGPT and generative AI as a student-driven innovation in higher education. Journal of Educational Technology & Society*, 26(3), 201-215. <https://doi.org/10.1016/j.procir.2023.05.002>
- Deng, J., y Lin, Y. (2023). *The benefits and challenges of ChatGPT: An overview. Frontiers in Computing and Intelligent Systems*, 2(2), 81-83. <https://doi.org/10.54097/fcis.v2i2.4465>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... y Wright, R. (2023). *"So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Eager, B., y Brunton, R. (2023). *Prompting higher education towards AI-augmented teaching and learning practice. Journal of University Teaching & Learning Practice*, 20(5), 1-19. <https://doi.org/10.53761/1.20.5.02>
- Ertmer, P. A., y Newby, T. J. (2013). *Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. Performance*

- Improvement Quarterly*, 26(2), 43-71. Recuperado de <https://onlinelibrary.wiley.com/doi/10.1002/piq.21143>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., y Wals, A. (2023). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460-474. <https://doi.org/10.1080/14703297.2023.2195846>
- García-Peñalvo, F. J. (2023). The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic? *Education in the Knowledge Society*, 24, e31279. <https://doi.org/10.14201/eks.31279>
- Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692. Recuperado de <https://www.mdpi.com/2227-7102/13/7/692>
- Hernández-Sampieri, R., y Mendoza, C. P. (2018). *Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta*. Ciudad de México, México: McGraw Hill.
- Jeon, J., y Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies*, 28(12), 15873-15892. Recuperado de <https://link.springer.com/article/10.1007/s10639-023-11834-1>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... y Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. Recuperado de <https://doi.org/10.1016/j.lindif.2023.102274>
- Khalil, M., y Er, E. (2023). Will ChatGPT get you caught? Rethinking of plagiarism detection. En Y. Ono y T. Kano (Eds.), *Proceedings of the International Conference on Human-Computer Interaction* (pp. 475-487). Cham: Springer. Recuperado de [https://link.springer.com/chapter/10.1007/978-3-031-34411-4\\_32](https://link.springer.com/chapter/10.1007/978-3-031-34411-4_32)
- Lai, C. Y., Cheung, K. Y., y Chan, C. S. (2023). Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: An extension of the technology acceptance , 5, 100178. <https://doi.org/10.1016/j.caeai.2023.100178>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Rajendran, J. A., y Pham, N. M. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *International Journal of Management Education*, 21(2), 100790. Recuperado de <https://www.sciencedirect.com/science/article/pii/S1472811723000289>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- Long, D., y Magerko, B. (2020). What is AI literacy? Competencies and design considerations. En CHI '20: Proceedings of the 2020 CHI Conference on Human

- Factors in Computing Systems (pp. 1-16). New York: ACM. Recuperado de <https://dl.acm.org/doi/10.1145/3313831.3376727>
- Memarian, B., y Doleck, T. (2023). *ChatGPT in education: Methods, potentials, and limitations. Computers in Human Behavior: Artificial Humans*, 1(2), 100022. <https://doi.org/10.1016/j.chbah.2023.100022>
- Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P. C., ... y Moore, J. H. (2023). *ChatGPT and large language models in academia: Opportunities and challenges. BioData Mining*, 16(1), 20. <https://doi.org/10.1186/s13040-023-00339-9>
- Mishra, P., y Koehler, M. J. (2006). *Technological pedagogical content knowledge: A framework for teacher knowledge. Teachers College Record*, 108(6), 1017-1054. Recuperado de <https://journals.sagepub.com/doi/10.1111/j.1467-9620.2006.00684.x>
- Mollick, E. R., y Mollick, L. (2023). *Assigning AI: Seven approaches for students, with prompts. The Wharton School Research Paper*. Recuperado de <https://arxiv.org/abs/2306.10052>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., y Qiao, M. S. (2021). *Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence*, 2, 100041. Recuperado de <https://www.sciencedirect.com/science/article/pii/S2666920X21000357>
- Noy, S., y Zhang, W. (2023). *Experimental evidence on the productivity effects of generative artificial intelligence. Science*, 381(6654), 187-192. Recuperado de <https://www.science.org/doi/10.1126/science.adh2586>
- Ouyang, F., y Jiao, P. (2021). *Artificial intelligence in education: The three paradigms. Computers and Education: Artificial Intelligence*, 2, 100020. Recuperado de <https://www.sciencedirect.com/science/article/pii/S2666920X2100014X>
- Perkins, M. (2023). *Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. Journal of University Teaching & Learning Practice*, 20(2), 1-24. <https://doi.org/10.53761/1.20.02.07>
- Rudolph, J., Tan, S., y Tan, S. (2023). *ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Sabzalieva, E., y Valentini, A. (2023). *ChatGPT and artificial intelligence in higher education: Quick start guide*. París: UNESCO, International Institute for Higher Education in Latin America and the Caribbean. Recuperado de <https://unesdoc.unesco.org/ark:/48223/pf0000385146>
- Sallam, M. (2023). *ChatGPT utility in healthcare education, research, and practice: Systematic review on the promising perspectives and valid concerns. Healthcare*, 11(6), 887. <https://doi.org/10.3390/healthcare11060887>

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
- Wardat, Y., Tashtoush, M., AlAli, R., y Saleh, S. (2024). *Artificial intelligence in education: Mathematics teachers' perspectives, practices, and challenges*. *Iraqi Journal for Computer Science and Mathematics*, 5(1), 60-77. Recuperado de: <https://ijcsm.researchcommons.org/ijcsm/vol5/iss1/20/>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., ... y Gašević, D. (2024). *Practical and ethical challenges of large language models in education: A systematic scoping review*. *British Journal of Educational Technology*, 55(1), 90-112. <https://doi.org/10.1111/bjet.13370>
- Zawacki-Richter, O., Marín, V. I., Bond, M., y Gouverneur, F. (2019). *Systematic review of research on artificial intelligence applications in higher education — Where are the , 16(1), 39*. <https://doi.org/10.1186/s41239-019-0171-0>
- Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. En M. Boekaerts, P. R. Pintrich y M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13-39). San Diego, CA: Academic Press. <https://doi.org/10.1016/B978-012109890-2/50031-7>

**Conflicto de intereses:** El autor declara que no existe conflicto de interés.

**Conflicto de intereses:**

Los autores declaran que no existe conflicto de interés