

Más allá de la respuesta correcta: el impacto del uso de chatbots de Inteligencia Artificial en habilidades del pensamiento crítico en el área de matemáticas con estudiantes de nivel secundaria

Beyond the correct answer: the impact of using Artificial Intelligence chatbots on critical thinking skills in mathematics with secondary school students

Liliana Karina Mora Quintero¹ y Edgar Escobar Gutiérrez²

¹Universidad Contemporánea de las Américas, lilipzy@hotmail.com, <https://orcid.org/0009-0005-9919-6377>, México

²Universidad Contemporánea de las Américas, edgar.e.g@unicla.edu.mx, <https://orcid.org/0009-0002-4237-6092>, México

Información del Artículo

Trazabilidad:

Recibido 25-04-2026

Revisado 26-04-2026

Aceptado 30-06-2026

Palabras Clave:

Pensamiento crítico
Inteligencia artificial
Chatbots

Keywords:

Critical thinking
Artificial intelligence
Chatbots

RESUMEN

El desarrollo del pensamiento crítico constituye una prioridad educativa ante la creciente incorporación de tecnologías digitales en el aula. En este contexto, el uso de chatbots basados en inteligencia artificial plantea oportunidades y desafíos para los procesos cognitivos relacionados con la resolución de problemas matemáticos. El presente estudio tuvo como objetivo analizar el efecto del uso guiado de un chatbot en el desarrollo de habilidades de pensamiento crítico matemático en estudiantes de tercer grado de secundaria. Se empleó un enfoque mixto con diseño cuasi experimental, integrado por un grupo control y un grupo experimental. Para la recolección de datos se utilizaron una prueba de desempeño, una escala Likert, una rúbrica analítica y entrevistas semiestructuradas. La intervención se desarrolló durante seis sesiones bajo un enfoque de mediación pedagógica. Los resultados mostraron que el grupo experimental presentó mayores avances en las habilidades de análisis e inferencia matemática, alcanzando un 62% de desempeño en niveles avanzados, frente al 29% observado en el grupo control. Asimismo, se identificó una evolución en el uso del chatbot, pasando de estrategias reproductivas a prácticas más reflexivas y orientadas a la verificación de respuestas, alcanzando un 80% en la categoría "siempre". Se concluye que el uso guiado de chatbots contribuye al fortalecimiento del pensamiento crítico matemático cuando se integra en estrategias didácticas orientadas a la reflexión, la evaluación y la argumentación.

ABSTRACT

The development of critical thinking is an educational priority given the increasing integration of digital technologies in the classroom. In this context, the use of AI-based chatbots presents both opportunities and challenges for the cognitive processes related to mathematical problem-solving. This study aimed to analyze the effect of guided chatbot use on the development of mathematical critical thinking skills in third-year secondary school students. A mixed-methods approach with a quasi-experimental design was employed, consisting of a control group and an experimental group. Data collection methods included a performance test, a Likert scale, an analytical rubric, and semi-structured interviews. The intervention was conducted over six sessions using a pedagogical mediation approach. The results showed that the experimental group demonstrated greater progress in mathematical analysis and inference skills, achieving 62% performance at advanced levels, compared to 29% in the control group. Furthermore, an evolution in chatbot use was identified, shifting from reproductive strategies to more reflective practices focused on verifying responses, reaching 80% in the "always" category. It is concluded that the guided use of chatbots contributes to strengthening mathematical critical thinking when integrated into teaching strategies focused on reflection, evaluation, and argumentation.

INTRODUCCIÓN

La incorporación de tecnologías basadas en inteligencia artificial (IA), particularmente los chatbots educativos, ha transformado significativamente los entornos de enseñanza y aprendizaje. Estas herramientas ofrecen respuestas inmediatas, acompañamiento personalizado y apoyo en la resolución de tareas académicas, lo que ha favorecido su creciente adopción en diversos niveles educativos, incluida la educación secundaria. No obstante, su integración plantea un desafío relevante para la investigación educativa: determinar si su uso contribuye al fortalecimiento de habilidades cognitivas de orden superior o si, por el contrario, fomenta formas de dependencia que limitan el desarrollo del pensamiento crítico. En este contexto, diversos estudios han señalado que la IA puede constituirse como una herramienta pedagógica valiosa cuando su implementación responde a objetivos didácticos claramente definidos y se encuentra mediada por procesos de reflexión y acompañamiento docente.

Aunque la literatura reconoce beneficios y riesgos asociados al uso de la inteligencia artificial en la educación, existe consenso en que los efectos de estas tecnologías dependen en gran medida de las condiciones pedagógicas bajo las cuales son utilizadas. Investigaciones recientes destacan que la retroalimentación inmediata, la personalización del aprendizaje y el apoyo durante la resolución de problemas pueden favorecer procesos de análisis, reflexión y metacognición (Erazo Molina et al., 2023; Mosqueda Chávez, 2024). Sin embargo, también se ha advertido que una utilización carente de orientación didáctica puede generar dependencia tecnológica y reducir la autonomía cognitiva del estudiantado (Correal Romero, 2025). Desde esta perspectiva, el presente estudio parte de la premisa de que los efectos observados no son inherentes a la tecnología, sino a las estrategias pedagógicas que acompañan su implementación.

A pesar del crecimiento de la investigación sobre inteligencia artificial aplicada a la educación, la mayor parte de los estudios se ha desarrollado en contextos universitarios o en escenarios generales de aprendizaje digital. En contraste, continúa siendo limitada la evidencia empírica acerca del impacto de los chatbots basados en inteligencia artificial generativa sobre el desarrollo del pensamiento crítico matemático en estudiantes de educación secundaria, particularmente en contextos latinoamericanos. Esta situación evidencia una brecha de conocimiento que dificulta comprender de qué manera estas herramientas influyen en procesos cognitivos complejos dentro de escenarios educativos reales.

En atención a esta problemática, la presente investigación tiene como objetivo analizar el efecto del uso de chatbots de inteligencia artificial en el desarrollo de habilidades de pensamiento crítico en estudiantes de tercero de secundaria durante la resolución de problemas matemáticos. De manera específica, se examinan dimensiones como la interpretación, el análisis, la evaluación, la argumentación y la autorregulación, con el propósito de generar evidencia que contribuya a orientar una integración pedagógica crítica, reflexiva y formativa de estas tecnologías en el aula.

MATERIALES Y MÉTODOS

La presente investigación se desarrolló bajo un enfoque mixto con diseño explicativo secuencial (QUAN→QUAL), caracterizado por el predominio de la fase cuantitativa y complementado con evidencia cualitativa orientada a profundizar en la comprensión de los procesos cognitivos asociados al uso de chatbots basados en inteligencia artificial (IA) durante la resolución de problemas matemáticos. Este enfoque permite integrar datos cuantitativos y cualitativos en un mismo estudio para obtener una comprensión más amplia y profunda del fenómeno investigado (Hernández-Sampieri & Mendoza, 2018).

La fase cuantitativa tuvo como propósito identificar el efecto del uso guiado de chatbots de IA sobre el desarrollo del pensamiento crítico matemático mediante la comparación de resultados obtenidos antes y después de una intervención educativa. Complementariamente, la fase cualitativa permitió explorar las experiencias de los estudiantes, las estrategias de razonamiento empleadas y los procesos de autorregulación y metacognición desarrollados durante la interacción con estas herramientas tecnológicas. Se utilizó un diseño cuasi experimental con grupo control y grupo experimental, estructurado mediante un esquema pretest–intervención–postest. Este tipo de diseño resulta apropiado en contextos educativos donde no es posible realizar una asignación completamente aleatoria de los participantes, debido a que los grupos suelen encontrarse previamente conformados (Hernández-Sampieri et al., 2014).

La variable independiente correspondió al uso guiado de chatbots de inteligencia artificial como herramienta de mediación cognitiva durante la resolución de problemas matemáticos. La variable dependiente fue el desarrollo del pensamiento crítico matemático, operacionalizado a partir de las

dimensiones de interpretación, análisis, evaluación, inferencia, explicación, autorregulación y metacognición propuestas por Facione (2007).

La hipótesis de investigación planteó que los estudiantes que participan en una estrategia de aprendizaje mediada por chatbots de inteligencia artificial presentan niveles superiores de pensamiento crítico matemático en comparación con aquellos que trabajan mediante estrategias tradicionales sin el uso de estas herramientas. La muestra estuvo conformada por 15 estudiantes de tercer grado de secundaria pertenecientes a la Escuela Valle de Bravo, distribuidos en un grupo experimental ($n = 8$) y un grupo control ($n = 7$). La selección de los participantes se realizó mediante un muestreo no probabilístico por conveniencia, debido a las condiciones de acceso institucional y disponibilidad de los grupos participantes. Este tipo de muestreo es frecuente en investigaciones educativas aplicadas en contextos reales de aula, donde la población se encuentra delimitada por condiciones institucionales específicas (Otzen & Manterola, 2017).

La recolección de datos se realizó mediante un sistema multimétodo integrado por cinco instrumentos diseñados a partir de una matriz de operacionalización alineada con las dimensiones del pensamiento crítico y los objetivos de investigación. El primer instrumento fue la Prueba de Desempeño en Pensamiento Crítico Matemático (PDPCM-IA), diseñada para evaluar habilidades de interpretación, análisis, evaluación, inferencia, explicación, autorregulación y metacognición a través de problemas abiertos contextualizados que requerían justificar procedimientos y argumentar decisiones durante la resolución de situaciones matemáticas.

El segundo instrumento consistió en una Rúbrica Analítica de Pensamiento Crítico, estructurada en cuatro niveles de desempeño y utilizada para valorar sistemáticamente las respuestas obtenidas en la PDPCM-IA. Los descriptores de la rúbrica fueron elaborados considerando los componentes centrales del pensamiento crítico propuestos por Facione (2007).

Como tercer instrumento se utilizó la Escala Comparativa de Pensamiento Crítico (ECPC-M), integrada por reactivos tipo Likert de cinco puntos orientados a medir la frecuencia con la que los estudiantes manifiestan comportamientos asociados al pensamiento crítico durante la resolución de problemas matemáticos. Esta escala fue aplicada en modalidad pretest y postest.

El cuarto instrumento correspondió a una Escala de Percepción sobre Uso de Chatbots, dirigida exclusivamente al grupo experimental y orientada a explorar dimensiones relacionadas con utilidad cognitiva, autonomía, pensamiento crítico, reflexión y riesgos percibidos durante la interacción con herramientas de IA.

Finalmente, se diseñó una Entrevista Semiestructurada compuesta por preguntas abiertas derivadas de los objetivos de investigación, con la finalidad de profundizar en las experiencias de aprendizaje, estrategias cognitivas y percepciones de los estudiantes respecto al uso de chatbots en contextos matemáticos.

Todos los instrumentos fueron sometidos a un proceso de validación por juicio de expertos conformado por especialistas en metodología de la investigación, tecnología educativa y enseñanza de las matemáticas. Posteriormente se realizó una prueba piloto con estudiantes de características similares a las de la muestra definitiva para verificar la claridad, pertinencia y coherencia de los reactivos. La confiabilidad de las escalas fue estimada mediante el coeficiente Alfa de Cronbach, obteniéndose valores superiores a .80, considerados adecuados para investigaciones educativas (George & Mallery, 2019).

La investigación se desarrolló en tres fases.

1. En la primera fase se aplicaron la PDPCM-IA y la ECPC-M a ambos grupos con el propósito de establecer una línea base de comparación respecto a las habilidades de pensamiento crítico matemático.
2. En la segunda fase se implementó una intervención didáctica dirigida exclusivamente al grupo experimental. La intervención se desarrolló durante seis sesiones consecutivas de aproximadamente 50 minutos cada una dentro del horario regular de clase. Las actividades fueron diseñadas bajo un enfoque de mediación pedagógica orientado al fortalecimiento del pensamiento crítico mediante el uso guiado de un chatbot de inteligencia artificial. La secuencia didáctica promovió progresivamente habilidades de interpretación, análisis, evaluación, inferencia y metacognición a través de actividades centradas en la identificación de errores, comparación de estrategias, verificación de resultados, resolución de problemas complejos y reflexión sobre los procesos de aprendizaje. Durante toda la intervención se enfatizó que el chatbot debía utilizarse como herramienta para contrastar procedimientos, justificar decisiones y evaluar respuestas, evitando su empleo como sustituto del razonamiento matemático. El docente asumió un papel de mediador, formulando preguntas orientadoras y promoviendo la argumentación crítica sin proporcionar respuestas directas. Las producciones escritas generadas por los estudiantes fueron recopiladas sistemáticamente para su posterior análisis. De manera paralela, el grupo control desarrolló los mismos contenidos curriculares y resolvió problemas matemáticos equivalentes

utilizando estrategias tradicionales de enseñanza y aprendizaje, sin acceso a herramientas de inteligencia artificial durante el periodo de intervención.

3. En la tercera fase se aplicaron nuevamente la PDPCM-IA y la ECPC-M a ambos grupos. Adicionalmente, el grupo experimental respondió la Escala de Percepción sobre Uso de Chatbots y participó en entrevistas semiestructuradas orientadas a profundizar en las experiencias y procesos cognitivos desarrollados durante la intervención.

El análisis de la información se desarrolló de acuerdo con la lógica explicativa secuencial propia del diseño mixto. En la fase cuantitativa se realizaron análisis descriptivos mediante frecuencias, porcentajes, medias y desviaciones estándar para caracterizar el comportamiento de las variables estudiadas. Posteriormente, se efectuaron análisis inferenciales con el propósito de comparar los resultados obtenidos entre el pretest y el posttest, así como entre el grupo experimental y el grupo control. Considerando el tamaño reducido de la muestra, se emplearon pruebas paramétricas o no paramétricas según los supuestos de normalidad observados en los datos. Asimismo, se calculó la consistencia interna de las escalas mediante el coeficiente Alfa de Cronbach.

La información cualitativa obtenida a través de las entrevistas y de las evidencias producidas durante la intervención fue analizada mediante análisis temático siguiendo las fases propuestas por Braun y Clarke (2006): familiarización con los datos, generación de códigos iniciales, búsqueda de temas, revisión de temas, definición y denominación de categorías y elaboración del informe interpretativo.

Finalmente, los resultados cuantitativos y cualitativos fueron integrados mediante un proceso de triangulación metodológica con el propósito de construir una interpretación comprehensiva sobre el impacto de los chatbots de inteligencia artificial en el desarrollo del pensamiento crítico matemático.

La investigación se desarrolló respetando los principios éticos establecidos para estudios educativos con participación de menores de edad. Se contó con la autorización formal de la institución educativa, así como con el consentimiento informado de los padres o tutores y el asentimiento voluntario de los estudiantes participantes. Asimismo, se garantizó la confidencialidad, anonimización y protección de la información recopilada, asegurando que los datos fueran utilizados exclusivamente con fines académicos y científicos.

El estudio se condujo conforme a los principios éticos establecidos en la Declaración de Helsinki para investigaciones con seres humanos (Asociación Médica Mundial, 2013) y siguiendo los criterios de respeto, beneficencia y justicia recomendados para la investigación educativa contemporánea.

RESULTADOS Y DISCUSIÓN

La presente sección presenta los hallazgos obtenidos mediante los instrumentos de recolección de datos aplicados durante la investigación. Los resultados se organizan de acuerdo con las dimensiones del pensamiento crítico evaluadas y con la naturaleza de la evidencia recopilada. En concordancia con el diseño mixto explicativo secuencial del estudio, se integran resultados cuantitativos y cualitativos con el propósito de ofrecer una comprensión más amplia del efecto del uso de chatbots basados en inteligencia artificial sobre el desarrollo del pensamiento crítico matemático. La exposición se estructura en cuatro apartados: desempeño en pensamiento crítico, evolución durante la intervención, percepción estudiantil y hallazgos cualitativos derivados de las entrevistas semiestructuradas.

Los resultados de la Prueba de Desempeño en Pensamiento Crítico Matemático (PDPCM-IA) permitieron comparar los niveles de desempeño alcanzados por el grupo experimental ($n = 8$) y el grupo control ($n = 7$) antes y después de la intervención educativa.

En la evaluación inicial (pretest), ambos grupos presentaron patrones de desempeño similares. El 75% de los estudiantes del grupo experimental y el 71% del grupo control se ubicaron en los niveles básico e intermedio de las habilidades evaluadas. Asimismo, la dimensión de evaluación registró los niveles más bajos de desempeño, observándose que el 62% del grupo experimental y el 57% del grupo control se concentraban en el nivel inicial.

Tras la implementación de la intervención, los resultados del posttest evidenciaron diferencias entre ambos grupos. El grupo experimental mostró una mejora en las dimensiones de análisis e inferencia, donde el 62% de los participantes alcanzó niveles avanzados de desempeño. En contraste, únicamente el 29% de los estudiantes del grupo control logró ubicarse en este nivel. De igual manera, mientras el grupo experimental presentó una transición hacia categorías de desempeño superiores, el grupo control permaneció predominantemente en el nivel intermedio (57%).

Estos resultados sugieren una evolución favorable en las habilidades de pensamiento crítico matemático del grupo experimental, particularmente en aquellas relacionadas con el análisis de información, la inferencia de soluciones y la evaluación de procedimientos durante la resolución de problemas matemáticos.

Tabla 1: Comparación de niveles de desempeño por habilidad (%)

Habilidad	Experimental Pre	Experimental Post	Control Pre	Control Post
Interpretación	50% básico	63% intermedio	57% básico	43% básico / 43% intermedio
Análisis	62% intermedio	62% avanzado	57% intermedio	57% intermedio
Evaluación	62% inicial	58% intermedio	57% inicial	57% básico
Inferencia	62% intermedio	62% avanzado	57% intermedio	57% intermedio
Explicación	50% básico	63% intermedio	57% básico	43% básico / 43% intermedio

Los resultados obtenidos evidencian un mayor avance en el grupo experimental, particularmente en las habilidades asociadas al análisis, la inferencia y la evaluación de procedimientos matemáticos. En comparación con el grupo control, los estudiantes que participaron en la intervención mediada por chatbots mostraron una mayor proporción de desempeños ubicados en niveles avanzados, lo que sugiere un fortalecimiento de las habilidades de pensamiento crítico vinculadas con la resolución de problemas.

Asimismo, durante el desarrollo de la intervención se observó una evolución progresiva en la forma en que los estudiantes interactuaron con el chatbot. Inicialmente, el uso de la herramienta estuvo orientado principalmente a la obtención de respuestas; sin embargo, conforme avanzaron las sesiones, se identificó una mayor tendencia a contrastar procedimientos, verificar resultados y analizar posibles inconsistencias entre las respuestas generadas por el sistema y las soluciones elaboradas por los propios estudiantes. Estas evidencias reflejan una participación más activa en los procesos de análisis y reflexión requeridos para la resolución de problemas matemáticos.

El análisis de la matriz de evaluación aplicada durante el proceso de intervención permitió identificar una evolución progresiva en el desempeño del grupo experimental a lo largo de las seis sesiones desarrolladas. Los resultados muestran un incremento gradual en las habilidades asociadas con el análisis, la evaluación de procedimientos y la verificación de resultados durante la resolución de problemas matemáticos.

En las primeras sesiones se observó que los estudiantes utilizaban el chatbot principalmente como una fuente de consulta inmediata. Conforme avanzó la intervención, se registró una mayor participación en actividades relacionadas con la comparación de procedimientos, la identificación de inconsistencias y la revisión de respuestas obtenidas. Este cambio fue particularmente evidente a partir de la cuarta sesión, momento en el que comenzaron a manifestarse con mayor frecuencia conductas asociadas a la reflexión y al análisis de las estrategias de resolución.

Asimismo, se identificaron avances en las dimensiones de verificación y autorregulación. Aunque estas habilidades no se consolidaron de manera homogénea en todos los participantes, se observaron evidencias de una mayor tendencia a comprobar resultados, revisar procedimientos y justificar decisiones durante la resolución de problemas matemáticos.

Por su parte, el grupo control mostró variaciones menos pronunciadas a lo largo del periodo de estudio. Los resultados obtenidos indican que la mayoría de los estudiantes permaneció en niveles de desempeño similares a los observados en la evaluación inicial, particularmente en las dimensiones relacionadas con el análisis, la evaluación y la inferencia. Asimismo, no se identificaron cambios significativos en las estrategias de resolución empleadas durante las actividades desarrolladas.

Los registros obtenidos mediante la matriz de seguimiento muestran una tendencia ascendente sostenida en el desempeño del grupo experimental a lo largo de las seis sesiones de intervención. El puntaje promedio pasó de 5.8 puntos en la primera sesión a 13.1 puntos en la sesión final, lo que representa un incremento del 125% respecto al valor inicial. Los incrementos más pronunciados se observaron a partir de la cuarta sesión, coincidiendo con una mayor frecuencia de conductas relacionadas con la comparación de estrategias, la verificación de resultados y la evaluación de procedimientos matemáticos.

En la ilustración 1 muestra el incremento progresivo de los puntajes obtenidos en la matriz de seguimiento aplicada durante las seis sesiones de intervención. Se observa una tendencia ascendente sostenida, con un crecimiento más evidente a partir de la cuarta sesión.

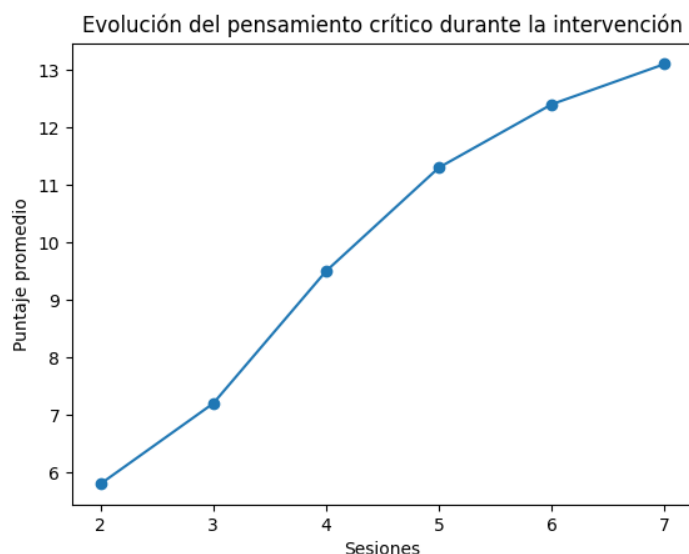


Fig. 1: Evolución del pensamiento crítico durante la intervención

Se observa una tendencia ascendente sostenida, con un crecimiento más marcado a partir de la sesión 4, lo que sugiere un efecto acumulativo de la intervención.

Los resultados obtenidos mediante la Escala Comparativa de Pensamiento Crítico (ECPC-M) y la Escala de Percepción sobre el Uso de Chatbots proporcionan información complementaria acerca de la autopercepción de los estudiantes respecto a las habilidades cognitivas empleadas durante la resolución de problemas matemáticos. En el grupo experimental, las respuestas se concentraron en las categorías de mayor frecuencia para los indicadores relacionados con la verificación de respuestas y la identificación de errores, ambos ubicados predominantemente en la opción siempre. Asimismo, los indicadores vinculados con la comparación de estrategias, la argumentación de procedimientos y la modificación de estrategias ante la detección de errores se situaron principalmente en la categoría frecuentemente. Estos resultados sugieren que los estudiantes perciben una participación recurrente en procesos asociados con el análisis, la evaluación y la revisión de procedimientos durante la resolución de problemas matemáticos. De igual manera, la frecuencia reportada en las categorías superiores indica una mayor conciencia sobre las estrategias empleadas para verificar resultados, identificar inconsistencias y justificar decisiones durante el proceso de resolución.

Tabla 2: Percepción de habilidades de pensamiento crítico durante la resolución de problemas: comparación entre grupos

Indicador de pensamiento crítico	Grupo experimental (frecuencia)	Grupo control (registros cualitativos)
Verificación de respuestas	Siempre	Ocasional
Identificación de errores	Siempre	Esporádica
Comparación de estrategias	Frecuentemente	Poco sistemática
Argumentación de procedimientos	Frecuentemente	No consolidada
Cambio de estrategia ante el error	Frecuentemente	Esporádico

Los resultados de la Escala Comparativa de Pensamiento Crítico (ECPC-M) y de la Escala de Percepción sobre el Uso de Chatbots proporcionan información complementaria acerca de la autopercepción de los

estudiantes respecto a las habilidades cognitivas empleadas durante la resolución de problemas matemáticos. En el grupo experimental, las respuestas se concentraron en las categorías de mayor frecuencia para los indicadores relacionados con la verificación de respuestas y la identificación de errores, ambos ubicados predominantemente en la opción siempre. Asimismo, los indicadores asociados con la comparación de estrategias, la argumentación de procedimientos y la modificación de estrategias ante la detección de errores se situaron principalmente en la categoría frecuentemente.

La escala Likert aplicada al grupo experimental mostró una alta frecuencia en habilidades relacionadas con el pensamiento crítico. En particular, el 80% de los participantes reportó que verifica las respuestas obtenidas en la categoría siempre, mientras que el 78% indicó identificar errores con la misma frecuencia. De igual manera, la comparación de estrategias fue reportada como frecuente por el 65% de los estudiantes y la argumentación de procedimientos por el 60%. Por otra parte, la Escala de Percepción sobre el Uso de Chatbots evidenció una valoración favorable de la herramienta como apoyo para la comprensión y resolución de problemas matemáticos. Las respuestas se distribuyeron principalmente entre las categorías de vez en cuando y frecuentemente, lo que indica que los estudiantes perciben su utilidad de manera diferenciada según el tipo de actividad desarrollada y las necesidades específicas de aprendizaje.

Asimismo, los indicadores relacionados con el cuestionamiento, la reflexión sobre procedimientos y la revisión de respuestas alcanzaron frecuencias elevadas dentro de las categorías superiores de la escala, lo que evidencia una percepción positiva respecto al papel del chatbot como recurso de apoyo durante el proceso de resolución de problemas. En conjunto, los resultados obtenidos muestran que los estudiantes del grupo experimental reportan una participación recurrente en actividades relacionadas con la verificación, el análisis y la revisión de procedimientos matemáticos durante el desarrollo de la intervención.

El análisis cualitativo de las evidencias recopiladas durante la intervención revela una transición claramente diferenciada en la forma en que ambos grupos interactuaron con el chatbot. En el grupo experimental, esta evolución siguió una trayectoria progresiva y consistente. Durante las primeras sesiones predominó un uso de carácter reproductivo, en el que los estudiantes consultaban la herramienta de manera inmediata y transcribían las respuestas obtenidas sin realizar procesos de reflexión o verificación. No obstante, a partir de la cuarta sesión se observó un cambio cualitativo sostenido: los participantes comenzaron a contrastar las respuestas generadas por el chatbot con sus propios procedimientos, identificar discrepancias y formular preguntas orientadas a comprender el origen de las diferencias encontradas. Este desplazamiento, desde una interacción predominantemente pasiva hacia una participación analítica y reflexiva, no ocurrió de manera simultánea en todos los estudiantes; sin embargo, la tendencia fue consistente y se consolidó hacia el final de la intervención.

En contraste, el grupo control mantuvo un patrón de uso heterogéneo e instrumental durante todo el periodo de estudio, sin evidenciar una evolución comparable. Mientras que el grupo experimental desarrolló prácticas sistemáticas de comparación, verificación y validación de resultados, en el grupo control predominó el uso directo de la herramienta sin un cuestionamiento posterior. Las respuestas generadas por el chatbot eran aceptadas o descartadas, pero rara vez sometidas a un análisis crítico. Esta diferencia no puede atribuirse al acceso a la tecnología, ya que ambos grupos dispusieron de la misma herramienta, sino a la ausencia de una mediación pedagógica orientada a promover procesos de análisis y evaluación. En consecuencia, los resultados sugieren que el potencial formativo de una herramienta basada en inteligencia artificial depende en gran medida del andamiaje didáctico que acompaña su integración en el aula.

En el grupo experimental, el chatbot fue percibido como un recurso cognitivamente útil, aunque su valoración no fue homogénea ni automática. Los estudiantes aprendieron a identificar los momentos en que la herramienta contribuía efectivamente a su comprensión y aquellos en los que su consulta resultaba innecesaria, lo que constituye un indicador relevante de desarrollo metacognitivo. Los ítems de la Escala de Percepción relacionados con el cuestionamiento activo y la reflexión se concentraron en la categoría *frecuentemente*, lo que sugiere que la interacción con el chatbot estuvo asociada, para la mayoría de los participantes, con la activación de procesos de análisis sobre la validez de sus propios procedimientos, más que con la recepción pasiva de respuestas.

Asimismo, la identificación de errores en las respuestas proporcionadas por el chatbot no generó una pérdida de confianza en la herramienta; por el contrario, promovió prácticas más activas de verificación y comprobación. Este comportamiento explica los elevados porcentajes observados en los indicadores de verificación e identificación de errores, los cuales alcanzaron el 80 % y el 78 %, respectivamente, en la categoría *siempre*.

Por su parte, los registros cualitativos del grupo control evidencian patrones de uso heterogéneos con una tendencia predominantemente instrumental. Algunos estudiantes recurrieron al chatbot de manera inmediata ante cualquier dificultad, sin intentar resolver previamente la tarea ni verificar la coherencia de la respuesta obtenida, mientras que otros optaron por evitar su uso de forma casi sistemática. En ninguno de los casos se identificaron evidencias consistentes de verificación, comparación de estrategias o cuestionamiento crítico de la información proporcionada. Esta divergencia entre grupos no puede explicarse

por diferencias en las capacidades cognitivas iniciales, ya que los resultados del pretest mostraron niveles de desempeño comparables. Más bien, los hallazgos sugieren que la diferencia radica en la presencia o ausencia de una mediación pedagógica orientada a promover procesos reflexivos. En este sentido, los resultados refuerzan la tesis central del estudio: el potencial educativo del chatbot no reside exclusivamente en la tecnología, sino en las condiciones pedagógicas que estructuran su utilización.

El análisis de las entrevistas semiestructuradas permitió profundizar en las experiencias cognitivas, las decisiones estratégicas y las percepciones subjetivas relacionadas con el uso del chatbot en una submuestra del grupo experimental. La información fue analizada mediante un proceso de codificación temática, que permitió identificar patrones recurrentes y complementar los hallazgos obtenidos a través de los instrumentos cuantitativos y del análisis del proceso de intervención.

- **Transformación en la resolución de problemas:** Los participantes describieron de manera consistente cambios en su forma de abordar las tareas, incorporando estrategias de revisión y contraste que no estaban presentes antes de la intervención. La resolución de problemas dejó de seguir una lógica lineal, basada en la aceptación de una única respuesta, para adoptar una dinámica iterativa caracterizada por la elaboración de una solución inicial, su comparación con las respuestas del chatbot y la evaluación posterior de la coherencia de los resultados obtenidos. Esta transformación se refleja en expresiones como: *"Antes solo resolvía como me acordaba, ahora primero lo intento y luego veo si coincide con lo que dice el chatbot"* (E1) y *"Ya no me voy con la primera respuesta, trato de ver si tiene sentido"* (E3). Estos hallazgos son consistentes con lo planteado por Schoenfeld (1992), quien define el monitoreo durante la resolución de problemas como una actividad metacognitiva que implica la evaluación continua de las propias estrategias en función de los objetivos planteados.
- **Uso estratégico y selectivo del chatbot:** Las entrevistas evidenciaron un patrón de uso cada vez más selectivo y estratégico de la herramienta, concentrado principalmente en situaciones de duda, validación o confirmación de procedimientos. Los estudiantes manifestaron haber desarrollado criterios para determinar cuándo resultaba pertinente consultar el chatbot y cuándo era preferible resolver la tarea de manera autónoma. Al respecto, señalaron: *"Lo uso cuando no entiendo un paso, no para todo el problema"* (E3) y *"Primero lo intento solo y si no puedo, entonces le pregunto"* (E2). Este comportamiento contrasta con el uso inicial reportado durante las primeras sesiones, caracterizado por consultas inmediatas y dependencia de las respuestas generadas por la herramienta: *"Al inicio sí copiaba, porque pensé que siempre estaba bien"* (E1). La transición observada desde una conducta dependiente hacia una utilización autónoma y estratégica constituye un indicador relevante de desarrollo metacognitivo.
- **Reconocimiento crítico del error y conciencia de verificación:** Uno de los hallazgos más relevantes fue el reconocimiento, por parte de los estudiantes, de que el chatbot puede cometer errores. Lejos de generar desconfianza generalizada, esta constatación favoreció el desarrollo de prácticas de evaluación crítica y verificación sistemática. Los participantes expresaron esta idea mediante afirmaciones como: *"Me di cuenta que a veces el chatbot se equivoca, entonces tengo que revisar"* (E1) y *"Si no lo compruebo, no sé si está bien"* (E3). La conciencia sobre la falibilidad de los sistemas automatizados activó la necesidad de contrastar y evaluar la información recibida, lo que constituye una manifestación directa de la habilidad de evaluación descrita por Facione (2007). Este resultado adquiere especial relevancia en el contexto del debate sobre la dependencia tecnológica, ya que sugiere que una integración pedagógicamente mediada de la inteligencia artificial puede fortalecer, en lugar de debilitar, la autonomía crítica del estudiante.
- **Indicios de reflexión metacognitiva:** Finalmente, se identificaron testimonios que evidencian procesos de reflexión sobre el propio aprendizaje, particularmente en relación con la comprensión de los procedimientos utilizados para resolver problemas. Los estudiantes señalaron: *"Ahora pienso más en por qué hago cada paso"* (E2) y *"Antes solo seguía pasos, ahora trato de entenderlos"* (E1). Aunque la profundidad de estas reflexiones fue variable entre los participantes, situación coherente con los niveles intermedios observados en la dimensión de metacognición evaluada mediante la rúbrica analítica, su presencia sugiere el inicio de procesos de autorregulación que trascienden la mera ejecución algorítmica. En este sentido, Goldstein y Calero (2022) sostienen que la metacognición implica no solo reconocer lo que se sabe, sino también tomar decisiones estratégicas orientadas a mejorar la comprensión y el aprendizaje, planteamiento que describe de manera precisa la trayectoria observada en los testimonios analizados.

La triangulación de los datos obtenidos mediante los cinco instrumentos permite identificar un patrón convergente que otorga solidez interpretativa a los hallazgos. Los tres ejes de convergencia se presentan a continuación.

Tabla 3: Categorías emergentes comparadas: grupo experimental y grupo control.

Categoría emergente	Grupo experimental	Grupo control
Uso del chatbot	Progresivo y reflexivo	Heterogéneo e instrumental
Desarrollo del pensamiento crítico	Avance en análisis e inferencia; mejora en evaluación	Limitado; comprensión básica
Verificación y autorregulación	Incorporación gradual y sistemática	Esporádica y no sistemática

En primer lugar, los datos de la matriz de intervención, la escala de percepción y la entrevista coinciden en señalar una transición en el grupo experimental desde un uso reproductivo hacia un uso estratégico-reflexivo del chatbot. En segundo lugar, los resultados del posttest, las evidencias del proceso y los testimonios cualitativos convergen en mostrar avances significativos en análisis e inferencia, así como una incorporación progresiva de la verificación como práctica cognitiva. Finalmente, tanto los datos cuantitativos como los cualitativos revelan que el estudiantado desarrolla una conciencia creciente sobre la necesidad de evaluar y comprobar la información, especialmente cuando proviene de fuentes automatizadas.

Estas convergencias robustecen la interpretación de los resultados y permiten afirmar que la intervención didáctica mediada por el uso guiado del chatbot constituye un factor asociado al desarrollo del pensamiento crítico en el contexto específico de la resolución de problemas matemáticos en educación secundaria. En este sentido, los resultados indican que la intervención didáctica mediada por el uso guiado del chatbot contribuye de manera relevante al desarrollo del pensamiento crítico, particularmente en procesos de análisis, verificación y argumentación.

Los hallazgos de esta investigación aportan evidencia empírica situada sobre la relación entre el uso pedagógicamente mediado de chatbots de inteligencia artificial y el desarrollo de habilidades de pensamiento crítico en estudiantes de educación secundaria en el área de matemáticas. La discusión que sigue examina estos resultados a la luz de la literatura existente, identifica convergencias y divergencias con estudios previos, analiza las implicaciones teóricas y prácticas derivadas, y explicita las limitaciones del estudio junto con posibles líneas de investigación futura.

El hallazgo más relevante de este estudio es que el uso guiado del chatbot, dentro de una intervención pedagógica estructurada, se asocia con mayores niveles de desarrollo del pensamiento crítico en comparación con el uso libre de la misma herramienta. Este resultado converge con las advertencias de Cabero-Almenara et al. (2023), quienes señalan que la integración de tecnologías digitales promueve el pensamiento crítico cuando se orienta hacia la reflexión y la construcción activa del conocimiento, y no cuando se limita a la automatización de respuestas. En este sentido, la diferencia observada entre el grupo experimental y el grupo control no parece explicarse por la disponibilidad tecnológica en sí misma, sino por la calidad del diseño didáctico que acompañó su utilización.

Esta interpretación se ve reforzada por los datos cualitativos obtenidos durante la intervención. Los estudiantes del grupo experimental describieron un proceso de transformación en su forma de resolver problemas que implicó revisar, comparar y cuestionar las respuestas obtenidas, conductas que no emergieron de manera consistente en condiciones de uso libre. Por el contrario, en el grupo control predominó una interacción de carácter instrumental, con escasas evidencias de análisis o validación. Este contraste evidencia la relevancia de la mediación docente como condición necesaria para que la tecnología contribuya al fortalecimiento de procesos cognitivos de orden superior.

Asimismo, este hallazgo dialoga directamente con la hipótesis planteada al inicio del estudio, según la cual los estudiantes que utilizan chatbots como apoyo en la resolución de problemas, bajo condiciones de mediación pedagógica, alcanzan niveles más altos de pensamiento crítico en dimensiones como el análisis y la inferencia. La evidencia disponible respalda parcialmente esta hipótesis, ya que los mayores avances se observaron precisamente en dichas dimensiones; sin embargo, las habilidades de evaluación, argumentación y autorregulación mostraron progresos más moderados, lo que sugiere que el desarrollo integral del pensamiento crítico requiere intervenciones más prolongadas y sistemáticas.

Los resultados muestran que no todas las dimensiones del pensamiento crítico se desarrollan con la misma velocidad ni en el mismo grado. Las habilidades de análisis e inferencia registraron los mayores avances, mientras que la evaluación, la argumentación y la autorregulación presentaron progresos más lentos y

heterogéneos. Esta diferencia requiere una interpretación que vaya más allá de la simple descripción de los resultados.

Las habilidades de análisis e inferencia parecen activarse de manera relativamente directa durante la interacción con el chatbot. Comparar la respuesta propia con la proporcionada por el sistema, identificar diferencias en los procedimientos y considerar explicaciones alternativas son conductas que la dinámica de interacción con la herramienta puede favorecer, especialmente cuando el estudiante detecta discrepancias entre ambas respuestas. En contraste, la evaluación entendida como la capacidad para valorar la validez y pertinencia de una afirmación o procedimiento y la autorregulación entendida como la capacidad para monitorear y ajustar el propio proceso de aprendizaje requieren niveles más elevados de reflexión crítica que difícilmente emergen de forma espontánea, por lo que demandan un andamiaje pedagógico explícito y sostenido.

Estos resultados son consistentes con lo planteado por Sandoval Peña et al. (2025), quienes señalan que la mediación tecnológica debe fortalecer deliberadamente la autorregulación y la reflexión crítica para evitar la dependencia cognitiva y favorecer aprendizajes significativos. Desde esta perspectiva, los diseños de intervención que incorporen chatbots en el aula deberían incluir actividades específicamente orientadas al desarrollo de la evaluación crítica y la metacognición, en lugar de asumir que estas habilidades surgirán automáticamente como consecuencia del uso de la herramienta.

En el ámbito de la educación matemática, estos hallazgos coinciden con los reportados por Huamán Valencia et al. (2025), quienes sostienen que las estrategias de resolución de problemas favorecen el desarrollo de competencias cognitivas complejas al promover la exploración de alternativas y la argumentación fundamentada. En este contexto, el chatbot puede interpretarse como un recurso que amplía las oportunidades de exploración cognitiva mediante la confrontación con soluciones alternativas que enriquecen el razonamiento del estudiante cuando son aprovechadas pedagógicamente.

Uno de los hallazgos cualitativos más relevantes de este estudio es que el reconocimiento de la falibilidad del chatbot por parte de los estudiantes actuó como un detonador del pensamiento crítico. Lejos de generar una pérdida de confianza en la herramienta, la constatación de que el sistema automatizado puede equivocarse promovió la necesidad de verificar la información recibida y, con ello, el ejercicio de la evaluación como habilidad cognitiva. Este resultado posee importantes implicaciones para el debate sobre la dependencia tecnológica en educación.

Una de las preocupaciones más frecuentes en la literatura sobre inteligencia artificial y educación secundaria es que el acceso inmediato a respuestas automatizadas disminuya la disposición de los estudiantes para razonar de manera independiente, generando lo que Correal Romero (2025) denomina dependencia cognitiva. Sin embargo, los datos obtenidos en esta investigación matizan esta preocupación. Cuando los estudiantes participan en experiencias pedagógicamente mediadas que promueven la comparación, la verificación y el cuestionamiento de las respuestas generadas por el chatbot, la dependencia cognitiva no constituye el resultado predominante. Por el contrario, la conciencia de que el sistema puede cometer errores parece contribuir significativamente a la activación del pensamiento crítico, al exigir que el estudiante ejerza su propio juicio para validar la información obtenida.

Esta interpretación coincide con la definición de pensamiento crítico propuesta por Facione (2007), quien lo describe como un juicio autorregulado y deliberado que conduce a procesos de interpretación, análisis, evaluación e inferencia. La experiencia de identificar errores en las respuestas del chatbot obliga al estudiante a ejercer dicho juicio de manera auténtica, ya que no puede delegar completamente la verificación en una fuente externa. En este sentido, la tecnología puede constituirse en un recurso que favorece el desarrollo de la autonomía cognitiva del estudiante.

Desde una perspectiva teórica, los resultados de este estudio contribuyen a precisar las condiciones bajo las cuales la mediación tecnológica basada en inteligencia artificial puede favorecer el desarrollo del pensamiento crítico en estudiantes de educación secundaria. En primer lugar, se confirma que la mera disponibilidad de la herramienta no es suficiente para producir cambios significativos; la calidad del diseño pedagógico que estructura su utilización resulta determinante. En segundo lugar, se identifican habilidades que parecen responder de manera más inmediata a la mediación con chatbots análisis e inferencia y otras que requieren intervenciones pedagógicas más específicas evaluación, argumentación y metacognición. Esta diferenciación aporta una comprensión más precisa de los procesos involucrados en el desarrollo del pensamiento crítico.

Desde una perspectiva práctica, los hallazgos ofrecen orientaciones concretas para el diseño de experiencias de aprendizaje que incorporen chatbots en el aula. La evidencia sugiere que estas experiencias deberían estructurarse en torno a tres principios fundamentales: primero, promover la confrontación explícita entre las respuestas elaboradas por el estudiante y las proporcionadas por el chatbot; segundo, incorporar actividades sistemáticas de verificación que conviertan la evaluación en una práctica habitual; y tercero, fomentar la discusión colectiva sobre los errores y limitaciones de la herramienta como una oportunidad para fortalecer el juicio crítico, la alfabetización digital y la evaluación de la calidad de las fuentes de

información. Estas competencias resultan especialmente relevantes en el marco de la Nueva Escuela Mexicana, la cual promueve la formación de ciudadanos capaces de analizar y evaluar críticamente su realidad (SEP, 2022).

Respecto a la formación docente, los resultados respaldan las observaciones de Sossa Páez (2025) sobre la necesidad de una preparación contextualizada que garantice que la inteligencia artificial contribuya a la formación de ciudadanos críticos y éticos. En consecuencia, los docentes no solo requieren conocimientos técnicos sobre el funcionamiento de estas herramientas, sino también competencias didácticas que les permitan diseñar estrategias de mediación orientadas al fortalecimiento de la autonomía cognitiva del estudiante.

Este estudio presenta diversas limitaciones que deben considerarse al interpretar sus resultados. En primer lugar, el tamaño reducido de la muestra 15 estudiantes pertenecientes a un mismo centro escolar limita la generalización estadística de los hallazgos. Dado el carácter exploratorio de la investigación, los resultados deben entenderse como evidencia contextualizada y preliminar que requiere ser corroborada mediante estudios desarrollados en contextos más amplios y con muestras representativas.

En segundo lugar, el periodo de intervención fue relativamente breve para promover el desarrollo pleno de habilidades metacognitivas complejas como la evaluación y la autorregulación, cuya consolidación demanda experiencias de aprendizaje más prolongadas y diversas. Esta situación podría explicar parcialmente los avances más moderados observados en dichas dimensiones en comparación con las habilidades de análisis e inferencia.

En tercer lugar, el diseño cuasi experimental sin asignación aleatoria introduce una posible fuente de sesgo que no puede descartarse completamente, aunque la equivalencia inicial observada entre los grupos en el pretest atenúa esta limitación. Asimismo, el uso de un muestreo por conveniencia restringe la diversidad de contextos representados en el estudio.

A partir de estas limitaciones, se identifican diversas líneas de investigación futura. En primer lugar, resulta necesario replicar el estudio con muestras más amplias y procedimientos de asignación aleatoria que permitan incrementar la robustez metodológica de los hallazgos. En segundo lugar, sería pertinente analizar los efectos diferenciales de distintos tipos de chatbots, considerando variables como el grado de especialización matemática y la capacidad de ofrecer retroalimentación adaptativa. En tercer lugar, investigaciones longitudinales permitirían examinar la persistencia de los efectos observados y su posible transferencia a otros ámbitos del conocimiento. Finalmente, se considera relevante profundizar en el papel mediador del docente, identificando las estrategias específicas de andamiaje que potencian el uso del chatbot como herramienta para el desarrollo del pensamiento crítico.

Los resultados de esta investigación también generan reflexiones relevantes desde una perspectiva ética. La incorporación de chatbots basados en inteligencia artificial en contextos educativos formales implica responsabilidades pedagógicas e institucionales que trascienden el rendimiento académico. Entre ellas destaca la necesidad de establecer lineamientos claros sobre cuándo, cómo y con qué finalidad deben utilizarse estas herramientas, favoreciendo que los estudiantes desarrollen criterios propios de uso y eviten reducir su función a la simple obtención de respuestas.

En este sentido, los resultados obtenidos son alentadores. Cuando la interacción con el chatbot se encuentra adecuadamente mediada desde el punto de vista pedagógico, los estudiantes no solo fortalecen habilidades de pensamiento crítico vinculadas al contenido matemático, sino que también desarrollan una actitud de escepticismo funcional frente a la información automatizada. Esta disposición a verificar, cuestionar y evaluar críticamente las fuentes constituye una competencia transversal fundamental para el siglo XXI y representa uno de los principales aportes formativos que puede ofrecer una integración crítica y responsable de la inteligencia artificial en la educación secundaria.

CONCLUSIÓN

Esta investigación se propuso analizar el impacto del uso de chatbots de inteligencia artificial en el desarrollo de habilidades de pensamiento crítico en estudiantes de tercer grado de secundaria durante la resolución de problemas matemáticos. Los hallazgos obtenidos permiten trazar un panorama claro sobre las condiciones en las que la tecnología puede constituirse como un recurso genuinamente formativo, y no simplemente como un mecanismo de acceso a respuestas.

El primer hallazgo central que se desprende del estudio es que el uso guiado del chatbot, integrado en un diseño didáctico con intención pedagógica explícita, se asocia de manera consistente con un mayor desarrollo del pensamiento crítico en comparación con su uso libre y no mediado. Este resultado no es trivial: confirma que la herramienta tecnológica, por sí sola, carece de valor formativo. Es el diseño de la intervención la estructura de las tareas, el rol orientador del docente y la secuencia progresiva de actividades lo que convierte al chatbot en un catalizador de procesos cognitivos superiores. El grupo experimental, que participó en una intervención organizada en seis sesiones bajo principios de mediación pedagógica, alcanzó

niveles avanzados en las dimensiones de análisis e inferencia en un 62% de los casos, frente al 29% registrado en el grupo control, donde el uso de la herramienta quedó a criterio exclusivo del estudiante.

En segundo lugar, el estudio identifica que no todas las habilidades del pensamiento crítico responden con la misma velocidad ni profundidad a la intervención. El análisis y la inferencia habilidades que se activan de forma casi natural al confrontar la propia solución con la respuesta generada por el sistema mostraron los avances más significativos. La evaluación, la argumentación y la autorregulación, en cambio, requieren un andamiaje pedagógico más deliberado y sostenido para consolidarse. Este hallazgo tiene implicaciones prácticas directas: los diseños de intervención que incorporen IA en el aula deben atender específicamente estas dimensiones mediante actividades que promuevan la verificación sistemática, la justificación argumentada y la reflexión metacognitiva, sin asumir que estas prácticas emergerán de manera espontánea. Un tercer resultado de especial relevancia es el que se refiere al papel que juega la conciencia sobre la falibilidad del chatbot en el desarrollo del pensamiento crítico. Contrario a lo que podría suponerse, el reconocimiento de que el sistema automatizado comete errores no generó desconfianza ni abandono de la herramienta, sino que se convirtió en un detonador de evaluación activa. Cuando los estudiantes descubrieron que las respuestas del chatbot podían ser incorrectas, se vieron obligados a ejercer su propio juicio para verificar, corregir y fundamentar. Esta dinámica ilustra un mecanismo pedagógico de alto valor: la falibilidad del sistema, lejos de ser un obstáculo, puede convertirse en una oportunidad estructural para el desarrollo del pensamiento crítico, siempre que la intervención esté diseñada para aprovecharla.

En cuarto lugar, los datos evidencian que el uso del chatbot por parte del grupo experimental transitó, a lo largo de la intervención, desde una modalidad predominantemente reproductiva consultar y transcribir sin mediación reflexiva hacia una modalidad analítico-estratégica, caracterizada por la comparación de procedimientos, la detección de inconsistencias y la formulación de preguntas orientadas a la comprensión. Este tránsito no ocurrió de forma simultánea en todos los estudiantes ni en todas las sesiones, pero su trayectoria es clara y sostenida, y constituye en sí mismo un indicador de desarrollo metacognitivo. El grupo control, en contraste, mantuvo patrones heterogéneos e instrumentales durante todo el período, sin evidencias de procesos similares de evolución en el uso de la herramienta.

Desde una perspectiva más amplia, este estudio aporta evidencia empírica localizada sobre una pregunta que la literatura educativa ha planteado con insistencia creciente: ¿bajo qué condiciones la inteligencia artificial promueve el pensamiento crítico en lugar de sustituirlo? La respuesta que emerge de los datos no es simple ni binaria. La misma herramienta que, en ausencia de mediación pedagógica, favorece la dependencia cognitiva y el uso instrumental, puede convertirse, cuando se integra en un diseño didáctico orientado a la reflexión y la evaluación, en un instrumento de empoderamiento cognitivo. La clave no reside en la tecnología sino en la intencionalidad pedagógica con la que se utiliza.

Este hallazgo cobra especial relevancia en el contexto del sistema educativo mexicano, donde la Nueva Escuela Mexicana promueve la formación de ciudadanos con capacidad crítica para interpretar y evaluar su realidad. La integración responsable de herramientas de inteligencia artificial en el aula no puede reducirse a garantizar el acceso tecnológico; exige también que los docentes cuenten con formación específica para diseñar experiencias de aprendizaje en las que la IA amplifique el razonamiento humano, en lugar de reemplazarlo. En este sentido, la investigación apoya la necesidad de incluir el desarrollo de competencias tecnopedagógicas en los procesos de formación y actualización docente, particularmente en lo que respecta al uso de herramientas de IA generativa en educación básica.

En cuanto a las limitaciones del estudio, el tamaño reducido de la muestra y el carácter cuasi-experimental del diseño restringen la generalización de los resultados. Los hallazgos presentados deben entenderse como evidencia exploratoria situada, válida para orientar hipótesis y diseños de investigación más amplios, pero no como conclusiones definitivas sobre el efecto de los chatbots en el pensamiento crítico a escala poblacional. Asimismo, el período de intervención resultó insuficiente para el desarrollo completo de habilidades metacognitivas como la autorregulación, cuya consolidación requiere procesos de aprendizaje más prolongados.

Con todo, este estudio contribuye a precisar el mapa del desarrollo del pensamiento crítico en contextos mediados por IA, identificando qué habilidades responden con mayor celeridad, cuáles demandan intervención adicional y qué condiciones pedagógicas resultan determinantes. Sus resultados abren una agenda de investigación concreta: estudios con muestras más amplias, diseños longitudinales que permitan examinar la permanencia de los efectos, y análisis comparativos de distintos tipos de chatbots y modalidades de mediación docente. La pregunta sobre cómo integrar la inteligencia artificial en la educación de manera que fortalezca, y no erosione, la autonomía cognitiva del estudiante, sigue abierta. Esta investigación ofrece algunas respuestas parciales y, sobre todo, un conjunto de condiciones que permiten formular esa pregunta con mayor precisión.

REFERENCIAS

- Asociación Médica Mundial. (2013). Declaración de Helsinki: Principios éticos para las investigaciones médicas en seres humanos. <https://www.wma.net/es/policies-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Brookhart, S. M. (2018). *How to use rubrics effectively*. Educational Leadership, 65(4), 14–19. <https://www.ascd.org/el/articles/how-to-use-rubrics-effectively>
- Cabero-Almenara, J., Barroso-Osuna, J., Guillén-Gámez, F. D., & Palacios-Rodríguez, A. (2025). Creencias pedagógicas docentes y su aceptación de la inteligencia artificial en la educación superior: un estudio comparativo entre países. *Aula Abierta*, 54(3), 257–268. <https://doi.org/10.17811/rifie.21273>
- Cabero-Almenara, J., Llorente-Cejudo, C., & Guillén-Gámez, F. D. (2023). *La inteligencia artificial en la educación: análisis bibliométrico y tendencias*. Comunicar, 31(74), 9–21. <https://doi.org/10.3916/C74-2023-01>
<https://www.revistacomunicar.com/ojs/index.php/comunicar/article/view/1241>
- Clason, D. L., & Dormody, T. J. (1994). Analyzing data measured by individual Likert-type items. *Journal of Agricultural Education*, 35(4), 31–35. <https://doi.org/10.5032/jae.1994.04031>
- Comunicar. (2026). *Revisión sistemática de estudios sobre pensamiento crítico e inteligencia artificial en educación secundaria*. Comunicar, 84. Recuperado de <https://www.revistacomunicar.com/ojs/index.php/comunicar/article/download/116785/5212/29322>
- Correal Romero, T. F. (2025). *La influencia de la inteligencia artificial en el pensamiento crítico en los estudiantes de secundaria*. Dialéctica, 1(25), 150–177. <https://doi.org/10.56219/dialectica.v1i25.3882>
- Ennis, R. H. (2011). *The nature of critical thinking: An outline of critical thinking dispositions and abilities*. Recuperado de https://education.illinois.edu/docs/default-source/faculty-documents/robert-ennis/thenatureofcriticalthinking_51711_000.pdf
- Erazo Molina, S. del P., Bustillos Beltrán, C. K., Carvajal Jiménez, M. A., & Chingo Chiluisa, L. P. (2023). *La inteligencia artificial como mediadora en el desarrollo del pensamiento crítico: implicaciones pedagógicas y transformación del rol docente*. Revista Imaginario Social. <https://doi.org/10.59155/is.v9i1.354>
- Facione, P. A. (2007). *Pensamiento crítico: ¿Qué es y por qué es importante?* Insight Assessment. <https://eduteka.icesi.edu.co/pdfdir/PensamientoCriticoFacione.pdf>
- Goldstein, J., & Calero, C. I. (2022). ¿De qué hablamos cuando hablamos de metacognición en el aula? *Journal of Neuroeducation*, 3(1), 45–56. <https://doi.org/10.1344/joned.v3i1.39565>
- González, J. (2024). *Inteligencia artificial en educación: oportunidades y retos contemporáneos*. Revista de Tecnología Educativa, 20(1), 100–118. <https://revistatecnoseduca.org/index.php/rte/article/view/234>
- Hernández Sampieri, R., Fernández Collado, C., & Baptista Lucio, P. (2014). *Metodología de la investigación* (6ª ed.). McGraw-Hill Interamericana <http://187.191.86.244/rceis/registro/Metodolog%C3%ADa%20de%20la%20Investigaci%C3%B3n%20SAMPIERI.pdf>
- Huamán Valencia, C. N., Huamán Valencia, E. M., Huamán Valencia, D. M., & Ramírez Ríos, A. (2025). Estrategias efectivas para la resolución de problemas matemáticos en la educación secundaria: Una revisión sistemática. *Revista Tribunal*, 5(11), 129–142. <https://doi.org/10.59659/revistatribunal.v5i11.146>
- Lucana Wehr, Y. E., & Roldán Baluis, W. L. (2023). *Chatbot basado en inteligencia artificial para la educación escolar*. Horizontes, 15(3), 75–89. <https://revistahorizontes.org/index.php/revistahorizontes/article/view/1039>
- Mosqueda Chávez, E. (2024). *La inteligencia artificial como aliada del aprendizaje y el pensamiento crítico*. Revista Mexicana de Bachillerato a Distancia, 16(32). <https://doi.org/10.22201/cuaieed.20074751e.2024.32.89555>
- Olivares O., S. L., & Heredia Escorza, Y. (2012). Desarrollo del pensamiento crítico en estudiantes de educación media. *Revista Electrónica de Investigación Educativa*, 14(2), 87–104. <https://redie.uabc.mx/redie/article/view/299>

- Otzen, T., & Manterola, C. (2017). Técnicas de muestreo sobre una población a estudio. *International Journal of Morphology*, 35(1), 227–232. <https://scielo.conicyt.cl/pdf/ijmorphol/v35n1/art37.pdf>
- Sandoval Peña, J. M., López Delgado, C., & Pérez Rodríguez, S. (2025). *La metacognición en entornos educativos mediados por IA: desafíos y oportunidades*. *Revista Educativa Iberoamericana*, 12(1), 45–63. <https://doi.org/10.37471/rei.v12i1.215> <https://revistaeducativaiberoamericana.org/index.php/rei/article/view/215>
- Schoenfeld, A. (1992). *Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics*. En D. A. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 334–370). Macmillan. https://www.researchgate.net/publication/289963462_Learning_to_think_mathematically_Problem_solving_metacognition_and_sense_making_in_mathematics https://faculty.tarleton.edu/brawner/wp-content/uploads/sites/11/2023/02/Schoenfeld_MathThinking.pdf
- Secretaría de Educación Pública. (2022). *Plan de estudios para la educación preescolar, primaria y secundaria: Nueva Escuela Mexicana*. SEP. https://www.sev.gob.mx/educacion-basica/secundaria/wp-content/uploads/2025/09/Plan-de-Estudio-ISBN-ELECTRONICO.pdf?utm_source=chatgpt.com
- Secretaría de Educación Pública. (2022, 19 de agosto). *Acuerdo número 14/08/22 por el que se establece el Plan de Estudios para la educación preescolar, primaria y secundaria*. *Diario Oficial de la Federación*.
- Sossa Páez, J. A. (2025). *Inteligencia artificial en la educación secundaria: mediadora del pensamiento crítico-digital*. *Sinopsis Educativa*, 24(2), 289–298. https://revistas.upel.edu.ve/index.php/sinopsis_educativa/article/view/4317
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425–478. <https://doi.org/10.2307/30036540>