

Alfabetización crítica en IA generativa: cómo cambian las percepciones de futuros docentes tras una experiencia formativa

Helcio Soares Padilha^a 

Víctor López-Simó^b 

Agostinho Serrano^c 

Fabiane Fischer Figueiredo^a 

^a Universidade Luterana do Brasil, Programa de Pós-Graduação em Ensino de Ciências e Matemática, Canoas, Rio Grande do Sul, Brasil.

^b Universitat Autònoma de Barcelona, Departamento de Didáctica de la Matemática y las Ciencias Experimentales, Cerdanyola del Vallès, Cataluña, España.

^c Universidade de Caxias do Sul, Programa de Pós-Graduação em Ensino de Ciências e Matemática, Caxias do Sul, Rio Grande do Sul, Brasil.

RESUMEN

Contexto: la Inteligencia Artificial Generativa plantea nuevas demandas a la formación inicial de docentes de Ciencias y Matemáticas en planificación, evaluación crítica, ética y uso pedagógico responsable. **Objetivos:** analizar cómo cambian las percepciones de futuros docentes sobre IA generativa después de una experiencia formativa centrada en uso pedagógico, crítico y ético. Se examina también qué prácticas son percibidas como relevantes para la docencia y cómo perfiles de uso previo contextualizan estos cambios. **Diseño:** estudio cuantitativo con comparación pre-postaller y análisis exploratorio de perfiles de uso. **Ambiente y participantes:** 137 docentes en formación inicial con respuestas pareadas válidas. **Recolección y análisis de datos:** cuestionarios pre y postaller analizados con estadística descriptiva, prueba de Wilcoxon y análisis de conglomerados. **Resultados:** todos los ítems mostraron cambios estadísticamente significativos ($p < .001$). Los mayores fueron en ética ($\Delta=0.83$), reflexión epistemológica ($\Delta=0.73$) y comprensión de diferencias entre IA y búsqueda en internet ($\Delta=0.66$). Hallazgo notable: la utilidad percibida como estudiante universitario (media 4.49) fue superior a la percibida como docente (media 4.16), indicando cautela pedagógica ante la adopción. Las prácticas más indicadas para uso futuro fueron generación de actividades (86.1%) y creación de planes de clase (56.9%). **Conclusiones:** la experiencia formativa generó percepciones críticas y éticas sobre IA generativa. Los resultados subrayan la importancia de integrar ChatGPT en formación docente inicial como objeto de reflexión y decisión pedagógica fundamentada, no solo como herramienta técnica.

Palabras clave: Inteligencia Artificial Generativa; ChatGPT; formación inicial docente; alfabetización en IA; teoría de la difusión de innovaciones.

Autor de correspondencia: Helcio Soares Padilha Junior.

E-mail: helciopadilha@gmail.com

RESUMO

Contexto: a Inteligência Artificial Generativa coloca novas demandas à formação inicial de docentes de Ciências e Matemática no que diz respeito à planificação, avaliação crítica, ética e uso pedagógico responsável. **Objetivos:** analisar como mudam as percepções de futuros docentes sobre IA generativa após uma experiência formativa centrada em uso pedagógico, crítico e ético. Examina-se também quais práticas são percebidas como mais relevantes para a docência e como perfis de uso prévio de IA contextualizam estas mudanças. **Design:** estudo quantitativo com comparação pré e pós-oficina e análise exploratória de perfis de uso. **Ambiente e participantes:** 137 docentes em formação inicial com respostas pareadas válidas. **Coleta e análise de dados:** questionários pré e pós-oficina analisados com estatística descritiva, teste de Wilcoxon e análise de conglomerados. **Resultados:** todos os itens mostraram mudanças estatisticamente significativas ($p < .001$). As maiores foram em ética ($\Delta=0.83$), reflexão epistemológica ($\Delta=0.73$) e compreensão de diferenças entre IA e busca na internet ($\Delta=0.66$). Achado notável: a utilidade percebida como estudante universitário (média 4.49) foi superior à percebida como docente (média 4.16), indicando cautela pedagógica na adoção. As práticas mais indicadas para uso futuro foram geração de atividades (86.1%) e criação de planos de aula (56.9%). **Conclusões:** a experiência formativa gerou percepções críticas e éticas sobre IA generativa. Os resultados ressaltam a importância de integrar ChatGPT na formação docente inicial como objeto de reflexão e decisão pedagógica fundamentada, não apenas como ferramenta técnica.

Palavras-chave: Inteligência Artificial generativa; ChatGPT; formação inicial de professores; alfabetização em IA; teoria de difusão de inovações.

INTRODUCCIÓN

La Inteligencia Artificial (IA) generativa ocupa un lugar central en debates actuales sobre formación docente, especialmente en Ciencias y Educación Matemática. Los futuros profesores necesitan comprender no solo las posibilidades pedagógicas de estas herramientas, sino también sus límites conceptuales, éticos y didácticos. El uso responsable de ChatGPT y similares requiere que el docente maneje criterios de evaluación, revisión y adecuación pedagógica.

La formación inicial docente constituye un espacio estratégico para investigar la integración de tecnologías emergentes. Las creencias y experiencias que los futuros profesores construyen durante esta formación influyen en cómo incorporarán tecnologías en el aula (Zhang & Zhu, 2023). Con IA generativa, esta relevancia es doble: los futuros docentes la utilizan

como estudiantes universitarios y deberán decidir si adoptarla como docentes (Bae *et al.*, 2024).

El campo de IA generativa en formación docente está en expansión pero acompañado de incertidumbres sobre su uso pedagógico, ético e institucional. Bae *et al.* (2024) documentaron que experiencias con ChatGPT amplían conciencia y reducen ansiedad, aunque pocos participantes confirmaron intención de adopción pedagógica. Panday-Shukla (2025) halló beneficios similares en generación de ideas y eficiencia temporal, acompañados de preocupaciones sobre integridad académica, aprendizaje superficial y falta de orientaciones institucionales.

En Ciencias y Educación Matemática, la incorporación de IA demanda atención específica. Blonder *et al.* (2025) muestran que interacciones con herramientas generativas revelan conocimiento pedagógico del contenido en planificación, diseño de actividades y evaluación. Dilling & Herrmann (2024) documentan que futuros profesores de Matemática tienden a usar modelos de lenguaje de forma poco específica, con dificultades para evaluar la calidad matemática de respuestas generadas. Estos estudios subrayan que la formación docente debe ir más allá de demostrar herramientas y favorecer experiencias de uso crítico, situado y pedagógicamente fundamentado.

La alfabetización en IA ofrece un marco importante para esta discusión. Chiu *et al.* (2024) plantean que la formación debe incluir conocimiento sobre tecnología, impacto, ética, colaboración y autorreflexión. Esto significa que usar herramientas generativas requiere más que escribir *prompts* efectivos: implica reconocer limitaciones, identificar sesgos y evaluar confiabilidad de respuestas. Meegan & Young (2025) documentan que experiencias prácticas que combinan experimentación, reflexión y análisis crítico favorecen esta alfabetización en futuros docentes.

Este estudio adopta la Teoría de la Difusión de Innovaciones (TDI) (Rogers, 2003) para analizar cómo futuros docentes evalúan ChatGPT. Rogers (2003) identifica cinco atributos clave que median la adopción de una innovación: ventaja relativa, compatibilidad, complejidad, experimentabilidad y observabilidad. Estos atributos ayudan a comprender la utilidad percibida, las cautelas y las condiciones pedagógicas que median la adopción de IA generativa.

Este estudio analiza cambios en percepciones de futuros docentes de Ciencias y Matemáticas después de una experiencia formativa sobre uso pedagógico, crítico y ético de ChatGPT. Participaron 137 docentes en

formación inicial. El taller incluyó discusiones sobre usos responsables, demostraciones prácticas de planificación y creación de actividades, y exploración de posibilidades disciplinares específicas para Matemática, como el análisis de demanda cognitiva de tareas.

La pregunta central es: ¿cómo cambian las percepciones de futuros docentes sobre IA generativa después de una experiencia formativa centrada en su uso pedagógico, crítico y ético? El estudio también examina qué prácticas fueron percibidas como más relevantes para la docencia y cómo perfiles de uso previo contextualizan estos cambios.

La contribución central es articular formación inicial de docentes de Ciencias y Matemáticas con alfabetización en IA y TDI. El estudio documenta cómo experiencias formativas breves pueden generar cambios significativos en percepciones sobre utilidad, ética, límites y confianza en interacción con IA. Contribuye así a conceptualizar IA generativa como objeto de formación crítica y decisión pedagógica fundamentada en campos disciplinares específicos.

MARCO TEÓRICO

La formación inicial docente constituye un espacio estratégico para investigar cómo integrar tecnologías emergentes en la enseñanza de Ciencias y Educación Matemática. El desafío central es construir criterios pedagógicos, disciplinares y éticos para decidir cuándo, cómo y por qué utilizar herramientas digitales. En la educación STEM, Zhang & Zhu (2023) muestran que la formación de futuros docentes ha sido investigada en relación con programas formativos, prácticas pedagógicas, tecnologías educativas y enfoques interdisciplinarios. Esta perspectiva sitúa la preparación docente como un proceso que articula conocimiento disciplinar, diseño pedagógico y uso situado de tecnologías.

En la Educación Matemática, los estudios sobre integración tecnológica indican que el acceso a herramientas digitales requiere acompañamiento formativo. Vaillant *et al.* (2020), al analizar plataformas y herramientas digitales para la enseñanza de la Matemática en Uruguay, identificaron baja frecuencia de uso docente, aunque recursos como GeoGebra y plataformas adaptativas aparecieron como herramientas relevantes. Dos Santos *et al.* (2025), en una experiencia de formación de profesores de Matemática en Cabo Verde, también destacaron que la integración de tecnologías exige práctica, apoyo continuado y reflexión sobre la relación entre tecnología, pedagogía y contenido. Estos estudios refuerzan la necesidad de

comprender la adopción tecnológica como parte de un proceso formativo, situado y progresivo.

El marco TPACK contribuye a esta discusión al proponer que la integración tecnológica requiere articular conocimiento tecnológico, pedagógico y disciplinar (Koehler & Mishra, 2008). En el campo de la enseñanza de Ciencias, Blonder *et al.* (2025) actualizan esta cuestión al vincular la IA generativa con el conocimiento pedagógico del contenido, especialmente en la planificación de clases y en el análisis de interacciones entre docentes y sistemas generativos:

La constatación de que los programas de formación docente deben incluir la integración de herramientas de GenAI en la educación científica es crucial para garantizar que los profesores en formación inicial estén dotados de las habilidades necesarias para utilizar eficazmente estas herramientas en su desarrollo profesional y en sus necesidades personales de aprendizaje. (Blonder *et al.*, 2025, p. 1303, traducción propia)

La incorporación de IA generativa en la enseñanza de Ciencias y Matemáticas exige atención específica a aspectos disciplinares y pedagógicos. En Ciencias, estos usos pueden incluir planificación, generación de actividades, evaluación crítica y retroalimentación. Sin embargo, exigen precisión conceptual, adecuación didáctica y el papel mediador del profesor. En Matemáticas, requieren análisis de la calidad de las tareas, el nivel de demanda cognitiva, la resolución de problemas y las formas de argumentación.

Dilling & Herrmann (2024), al estudiar el uso de ChatGPT por futuros profesores de Matemática en la construcción de pruebas geométricas, documentan una brecha importante. Muchos estudiantes poseen experiencias limitadas con modelos de lenguaje y tienden a usarlos de forma poco específica:

En el estudio se hizo evidente que muchos estudiantes tienen concepciones erróneas sobre el principio de funcionamiento de la IA generativa y de los grandes modelos de lenguaje. Además, muchos no utilizan técnicas de prompting ni las capacidades conversacionales de los grandes modelos de lenguaje. (Dilling & Herrmann, 2024, p. 13, traducción propia)

Esta brecha sugiere que la familiaridad superficial con una herramienta no garantiza un uso matemáticamente productivo. La formación inicial docente requiere actividades de experimentación, análisis de respuestas, comparación

de resultados y formulación de *prompts* para que el futuro docente desarrolle criterios de evaluación y comprenda los límites de estas herramientas.

¿Por qué algunos docentes adoptan innovaciones tecnológicas y otros no? La TDI de Rogers (2003) ofrece un marco para comprender este fenómeno. Rogers define la difusión como un proceso social fundamental:

La difusión es el proceso mediante el cual una innovación se comunica, a través de determinados canales y a lo largo del tiempo, entre los miembros de un sistema social. Es un tipo especial de comunicación, ya que los mensajes se relacionan con nuevas ideas. (Rogers, 2003, p. 5, traducción propia)

La IA generativa puede interpretarse como una innovación en el contexto de la formación docente, con características de comunicación, temporalidad y variabilidad según el sistema social en el que se difunde. Su adopción depende de experiencias previas, valores, oportunidades de uso y formas de comunicación dentro del contexto educativo. Rogers (2003) enfatiza que la tasa de adopción de una innovación se relaciona con atributos percibidos. Los atributos específicos son ventaja relativa (¿es mejor que alternativas existentes?), compatibilidad (¿encaja con valores previos?), complejidad (¿qué tan fácil es de usar?), experimentabilidad (¿se puede probar sin riesgo?) y observabilidad (¿son visibles los resultados?). Un hallazgo central de la TDI es que: “Las percepciones de los individuos sobre los atributos de una innovación, y no los atributos clasificados objetivamente por expertos o agentes de cambio, afectan su tasa de adopción” (Rogers, 2003, p. 6, traducción propia).

Esta proposición de Rogers (2003) sitúa la percepción individual como elemento clave. La adopción de una innovación depende de cómo los sujetos la interpretan en sus contextos específicos, no de características objetivas de la innovación misma. En la educación, esto significa que la aceptación de herramientas como ChatGPT está mediada por percepciones sobre su utilidad, compatibilidad con la práctica pedagógica y complejidad de su manejo.

Bae *et al.* (2024) aplican la TDI para analizar percepciones de futuros docentes sobre IA generativa. Sus resultados muestran que después de discusiones sobre ChatGPT, los futuros docentes ampliaron su conciencia y conocimiento inicial de la herramienta, además de reducir la ansiedad frente a la IA. Sin embargo, pocos participantes confirmaron intención de adoptar herramientas generativas en sus prácticas. Los autores observan una brecha entre interés inicial y adopción pedagógica, aspecto que ha sido documentado en otras investigaciones sobre adopción tecnológica en educación.

Panday-Shukla (2025) también adopta la TDI para analizar percepciones de futuros docentes y formadores de profesores sobre IA generativa. Sus resultados indican que la adopción se vincula con percepciones de beneficio, compatibilidad con valores pedagógicos, esfuerzo requerido para aprender a usar la herramienta, posibilidad de experimentación y visibilidad de resultados. Futuros docentes reconocen ventajas como generación de ideas, apoyo al aprendizaje y economía de tiempo. Simultáneamente, expresan preocupaciones sobre integridad académica, aprendizaje superficial, dependencia excesiva de la herramienta y falta de orientaciones institucionales claras. La TDI permite integrar estas percepciones como componentes de un proceso dinámico de adopción.

La alfabetización en IA complementa el marco de la TDI al especificar qué conocimientos, habilidades y disposiciones son relevantes para que los docentes se relacionen críticamente con estas tecnologías. Chiu *et al.* (2024) diferencian alfabetización y competencia en IA:

La competencia en IA se define como “la confianza y la capacidad de una persona para explicar claramente cómo funcionan las tecnologías de IA y cómo impactan en la sociedad, así como para utilizarlas de manera ética y responsable, y comunicarse y colaborar eficazmente con ellas en cualquier contexto”. (Chiu *et al.*, 2024, p. 4, traducción propia)

La alfabetización en IA implica múltiples dimensiones: conocimiento sobre cómo funcionan las tecnologías de IA, comprensión de sus impactos sociales, capacidad de utilizarlas éticamente, conciencia de sus limitaciones y sesgos, y disposición para colaborar de forma responsable. Estas dimensiones trascienden el dominio técnico de una herramienta específica para incluir juicio crítico, reflexión ética y conciencia de contexto.

Meegan & Young (2025) documentan que la alfabetización en IA puede desarrollarse a través de experiencias formativas estructuradas. Los autores analizaron una intervención con IA generativa basada en aprendizaje por escenarios para futuros docentes e identificaron avances en autoeficacia y alfabetización en IA. La intervención incluyó participación activa, retroalimentación reflexiva y discusión ética sobre usos responsables. Los autores destacan el potencial de experiencias formativas prácticas para desarrollar confianza y criterios de evaluación en relación con tecnologías generativas.

La convergencia entre formación inicial docente, enseñanza de Ciencias y Educación Matemática, alfabetización en IA y TDI ofrece un marco multidimensional para investigar cómo los futuros docentes se aproximan a tecnologías emergentes. Comprender esta aproximación requiere examinar percepciones sobre utilidad y compatibilidad con valores pedagógicos (elementos de TDI), conocimiento sobre funcionamiento y limitaciones de las herramientas (elementos de alfabetización en IA), y articulación entre conocimiento disciplinar, pedagógico y tecnológico (elementos de TPACK). Estos marcos teóricos permiten interpretar la adopción de innovaciones tecnológicas como un proceso complejo, situado en contextos específicos de formación docente, donde múltiples dimensiones (percepciones, conocimientos, habilidades y valores) interactúan.

METODOLOGÍA

La investigación se desarrolló mediante un diseño cuantitativo, exploratorio y descriptivo, con aplicación de cuestionarios antes y después de una experiencia formativa sobre IA generativa (Gil, 2002; Flick, 2013). El estudio analizó cambios en las percepciones autoinformadas de futuros docentes sobre el uso pedagógico, crítico y ético de ChatGPT. Por tratarse de una comparación pre y post taller sin grupo de control, los resultados se interpretan como cambios observados en las respuestas de los participantes después de la experiencia formativa, sin atribución de causalidad.

El estudio se realizó con futuros docentes de educación primaria en asignaturas de Matemáticas y Ciencias Naturales vinculados a la Universitat Autònoma de Barcelona, durante el año 2024. Se consideraron 137 participantes con respuestas pareadas en los cuestionarios pre y post taller, pareadas mediante el NIU (Número de Identificación Universitaria), utilizado exclusivamente para vincular respuestas del mismo participante en ambos momentos, preservando anonimato según Fowler (2014). La investigación se inscribe en un proyecto más amplio sobre incorporación de IA generativa en prácticas pedagógicas de profesores y futuros profesores de Ciencias y Matemáticas en Brasil y España.

La experiencia formativa fue un taller presencial de dos horas, orientado a demostrar posibilidades de uso consciente, responsable y pedagógico de ChatGPT. El taller se organizó en cuatro momentos. El primero (veinte minutos) abordó el uso responsable de la IA, incluyendo verificación de precisión, adaptación de contenidos al contexto, cuestiones éticas como plagio y autoría, y la idea de que la IA es herramienta de apoyo, no sustituta del profesor. El segundo momento (cuarenta minutos) presentó estudios de caso y

demostraciones de ChatGPT en tareas docentes: creación de planes de clase, generación de actividades y ejercicios, producción de textos introductorios, uso como tutor, y generación de retroalimentación personalizada. El tercer momento fue dedicado a actividades prácticas, en las que los participantes utilizaron ChatGPT para experimentar planificación y producción didáctica, incluyendo creación de planes de clase, materiales interactivos, evaluaciones, retroalimentación, corrección de pruebas, guiones de clase y adaptación de contenidos para estudiantes con necesidades específicas. En Educación Matemática, se incorporaron ejemplos específicos sobre resolución de problemas mediante el método de Polya y análisis de demanda cognitiva de tareas. El cuarto momento consistió en una discusión final y aplicación del cuestionario postaller.

Los instrumentos de recolección de datos fueron dos cuestionarios en línea elaborados en Google Forms: uno pretaller y otro postaller. El cuestionario pretaller incluyó un bloque sobre experiencia previa con herramientas de IA generativa, con preguntas sobre frecuencia de uso, tiempo de interacción regular, herramientas utilizadas (ChatGPT, Microsoft Bing Chat, Google Gemini y otras), uso de extensiones o complementos, tareas realizadas con IA, experiencia con comunidades de aprendizaje y seguimiento de actualizaciones sobre IA. El cuestionario postaller incluyó un bloque de evaluación del desarrollo del taller, con ítems sobre relevancia del contenido, contribución de dinámicas para comprensión del tema y comprensión de funcionamiento y funcionalidades de ChatGPT. También incluyó preguntas de selección múltiple sobre las prácticas demostradas que más llamaron la atención y aquellas que los participantes usarían en su futuro trabajo docente.

Ambos cuestionarios compartían ocho ítems centrales de percepción, derivados del contenido específico del taller, que permitieron la comparación pre y postaller. Estos ítems abordaban: confianza para formular preguntas específicas o prompts; consideración de limitaciones de herramientas de IA; atención a sesgos lingüísticos, de género, sociales o socioculturales; comprensión de la diferencia entre información generada por IA y búsquedas generales en internet; conocimiento de pautas éticas relacionadas con el uso de IA; impacto de la IA en la forma de pensar sobre conocimiento, aprendizaje y resolución de problemas; utilidad de la IA para el aprendizaje universitario; y percepción de la IA como herramienta relevante para el futuro ejercicio docente. Las respuestas fueron registradas en escala Likert de cinco puntos, posteriormente codificada de 1 a 5 para el análisis estadístico.

El análisis de datos se realizó mediante el software jamovi (versión 2.5.4, The jamovi project, 2024) y se estructuró en etapas que corresponden con la presentación de resultados. En primer lugar, se realizó un análisis exploratorio de conglomerados para identificar perfiles iniciales de uso de IA generativa entre los participantes. Para ello, se utilizaron variables del cuestionario pretaller relacionadas con frecuencia general de uso, tiempo de uso, uso de ChatGPT, uso de otras herramientas de IA, uso de extensiones, experiencia con comunidades de aprendizaje y seguimiento de actualizaciones. Debido a que estas variables presentaban escalas distintas, fueron estandarizadas antes del análisis. El análisis exploratorio de conglomerados jerárquicos se realizó con distancia euclidiana y método Ward.D2, elegido después de comparar distintos métodos como el más adecuado para esta configuración de datos. Los conglomerados fueron interpretados como perfiles exploratorios de uso previo, sin pretensión de constituir una clasificación definitiva ni equivaler directamente a categorías de la TDI.

En segundo lugar, se tabularon las respuestas de selección múltiple del cuestionario postaller para identificar las prácticas demostradas que más llamaron la atención y aquellas que los participantes declararon que utilizarían en su trabajo docente. Este análisis permitió caracterizar qué posibilidades pedagógicas de ChatGPT resultaron más relevantes o cercanas a la futura práctica docente de los participantes.

En tercer lugar, se consolidaron las bases pre y postaller, preservando el pareamiento por NIU, y se calcularon estadísticas descriptivas para caracterizar las respuestas. Posteriormente, se compararon los ocho ítems comunes mediante la prueba de rangos con signo de Wilcoxon para muestras pareadas, considerando el carácter ordinal de los ítems tipo Likert y la estructura pre/post de los datos. También se calculó el tamaño del efecto mediante correlación *rank-biserial* pareada (Kerby, 2014), con el objetivo de estimar la magnitud de las diferencias observadas.

En cuarto lugar, para examinar cómo los cambios en las percepciones variaban según el perfil inicial de cada participante, se realizó un análisis de covarianza (ANCOVA) estratificado por los seis perfiles identificados, utilizando los valores pré-taller como covariables para controlar las diferencias iniciales. Este análisis permitió explorar si la experiencia formativa impactó de manera diferenciada según el nivel de familiaridad previo con herramientas de IA.

Respecto a las herramientas de IA distintas de ChatGPT, se observó que durante el período de la investigación surgían herramientas variadas de IA

generativa (como Microsoft Bing Chat, Google Gemini y otras plataformas de síntesis de textos y análisis de información). Dado que el uso de estas herramientas fue esporádico y la mayoría de los participantes se enfocaron en ChatGPT, estas fueron agrupadas en una única variable dicotómica para facilitar el análisis de conglomerados.

La investigación siguió procedimientos éticos relativos a participación voluntaria, consentimiento informado, anonimato y confidencialidad. Antes de responder los cuestionarios, los participantes accedieron a un término de consentimiento libre e informado en línea, en el que se explicitaban los objetivos del estudio, la voluntariedad de la participación, el uso del NIU únicamente para pareamiento de respuestas y la protección de identidad. El Comité de la Universidade Luterana do Brasil aprobó la investigación, con dictamen ético número 76342523.3.0000.5349.

RESULTADOS Y ANÁLISIS

Esta sección presenta los resultados en cuatro niveles: perfiles exploratorios, prácticas demostradas, cambios pre-post en percepciones, y cambios diferenciados según perfil inicial. El análisis fue realizado con 137 respuestas pareadas. Los ítems de percepción fueron codificados de 1 a 5, donde valores más altos indican mayor acuerdo con la afirmación.

Perfiles exploratorios de uso previo de IA generativa

El análisis exploratorio de conglomerados identificó seis perfiles iniciales de uso de IA generativa, presentados en la Tabla 1. La distribución mostró que 51 participantes (37.2%) correspondieron al perfil de uso moderado centrado en ChatGPT, representando el grupo más numeroso. Un segundo grupo de 32 participantes (23.4%) presentó uso diversificado de herramientas de IA generativa. Los 54 participantes restantes se distribuyeron entre perfiles de menor experiencia: experiencia baja o nula (9 participantes, 6.6%), uso inicial centrado en ChatGPT (18 participantes, 13.1%), uso con apoyo formativo (11 participantes, 8.0%) y uso ampliado con extensiones (16 participantes, 11.7%).

Tabla 1

Perfiles exploratorios de uso previo de IA generativa

Perfil	n	%	Descripción interpretativa
Experiencia baja o nula	9	6.6	Participantes con una experiencia previa muy limitada con herramientas de IA generativa

Uso inicial centrado en ChatGPT	18	13.1	Participantes con contacto inicial u ocasional con ChatGPT
Uso de ChatGPT con apoyo formativo	11	8.0	Participantes que utilizaron ChatGPT y declararon haber aprendido mediante tutoriales, comunidades u otros recursos de apoyo
Uso moderado centrado en ChatGPT	51	37.2	Participantes que utilizaron ChatGPT, aunque con un uso limitado de otras herramientas de IA o extensiones
Uso ampliado con extensiones	16	11.7	Participantes con patrones de uso más avanzados, incluyendo extensiones o recursos complementarios
Uso diversificado de herramientas de IA	32	23.4	Participantes que utilizaron ChatGPT y también declararon usar otras herramientas de IA
Total	137	100.0	

Esta distribución muestra una heterogeneidad clara en la experiencia previa de los participantes con herramientas de IA generativa. Mientras que aproximadamente 60% del grupo (82 participantes) presentó experiencia moderada a ampliada, aproximadamente 40% (55 participantes) se ubicó en etapas iniciales o de baja experiencia con estas tecnologías. Esta composición heterogénea del grupo caracteriza la muestra de participantes que experimentó la intervención formativa.

Prácticas demostradas en el taller

El cuestionario después del taller incluyó preguntas sobre las prácticas demostradas que más llamaron la atención de los participantes y aquellas que utilizarían en su futuro trabajo docente. Estas prácticas estaban directamente vinculadas a la estructura del taller, que incluyó demostraciones sobre creación de planes de clase, generación de actividades y ejercicios, creación de textos introductorios, uso de ChatGPT como tutor, generación de retroalimentación personalizada y corrección de evaluaciones. En los materiales del taller, estas posibilidades fueron presentadas como ejemplos de uso pedagógico, siempre acompañadas de la necesidad de revisión, adaptación y supervisión docente. En la Tabla 2 son presentados los datos.

Tabla 2

Prácticas demostradas que llamaron la atención y que se utilizarían en la enseñanza

Práctica demostrada	Llamó la atención n	Llamó la atención %	Lo utilizaría n	Lo utilizaría %
Generación de actividades y ejercicios	74	54.0	118	86.1
ChatGPT como tutor y generación de retroalimentación personalizada	67	48.9	38	27.7
Corrección de evaluaciones	62	45.3	53	38.7
Creación de planes de clase	54	39.4	78	56.9
Creación de textos introductorios y diapositivas	26	19.0	49	35.8

Una diferencia notable aparece entre la proporción de participantes que indicaron que ciertas prácticas les llamaron la atención versus aquellas que declararon usarlas en el futuro. Para la generación de actividades y ejercicios, ambas proporciones fueron similares y altas (54.0% versus 86.1%), mostrando mayor intención de uso futuro que de interés inicial. Para la creación de planes de clase, la intención de uso futuro (56.9%) fue considerablemente mayor que el interés inicial (39.4%), indicando que la práctica fue mejor comprendida o evaluada más positivamente una vez experimentada. En cambio, para prácticas como tutorización con ChatGPT y corrección de evaluaciones, la proporción de uso futuro fue menor que la de interés inicial, mostrando una brecha entre la atención generada por la demostración y la disposición potencial a adoptarla en la práctica futura.

Las prácticas específicamente vinculadas a Educación Matemática, como el uso de ChatGPT con el método de Polya o análisis de demanda cognitiva de tareas, fueron mencionadas por un número menor de participantes, correspondiendo a un recorte específico del taller de Educación Matemática que no fue necesariamente presente de la misma forma en todos los grupos. Esta especificidad disciplinar refleja el diseño del taller, que incluyó demostraciones tanto genéricas como específicas de disciplinas.

Cambios en las percepciones autoinformadas

La comparación entre respuestas pre y postaller fue realizada mediante la prueba de rangos con signo de Wilcoxon para muestras pareadas. Todos los ocho ítems de percepción presentaron aumento en las medias después del taller con diferencias estadísticamente significativas ($p < .001$), presentados en la Tabla 3. El tamaño del efecto fue estimado mediante correlación *rank-biserial* pareada.

Tabla 3

Comparaciones pre y postaller de las percepciones autodeclaradas

Dimensión	Media pre	Media post	Diferencia	p	Tamaño del efecto
Autoeficacia para la elaboración de prompts	3.53	4.07	0.55	< .001	0.710
Conciencia de las limitaciones	4.20	4.54	0.34	< .001	0.564
Conciencia de los sesgos	3.99	4.41	0.42	< .001	0.564
Diferencia entre IA y búsqueda en internet	3.65	4.31	0.66	< .001	0.785
Conocimiento de directrices éticas	3.22	4.05	0.83	< .001	0.814
Impacto epistemológico percibido	3.02	3.75	0.73	< .001	0.660
Utilidad como estudiante universitario	3.88	4.49	0.61	< .001	0.791
Utilidad como futuro docente	3.67	4.16	0.49	< .001	0.619

Los cambios en las medias variaron según el ítem de percepción. La diferencia más elevada se observó en conocimiento de directrices éticas, con cambio de 0.83 puntos, seguido por impacto epistemológico percibido con 0.73 puntos. Diferencia de medias de 0.66 puntos se observó en el ítem de diferencia entre IA y búsqueda en internet. El ítem de utilidad como estudiante universitario presentó diferencia de 0.61 puntos, alcanzando la media postaller más alta (4.49) entre todos los ítems analizados. La autoeficacia para elaboración de prompts mostró cambio de 0.54 puntos. Los ítems sobre sesgos y limitaciones presentaron diferencias de 0.42 y 0.34 puntos respectivamente.

Cabe observar que el ítem de conciencia de limitaciones presentó la media inicial más alta (4.20) entre todos los ítems, reduciendo así el margen posible de mejora, alcanzando 4.54 postaller. En el extremo opuesto, el ítem de impacto epistemológico percibido presentó la media inicial más baja (3.02), permitiendo mayor espacio de ganancia. El ítem de utilidad como futuro docente mostró la diferencia de medias más baja, con 0.49 puntos. A pesar de esta variabilidad en magnitudes de cambio, todos los ocho ítems presentaron mejoras estadísticamente significativas, indicando que la experiencia formativa generó cambios perceptibles en todas las dimensiones medidas.

Cambios diferenciados según perfil de uso previo

Para examinar la variabilidad en los cambios percibidos según el perfil inicial de cada participante, se estratificó el análisis por los seis perfiles identificados, presentados en la Tabla 4. Los participantes del perfil de experiencia baja o nula presentaron ganancia media de 1.15 puntos, equivalente a 39.4% de mejora relativa sobre su puntuación inicial. El perfil de uso inicial mostró ganancia de 0.81 puntos (23.0% de mejora). El perfil con apoyo formativo presentó cambio de 0.44 puntos (11.5% de mejora). El perfil moderado mostró cambio de 0.52 puntos (13.9% de mejora). El perfil con extensiones registró cambio de 0.43 puntos (11.4% de mejora). El perfil diversificado presentó cambio de 0.51 puntos (14.0% de mejora). La ganancia media total en la muestra fue de 0.59 puntos, correspondiente a 16.2% de mejora relativa.

Tabla 4

Cambios en percepciones según perfil de uso previo de IA generativa

Perfil de uso previo	n	Media pre	Media pos	Δ	$\Delta\%$
1. Experiencia baja o nula	9	2.92	4.07	1.15	39.4
2. Uso inicial, centrado en ChatGPT	18	3.52	4.33	0.81	23.0
3. Uso con apoyo formativo	11	3.84	4.28	0.44	11.5
4. Uso moderado (referencia)	51	3.73	4.25	0.52	13.9
5. Uso ampliado con extensiones	16	3.77	4.20	0.43	11.4
6. Uso diversificado de herramientas	32	3.65	4.16	0.51	14.0
Total (M)	137	3.64	4.23	0.59	16.2

Nota. Δ = diferencia media (media post - media pre). $\Delta\%$ = ganancia normalizada ((media post - media pre) / media pre $\times$ 100). Los datos corresponden al score total de percepciones (media de los 8 ítems comparados). Todos los cambios fueron estadísticamente significativos ($p < .001$) cuando se analizaron en conjunto (Tabla 3). La variación en magnitud de cambio según perfil reflejaría efectos de piso (mayores ganancias en puntuaciones basales bajas) y techo (ganancias limitadas en puntuaciones basales altas) comunes en investigaciones de intervención pedagógica.

Una pauta clara emerge del análisis por perfiles: los participantes con experiencia inicial más baja presentaron las ganancias absolutas más pronunciadas, mientras que aquellos con mayor diversificación en herramientas de IA presentaron cambios más moderados en valores absolutos. Los participantes del perfil de experiencia baja o nula ($n=9$) mostraron incremento de 1.15 puntos, mientras que el perfil diversificado ($n=32$) mostró incremento de 0.51 puntos, una diferencia de 0.64 puntos entre los extremos. Sin embargo, todos los grupos presentaron cambios estadísticamente significativos cuando se aplicó la prueba pareada (todos $p < .001$).

El análisis de covarianza (ANCOVA) fue realizado para examinar si las diferencias en cambios percibidos permanecían significativas después de controlar los valores pré-taller. Los resultados mostraron que después de controlar por valores iniciales, no hubo diferencias estadísticamente significativas en los valores posttaller entre los grupos ($F(1,134) = 2.37$, $p = .126$, $\eta^2p = .017$). Sin embargo, los valores pré-taller fueron predictores significativos de los valores posttaller ($F(1,134) = 30.02$, $p < .001$, $\eta^2p = .183$), explicando 18.3% de la varianza en puntuaciones posttaller. Análisis similares en literacia crítica ($F(1,134) = 3.56$, $p = .061$, $\eta^2p = .026$) y actitudes ($F(1,134) = 0.304$, $p = .582$, $\eta^2p = .002$) mostraron patrones equivalentes, donde los valores iniciales fueron predictores significativos pero las diferencias entre grupos posttaller no fueron estadísticamente significativas.

La distribución de cambios según perfil reveló una relación inversa entre el nivel de experiencia inicial y la magnitud del cambio absoluto. Los participantes con baja experiencia inicial operaron con mayor espacio de mejora en la escala, mientras que aquellos ya familiarizados con IA generativa alcanzaban puntuaciones más cercanas al máximo (5.00), limitando así el potencial de ganancia adicional. A pesar de estos cambios más moderados en valores absolutos, las ganancias relativas en porcentaje de mejora continuaron siendo de 11.4% a 14.0% para los grupos establecidos, comparado con 39.4% para el grupo con experiencia baja o nula, indicando que la experiencia

formativa produjo efectos diferenciados según la posición inicial de cada participante en la trayectoria de adopción de herramientas de IA generativa.

DISCUSIÓN

Los resultados de este estudio revelan cambios significativos en las percepciones de futuros docentes de Ciencias y Matemáticas sobre la IA generativa, particularmente en dimensiones éticas, epistemológicas y de utilidad. Estos cambios, aunque vinculados a una experiencia formativa breve, ofrecen evidencia sobre cómo la formación docente inicial puede catalizar una aproximación crítica, ética y pedagógicamente fundamentada a herramientas emergentes como ChatGPT.

Interpretación desde la TDI

La TDI de Rogers (2003) proporciona un marco interpretativo pertinente para comprender estos resultados. Rogers (2003) define la difusión como un proceso social mediante el cual una innovación se comunica a través de canales específicos y a lo largo del tiempo en un sistema social. En el contexto de este estudio, la IA generativa emerge como una innovación percibida por los futuros docentes, cuya adopción depende menos de características objetivas de la herramienta que de cómo los sujetos la interpretan en sus contextos específicos.

Un aspecto central de la TDI es que las percepciones de los individuos sobre los atributos de una innovación (no los atributos clasificados objetivamente por expertos) determinan su tasa de adopción (Rogers, 2003). Los datos de este estudio confirmaron esta proposición. El aumento en la percepción de utilidad como estudiante universitario (diferencia media 0.61, llegando a media 4.49) sugiere que los participantes reconocieron ventaja relativa clara en el contexto de su propio aprendizaje. Simultáneamente, la percepción de utilidad como futuro docente fue menor (diferencia media 0.49, media 4.16), revelando cautela ante la aplicación en contextos pedagógicos. Esta brecha (4.49 vs 4.16) es especialmente relevante porque indica que el reconocimiento de valor no es automático en todos los contextos de actuación. Los futuros docentes parecen distinguir entre "útil para mí como estudiante" y "útil para mí como docente", sugiriendo una conciencia incipiente de que la mediación pedagógica exige consideraciones adicionales a las del aprendizaje personal.

Rogers (2003) describe cinco atributos percibidos que influyen en la adopción: ventaja relativa, compatibilidad, complejidad, experimentabilidad y observabilidad. Los datos de este estudio muestran variabilidad en estos

atributos según el tipo de práctica. La generación de actividades y ejercicios fue indicada como práctica futura por 86.1% de los participantes, probablemente porque representa una práctica de baja complejidad percibida, fácilmente experimentable y con resultados observables. En cambio, el uso de ChatGPT como tutor y la corrección de evaluaciones generaron menor intención de uso futuro (27.7% y 38.7% respectivamente), sugiriendo que los futuros docentes perciben mayor complejidad, mayor necesidad de supervisión humana y mayor responsabilidad pedagógica en estas aplicaciones.

Rogers (2003) también propone que la adopción de una innovación ocurre en etapas: conocimiento, persuasión, decisión, implementación y confirmación. Los resultados de este estudio sitúan a los participantes principalmente entre las etapas de conocimiento y persuasión. Ampliaron su conocimiento sobre las posibilidades de ChatGPT (como muestran los aumentos significativos en todos los ítems), y varios indicaron disposición inicial hacia la adopción. Sin embargo, la presencia de cautela ante usos complejos sugiere que aún no han transitado completamente a la etapa de decisión para todas las aplicaciones. Esta lectura evita una interpretación excesivamente optimista de los datos y reconoce que el cambio de percepción es un primer paso, pero no garantiza adopción pedagógica estable ni efectiva en el aula.

Alfabetización en IA como complemento a la TDI

La alfabetización en IA, según Chiu *et al.* (2024), va más allá de la comprensión técnica e incluye la capacidad de explicar cómo funcionan estas tecnologías, reconocer sus impactos sociales, utilizarlas éticamente y colaborar efectivamente con ellas. Este concepto complementa la TDI al especificar qué dimensiones de conocimiento son necesarias para que la percepción favorable se traduzca en uso crítico y responsable.

Los cambios más pronunciados en este estudio ocurrieron precisamente en dimensiones de alfabetización en IA. El conocimiento de pautas éticas mostró el cambio más alto (diferencia media 0.83), seguido por el impacto epistemológico percibido (0.73) y la comprensión de la diferencia entre IA y búsquedas generales en internet (0.66). Estos cambios reflejan exactamente lo que Chiu *et al.* (2024) definen como competencia en IA: comprender críticamente cómo funciona la tecnología, reconocer sus límites y actuar éticamente.

El aumento en la conciencia de limitaciones (diferencia media 0.34) y sesgos (0.42) es particularmente relevante, aunque las magnitudes fueron

menores. Esto se debe a que los participantes ya presentaban conciencia inicial moderada-alta sobre limitaciones (media inicial 4.20). A pesar de este efecto "techo", los cambios continuaron siendo estadísticamente significativos, indicando que incluso para participantes con conciencia previa, la experiencia formativa profundizó esa comprensión.

Este punto es central para la formación de profesores de Ciencias y Matemáticas. Las respuestas generadas por IA pueden parecer coherentes y convincentes, pero contener errores conceptuales, simplificaciones pedagógicamente inadecuadas o sesgos culturales (Dilling & Herrmann, 2024). La capacidad de evaluar críticamente estas respuestas constituye una condición para su uso pedagógico responsable. En esta dirección, la IA generativa debe ser incorporada a la formación docente como objeto de análisis crítico y como recurso mediado por el conocimiento profesional del profesor.

Dimensiones de cambio y sus implicaciones pedagógicas

El aumento en el impacto epistemológico percibido (media inicial 3.02, postaller 3.75) sugiere que la experiencia formativa logró ir más allá de la capacitación instrumental. Los participantes indicaron cambios en cómo conciben la relación entre IA, conocimiento y aprendizaje. Esto dialoga directamente con perspectivas que enfatizan que las tecnologías generativas no son simplemente herramientas neutrales, sino que transforman formas de pensar, aprender y producir conocimiento (Blonder *et al.*, 2025). La formación docente debe preparar profesores para que reconozcan y naveguen activamente estas transformaciones epistemológicas.

La autoeficacia para formular *prompts* aumentó de 3.53 a 4.07 (diferencia 0.54), indicando que los participantes ganaron confianza en interactuar intencionalmente con ChatGPT. Este resultado es importante porque la formulación de preguntas específicas es una práctica clave para obtener respuestas de calidad pedagógica (Dilling & Herrmann, 2024). Sin embargo, los mismos autores también advierten que muchos futuros profesores tienden a usar modelos de lenguaje de forma poco específica, sin técnicas deliberadas de *prompting*. Los resultados de este estudio sugieren que una experiencia formativa breve puede mejorar esta competencia, aunque probablemente requiera práctica y retroalimentación sostenida en contextos de enseñanza real.

Variabilidad según perfil de uso previo

Un hallazgo notable es cómo los cambios variaron según la experiencia inicial con IA generativa. Los participantes con baja experiencia o nula

presentaron ganancias absolutas más pronunciadas (1.15 puntos, 39.4% de mejora), mientras que usuarios diversificados mostraron cambios más moderados (0.51 puntos, 14.0% de mejora). La TDI ayuda a interpretar este patrón: individuos con puntuaciones iniciales más bajas operan con mayor "espacio de mejora" en la escala, mientras que aquellos ya familiarizados con la temática alcanzan puntuaciones cercanas al máximo, limitando el potencial de ganancia adicional. Este efecto es bien documentado en investigaciones de intervención.

Sin embargo, lo importante es que todos los perfiles, incluyendo usuarios establecidos, mostraron cambios estadísticamente significativos. Esto sugiere que la experiencia formativa fue particularmente efectiva para introducir la IA generativa como temática de reflexión crítica, pedagógica y ética a futuros docentes con poca experiencia previa, mientras que simultáneamente contribuyó a profundizar o matizar percepciones en usuarios ya familiarizados. Esto tiene implicaciones para el diseño de programas formativos: deben contemplar heterogeneidad en familiaridad sin sacrificar rigor crítico para ningún grupo.

Implicaciones disciplinares: Educación Matemática y Ciencias

En Educación Matemática, los resultados indican la necesidad de usos más situados y específicamente diseñados. Dilling & Herrmann (2024) documentan que futuros profesores tienden a utilizar ChatGPT de forma genérica para tareas Matemáticas, sin explorar técnicas específicas de *prompting* o sin verificar la validez matemática de las respuestas. La incorporación en este taller del método de Polya y de la discusión sobre demanda cognitiva de tareas Matemáticas señala una posibilidad pedagógica relevante: usar ChatGPT no como proveedor automático de ejercicios, sino como herramienta para generar ejercicios que luego son analizados críticamente según su calidad didáctica, nivel de demanda cognitiva, y validez de los procesos de resolución. Este uso exige que el profesor desarrolle criterios específicos sobre qué hace "buena" una tarea Matemática, independientemente de quién la generó.

En Educación Científica, la contribución se relaciona con dimensiones de planificación, explicación conceptual, adaptación de actividades y producción de retroalimentación. Blonder *et al.* (2025) argumentan que la IA generativa puede favorecer la reflexión sobre conocimientos pedagógicos, tecnológicos y disciplinares en la formación de profesores de Ciencias. Los aumentos observados en autoeficacia para formular *prompts* y en percepción de utilidad docente señalan una apertura para repensar cómo los profesores

pueden mediar entre el conocimiento científico, el pedagógico y las potencialidades de herramientas generativas. Simultáneamente, la cautela ante la tutorización automatizada sugiere conciencia de que la mediación pedagógica en Ciencias exige relación directa entre profesor y estudiante, no puede ser completamente delegada a IA.

Tensiones importantes: interés pedagógico versus cautela responsable

Un aspecto fundamental de los resultados es la tensión entre disposición favorable y cautela pedagógica. La generación de actividades y ejercicios fue indicada por 54.0% como práctica que llamó la atención, pero por 86.1% como posible uso futuro, mostrando alineación entre interés y adopción potencial. Sin embargo, para prácticas como tutorización (67 participantes, 48.9% interesados, pero solo 38, 27.7% como uso futuro) y corrección de evaluaciones (62 interesados, 45.3%, pero 53 como uso futuro, 38.7%), la brecha es considerable. Esta divergencia no es negativa; es evidencia de pensamiento crítico. Los futuros docentes reconocen el potencial de la IA generativa, pero simultáneamente perciben que ciertos usos involucran responsabilidades pedagógicas específicas (evaluación del aprendizaje, acompañamiento personalizado de estudiantes) que no pueden ser completamente delegadas a sistemas automáticos.

Esta cautela se alinea con Bae *et al.* (2024), quienes observaron que experiencias con ChatGPT amplían la conciencia y reducen la ansiedad de futuros profesores, pero la intención de adopción pedagógica permanece limitada. Los autores interpretan esto no como fracaso formativo, sino como evidencia de que los docentes están considerando seriamente las implicaciones pedagógicas de cada uso. De manera similar, Panday-Shukla (2025) encontró que la adopción de IA generativa está mediada por percepciones de compatibilidad con valores pedagógicos, preocupaciones sobre integridad académica, y necesidad de orientaciones institucionales claras.

Un aspecto crítico aquí es la cuestión de la sustitución del profesor. Los datos sugieren que los futuros docentes no interpretan la IA generativa como sustituta, sino como herramienta que requiere mediación docente. La percepción más baja de utilidad como futuro docente en comparación con estudiante (4.16 vs 4.49) puede reflejar esta comprensión: el profesor no puede simplemente usar ChatGPT "como es", sino que debe traducir, adaptar, validar y contextualizar sus respuestas. Este es precisamente el rol insustituible del docente que requiere formación específica.

Limitaciones y consideraciones metodológicas

Este estudio presenta limitaciones que deben ser consideradas. El diseño pre-post sin grupo de control no permite atribuir causalidad a la experiencia formativa de forma definitiva; es posible que otros factores (conversaciones con pares, exposición mediática a IA, o simplemente el paso del tiempo) hayan contribuido a los cambios observados. Los datos fueron autoinformados, basados en percepciones declaradas de los participantes, no en observación directa de prácticas docentes reales. La experiencia formativa fue de corta duración (dos horas), en un contexto institucional específico, con una muestra de 137 participantes de educación primaria. Por estas razones, los resultados deben ser interpretados con cautela respecto a procesos de adopción estable y a implementación efectiva en el aula.

Adicionalmente, el seguimiento fue inmediato postaller. No se conoce si los cambios en percepciones se mantienen en el tiempo, ni se observó si la disposición favorable a adoptar ciertas prácticas se tradujo efectivamente en su uso en contextos de enseñanza real. Este es un punto especialmente importante dado que la literatura sobre difusión de innovaciones muestra que la brecha entre intención de uso y adopción real puede ser sustancial.

Implicaciones para la formación docente inicial

Los hallazgos de este estudio tienen implicaciones claras para diseñar programas de formación docente inicial que integren IA generativa. En primer lugar, la formación debe ser explícitamente crítica y ética, no meramente instrumental. El hecho de que el mayor cambio ocurriera en conocimiento de pautas éticas (0.83) indica que los docentes en formación valoran y responden a orientaciones éticas, cuando estas son ofrecidas de forma clara y integrada.

En segundo lugar, la formación debe promover experimentación situada y reflexión pedagógica. Los aumentos en autoeficacia para formular *prompts* y en impacto epistemológico percibido sugieren que la práctica deliberada con la herramienta, combinada con discusión sobre sus implicaciones, cataliza aprendizaje más profundo que la exposición meramente informativa.

En tercer lugar, la heterogeneidad de experiencias previas con IA requiere que los programas formativos diferencien, sin sacrificar rigor. Algunos participantes llegaban sin experiencia; otros con familiaridad considerable. Un programa efectivo debe servir ambos grupos, reconociendo que los principios éticos, críticos y pedagógicos trascienden el nivel técnico de familiaridad.

En cuarto lugar, la formación debe reconocer las diferencias disciplinares. Matemática y Ciencias tienen sus propias preocupaciones sobre calidad, validez del conocimiento y mediación pedagógica. Formaciones genéricas sobre "IA en educación" pierden esta especificidad. Los resultados de este estudio, aunque derivados de una experiencia formativa general, subrayan la importancia de traducir los principios a contextos disciplinares específicos.

Contribución a la convergencia teórica

Este estudio contribuye a la convergencia entre tres campos importantes: formación inicial docente, alfabetización en IA, y TDI. Muestra que una experiencia formativa breve puede catalizar cambios significativos en percepciones relevantes para esta convergencia. Específicamente, evidencia que cambios en comprensión crítica (ética, epistemología, evaluación de calidad) ocurren simultáneamente a cambios en disposición hacia la adopción pedagógica, aunque esta disposición sea diferenciada según el tipo de aplicación.

Los resultados apoyan la integración de la IA generativa en la formación inicial de docentes de Ciencias y Matemáticas como objeto de experimentación, análisis crítico y decisión pedagógica fundamentada. Esto no significa aceptación acrítica de la tecnología, sino desarrollo de capacidad de los futuros profesores para evaluar cuándo, cómo y por qué utilizar ChatGPT en favor del aprendizaje, siempre considerando la calidad didáctica, los límites de la herramienta, la evaluación responsable del aprendizaje y el rol insustituible de la mediación docente.

CONCLUSIONES

Este estudio partió de una pregunta fundamental: cómo cambian las percepciones de futuros docentes sobre IA generativa después de una experiencia formativa breve. La respuesta que emergen de los datos es clara. Los futuros profesores que vivencian ChatGPT a través de una formación estructurada desarrollan una relación con la tecnología que es radicalmente diferente de la que tendrían por su cuenta. No es simplemente que aprendan a usarla. Es que empiezan a cuestionarla.

Los cambios observados en conocimiento ético, reflexión epistemológica y comprensión de límites sugieren que la formación docente inicial es un momento donde se decide algo fundamental: si los profesores que emergerán de nuestros programas verán IA como herramienta neutra a ser consumida o como artefacto cultural que modela la forma en que pensamos,

aprendemos y enseñamos. La magnitud de cambio en ética (0.83 puntos) respecto a otras dimensiones señala que cuando se les ofrece orientación clara sobre responsabilidad, los futuros docentes responden. Esto importa porque sugiere que hay agencia en el diseño curricular para cultivar pensamiento crítico sobre tecnología desde el inicio de la formación profesional.

La brecha entre utilidad percibida como estudiante y como docente (4.49 versus 4.16) reveló algo que la literatura había insinuado pero que este estudio documenta con claridad. Los futuros profesores no rechazan la IA. Pero reconocen que su aplicación en aula no es asunto técnico. Implica decisiones sobre evaluación, sobre mediación pedagógica, sobre el tipo de relación que quieren con sus estudiantes. Esta cautela no es ignorancia. Es pensamiento. Es una forma de resistencia consciente contra la tentación de dejar que la eficiencia tecnológica reemplace la complejidad relacional de la enseñanza. Programas formativos que cultiven esta capacidad de discernimiento están cumpliendo un propósito educativo que trasciende competencias digitales.

El estudio documenta también que estos cambios ocurrieron de forma diferenciada según experiencia previa con IA. Participantes sin familiaridad previa experimentaron ganancias más pronunciadas, mientras usuarios establecidos mostraron cambios más moderados. Este patrón es esperado y revelador. Indica que la formación efectivamente genera cambios en percepción, sin importar el punto de partida. Sugiere además que las instituciones pueden diseñar experiencias que sirvan a públicos heterogéneos sin sacrificar profundidad crítica para quienes ya tienen familiaridad con tecnología.

La contribución central de esta investigación reside en documentar que estudiantes de formación inicial en educación pueden, en breves espacios de tiempo, desarrollar una relación sofisticada con tecnologías emergentes cuando esas experiencias integran explícitamente dimensiones éticas y reflexivas. Otros estudios han mostrado que docentes pueden aprender a usar IA. Este estudio avanza mostrando que pueden aprender a pensarla críticamente. Es una diferencia que importa porque los docentes que piensan críticamente sobre tecnología transmiten esa capacidad a sus estudiantes. Los que la adoptan sin reflexión transmiten consumo acrítico.

Quedan preguntas sin responder que orientarán investigaciones futuras. La más urgente es si estos cambios en percepción se traducen en cambios reales en cómo los futuros profesores efectivamente usan IA una vez en contextos de enseñanza. También resulta crítico comprender si estas percepciones se mantienen o se erosionan con el tiempo. Es posible que futuros docentes, enfrentados a presiones reales de cobertura curricular, falta de apoyo

institucional y carencia de orientaciones claras en sus escuelas, gradualmente abandonen el pensamiento crítico que desarrollaron en su formación inicial y reviertan a usos más acríticos de la tecnología.

Para que las percepciones que documentamos aquí se traduzcan en práctica sostenida, la formación inicial no puede ser acto aislado. Requiere que las instituciones educativas donde esos futuros docentes trabajarán creen espacios de reflexión continuada, que ofrezcan recursos pedagógicos sobre IA crítica, y que las políticas educativas proporcionen orientaciones claras sin prescindir del juicio profesional docente. De lo contrario, corremos el riesgo de preparar profesores reflexivos para entornos que no valorarán esa reflexión.

Este estudio contribuye a una conversación más amplia sobre cómo la formación docente responde a innovaciones tecnológicas emergentes. Sugiere que es posible cultivar una generación de educadores que no ignoren la IA, que no la adopten ciegamente, sino que la incorporen como parte de una práctica profesional informada, ética y pedagógicamente fundamentada. Es responsabilidad de las instituciones de formación docente perseguir esa posibilidad.

DECLARACIONES SOBRE LAS CONTRIBUCIONES DE LOS AUTORES

Helcio Soares Padilha y Agostinho Serrano concibieron la idea e indicaron el marco teórico. Helcio Soares Padilha desarrolló la investigación y recopiló los datos. Víctor López-Simó indicó la metodología. Helcio Soares Padilha y Víctor López-Simó analizaron los datos. Fabiane Fischer Figueiredo desarrolló el estado del arte y revisó la metodología y la discusión. Todos los autores participaron activamente en la discusión de los resultados, revisaron y aprobaron la versión final del trabajo.

DECLARACIÓN DE DISPONIBILIDAD DE DATOS

Los datos que respaldan los resultados de este estudio estarán disponibles a solicitud razonable del autor correspondiente, Helcio Soares Padilha.

AGRADECIMIENTOS Y FINANCIACIÓN

El presente trabajo fue realizado con el apoyo de la Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiación 001.

REFERENCIAS

- Bae, H., Hur, J., Park, J., Choi, G. W., & Moon, J. (2024). Pre-service teachers' dual perspectives on generative AI: Benefits, challenges, and integration into their teaching and learning. *Online Learning*, 28(3), 131–156. <https://doi.org/10.24059/olj.v28i3.4543>
- Blonder, R., Feldman-Maggor, Y., & Rap, S. (2025). Are they ready to teach? Generative AI as a means to uncover pre-service science teachers' PCK and enhance their preparation program. *Journal of Science Education and Technology*, 34(6), 1301–1310. <https://doi.org/10.1007/s10956-024-10180-2>
- Chiu, T. K. F., Ahmad, Z., Ismailov, M., & Sanusi, I. T. (2024). What are artificial intelligence literacy and competency? A comprehensive framework to support them. *Computers and Education Open*, 6, Article 100171. <https://doi.org/10.1016/j.caeo.2024.100171>
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE.
- Dilling, F., & Herrmann, M. (2024). Using large language models to support pre-service teachers' mathematical reasoning: An exploratory study on ChatGPT as an instrument for creating mathematical proofs in geometry. *Frontiers in Artificial Intelligence*, 7, Article 1460337. <https://doi.org/10.3389/frai.2024.1460337>
- Dos Santos Dos Santos, J. M., Silveira, A. P. R., Breda, A. M. R. d., & Lavicza, Z. (2025). Challenges in science, technology, engineering, arts, and mathematics education in Cape Verde: A study of a mathematics teacher training project. *Education Sciences*, 15(1), Article 81. <https://doi.org/10.3390/educsci15010081>
- Flick, U. (2013). *Introdução à metodologia de pesquisa: Um guia para iniciantes*. Penso.
- Fowler, F. J., Jr. (2014). *Survey research methods* (5th ed.). SAGE Publications.
- Gil, A. C. (2002). *Como elaborar projetos de pesquisa* (4th ed.). Atlas.
- Kerby, D. S. (2014). The simple difference formula: An approach to teaching nonparametric correlation. *Comprehensive Psychology*, 3, 2165–2228.
- Koehler, M. J., & Mishra, P. (2008). Introducing technological pedagogical content knowledge. In AACTE Committee on Innovation and Technology (Eds.), *Handbook of technological pedagogical content knowledge (TPCK) for educators* (pp. 3–29). Routledge.
- Meegan, J., & Young, K. (2025). It's a tool not a crutch: A pilot generative AI intervention to enhance pre-service teachers' self-efficacy and AI

- literacy. *Technology, Knowledge and Learning*, 30, 1895–1916.
<https://doi.org/10.1007/s10758-025-09875-1>
- Panday-Shukla, P. (2025). Exploring generative artificial intelligence in teacher education. *Teaching and Teacher Education*, 165, Article 105088. <https://doi.org/10.1016/j.tate.2025.105088>
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- The jamovi project (2024). *jamovi*. (Version 2.5.4) [Computer software]. Retrieved from <https://www.jamovi.org>.
- Vaillant, D., Rodríguez Zidán, E., & Bentancor Biagas, G. (2020). Uso de plataformas e ferramentas digitais para o ensino da Matemática. *Ensaio: Avaliação e Políticas Públicas em Educação*, 28(108), 718–740. <https://doi.org/10.1590/S0104-40362020002802241>
- Zhang, Y., & Zhu, J. (2023). STEM pre-service teacher education: A review of research trends in the past ten years. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), Article em2292. <https://doi.org/10.29333/ejmste/13300>