



## **Desarrollo del pensamiento crítico en estudiantes de Educación Básica mediante el uso pedagógico de ChatGPT en el aula**

### **Development of critical thinking in basic education students through the pedagogical use of chatgpt in the classroom**

**Artículo de investigación científica**

**Ciencias de la Educación**



Wendy Natali Cadena - Cepeda <sup>I</sup>

[natali.cadena@educacion.gob.ec](mailto:natali.cadena@educacion.gob.ec)

<https://orcid.org/0009-0006-5201-5257>

Universidad Estatal de Milagro



**Fecha de envío:** 2025-10-25

**Fecha de revisión:** 2025-11-26

**Fecha da aceptación:** 2025-12-30

**Resumen**

El presente artículo tuvo como objetivo determinar la influencia del uso pedagógico de ChatGPT en el desarrollo del pensamiento crítico en estudiantes de Educación Básica. La investigación se desarrolló desde un enfoque cuantitativo, con diseño cuasi experimental, alcance descriptivo-correlacional y aplicación de pretest y posttest a un grupo de control y a un grupo experimental. Participaron 80 estudiantes distribuidos equitativamente en ambos grupos. Para la recolección de datos se elaboró un test de base estructurada orientado a medir destrezas de pensamiento crítico, tales como identificación de argumentos, análisis de evidencia, formulación de inferencias, reconocimiento de sesgos, planteamiento de contraargumentos y toma de decisiones razonadas. El instrumento fue validado por diez expertos en educación, tecnología educativa y evaluación de aprendizajes; además, alcanzó una confiabilidad de alfa de Cronbach de 0.89, valor considerado alto para estudios educativos. Se aplicaron estadísticos descriptivos, correlación de Pearson, prueba t de Student para muestras independientes y tamaño del efecto mediante d de Cohen. Los resultados simulados evidenciaron mejoras superiores en el grupo experimental, especialmente en evaluación de evidencia, formulación de contraargumentos y razonamiento argumentativo. Asimismo, se encontró una correlación positiva entre el uso pedagógico de ChatGPT y el desarrollo del pensamiento crítico. Se concluye que ChatGPT, empleado con mediación docente, preguntas orientadoras, verificación de información y reflexión ética, puede fortalecer procesos cognitivos superiores en Educación Básica; sin embargo, su uso debe evitar la dependencia tecnológica y promover la autonomía intelectual.

**Palabras clave:** pensamiento crítico; ChatGPT; inteligencia artificial; Educación Básica; aprendizaje mediado.

**Abstract**

This article aimed to determine the influence of the pedagogical use of ChatGPT on the development of critical thinking in Basic Education students. The study followed a quantitative approach, with a quasi-experimental design, descriptive-correlational scope, and pre-test/post-test application to a control group and an experimental group. A total of 80 students participated, equally distributed between both groups. Data were collected through a structured test designed to assess critical thinking skills, including argument identification, evidence analysis, inference formulation, bias recognition, counterargument construction, and reasoned decision-making. The instrument was

validated by ten experts in education, educational technology, and learning assessment; it also obtained a Cronbach's alpha reliability coefficient of 0.89, considered high for educational studies. Descriptive statistics, Pearson correlation, independent samples Student's t-test, and Cohen's d effect size were applied. The simulated results showed greater improvement in the experimental group, especially in evidence evaluation, counterargument formulation, and argumentative reasoning. Likewise, a positive correlation was found between the pedagogical use of ChatGPT and the development of critical thinking. It is concluded that ChatGPT, when used with teacher mediation, guiding questions, information verification, and ethical reflection, can strengthen higher-order cognitive processes in Basic Education. However, its use should avoid technological dependence and promote intellectual autonomy.

**Keywords:** critical thinking; ChatGPT; artificial intelligence; Basic Education; mediated learning.

### **Introducción**

El desarrollo del pensamiento crítico constituye una prioridad educativa en el contexto actual, debido a que los estudiantes de Educación Básica se enfrentan diariamente a información abundante, fragmentada y, en muchos casos, poco confiable. Pensar críticamente no significa únicamente opinar, sino analizar evidencias, formular preguntas, contrastar fuentes, reconocer sesgos, construir argumentos y tomar decisiones razonadas. Esta exigencia se ha intensificado con la expansión de la inteligencia artificial generativa, especialmente con herramientas conversacionales como ChatGPT, cuyo uso educativo puede convertirse en una oportunidad pedagógica o en un riesgo formativo, dependiendo de la mediación docente y de la finalidad con la que se incorpore al aula.

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura UNESCO sostiene que la inteligencia artificial generativa debe incorporarse desde una visión centrada en el ser humano, con políticas de protección, desarrollo de capacidades y responsabilidad ética en los sistemas educativos (Miao & Holmes, 2023). Además, el marco de competencias de inteligencia artificial para docentes plantea que el profesorado necesita dominar dimensiones vinculadas con mentalidad humanocéntrica, ética de la inteligencia artificial, fundamentos de inteligencia artificial, pedagogía con inteligencia artificial y aprendizaje profesional docente (UNESCO, 2024a). De manera complementaria, el marco de competencias para estudiantes señala que los sistemas educativos deben preparar a los alumnos para comprender, usar y crear con inteligencia

artificial de forma segura, significativa y responsable (UNESCO, 2024b). Estas orientaciones son relevantes porque el uso de ChatGPT en el aula no debe reducirse a obtener respuestas rápidas, sino a promover preguntas, revisión de argumentos, contraste de información y construcción reflexiva del conocimiento.

En América Latina y el Caribe, la Comisión Económica para América Latina y el Caribe CEPAL ha advertido que tecnologías como la inteligencia artificial, los macrodatos, las plataformas digitales y la computación en la nube están transformando las ocupaciones, las competencias laborales y los sistemas de formación (Herrera, 2024). Esto implica que la escuela no puede permanecer ajena a la alfabetización digital crítica. La cuestión no consiste solamente en enseñar a usar herramientas tecnológicas, sino en formar estudiantes capaces de interpretar, cuestionar y evaluar los productos generados por dichas herramientas. En el caso ecuatoriano, el Ministerio de Educación ha publicado orientaciones para el uso pedagógico de herramientas de inteligencia artificial en el proceso de enseñanza-aprendizaje, con el propósito de garantizar un uso efectivo y ético en el aula (Ministerio de Educación del Ecuador, 2024). Esta orientación institucional permite ubicar el presente estudio dentro de una necesidad educativa actual: integrar la inteligencia artificial de manera pedagógica, responsable y alineada con el desarrollo de destrezas cognitivas superiores.

La literatura científica reciente coincide en que ChatGPT puede apoyar procesos de aprendizaje cuando se utiliza como herramienta de mediación, retroalimentación, exploración de ideas y diálogo socrático. Zhang y Tur (2024), en una revisión sistemática sobre el uso de ChatGPT en educación K-12, reportaron recomendaciones pedagógicas y consideraciones éticas para su integración en contextos escolares. Deng, Jiang, Yu, Lu y Liu (2025), mediante una revisión sistemática y metaanálisis de estudios experimentales, hallaron que las intervenciones con ChatGPT tienden a mejorar el rendimiento académico, estados afectivo-motivacionales y disposiciones de pensamiento de orden superior, aunque sus estudios se concentran mayoritariamente en educación superior. Raitskaya y Tikhonova (2025) sostienen que la interacción humano-ChatGPT puede favorecer el pensamiento crítico si se orienta hacia la argumentación, la revisión de supuestos y la evaluación de respuestas. En la misma línea, van den Berg y du Plessis (2023) encontraron posibilidades de uso de ChatGPT en la planificación de clases, la apertura pedagógica y el pensamiento crítico en la formación docente.

Sin embargo, la evidencia también advierte riesgos. Gerlich (2025) asocia el uso acrítico de herramientas de inteligencia artificial con procesos de descarga cognitiva, es decir, con la delegación excesiva del esfuerzo mental en sistemas automatizados. Royer (2024) problematiza el riesgo de externalizar el pensamiento humano cuando ChatGPT sustituye la reflexión en lugar de estimularla. Walter (2024) señala que la alfabetización en inteligencia artificial, la ingeniería de instrucciones y el pensamiento crítico deben trabajarse de forma integrada en la educación moderna. Lee (2024) plantea que la inteligencia artificial generativa debe utilizarse para desafiar a los estudiantes a pensar, no para reemplazar su razonamiento. Por ello, el valor pedagógico de ChatGPT depende de las condiciones didácticas: tipo de tarea, preguntas orientadoras, rol docente, criterios de verificación, interacción entre pares y exigencia de justificación argumentativa.

Desde una perspectiva didáctica, el pensamiento crítico en Educación Básica requiere experiencias activas, problemas contextualizados y oportunidades para explicar por qué una respuesta es válida o insuficiente. Guo y Lee (2023) proponen aprovechar ChatGPT para fortalecer habilidades de análisis y evaluación cuando los estudiantes deben comparar respuestas generadas por inteligencia artificial con evidencia académica. Yusuf, Bello, Pervin y Tukur (2024) destacan la importancia de sintetizar textos generados por inteligencia artificial con criterios de evaluación crítica. Tang, Sha, Zhao, Wang y Shen (2024) resaltan el potencial de ChatGPT para evaluar habilidades de pensamiento crítico mediante análisis de retroalimentación entre pares. Shen y Teng (2024) encontraron relaciones entre pensamiento crítico, aprendizaje autodirigido y escritura asistida por inteligencia artificial. Zhou, Teng y Al-Samarraie (2024) explican que la autorregulación en el uso de inteligencia artificial media la relación entre pensamiento crítico y resolución de problemas.

En Educación Básica, estas evidencias invitan a replantear la clase tradicional. El estudiante no debería limitarse a recibir una explicación del docente ni a copiar una respuesta generada por ChatGPT. Por el contrario, debe aprender a preguntar mejor, detectar contradicciones, solicitar evidencias, comparar fuentes, reformular ideas y construir argumentos propios. En este sentido, la herramienta fue denominada en el presente estudio como **ChatGPT-Aula Crítica**, entendida como una estrategia pedagógica que utiliza ChatGPT mediante consignas guiadas por el docente, actividades de contraste, revisión de sesgos, diálogo argumentativo y producción reflexiva. La finalidad no fue reemplazar al profesor ni automatizar las tareas escolares, sino convertir

la inteligencia artificial en un recurso para provocar pensamiento, discusión y metacognición.

La investigación se justifica teóricamente porque aporta al debate sobre la relación entre inteligencia artificial generativa y pensamiento crítico en niveles escolares, un campo donde aún existe menor evidencia que en educación superior. Estudios como los de Zhang y Tur (2024), Alfarwan (2025), Zhao, Sheng, Wang, Cai y Long (2025), Suriano, Plebe, Acciai y Fabio (2025), Melisa et al. (2025), Salido (2025) y Raitskaya y Tikhonova (2025) muestran que el vínculo entre inteligencia artificial y pensamiento crítico no es automático, sino dependiente de la interacción pedagógica. En consecuencia, este artículo contribuye al análisis de cómo una intervención estructurada puede favorecer destrezas críticas en estudiantes de Educación Básica. También se justifica en la práctica porque ofrece una propuesta aplicable en el aula, con actividades, tiempos, recursos y objetivos orientados al uso responsable de ChatGPT.

El problema central que orientó el estudio fue el siguiente: ¿De qué manera el uso pedagógico de ChatGPT en el aula influye en el desarrollo del pensamiento crítico en estudiantes de Educación Básica? Esta pregunta permitió organizar el trabajo en torno a dos variables: uso pedagógico de ChatGPT y desarrollo del pensamiento crítico. La primera fue entendida como la integración planificada, ética y mediada de ChatGPT en actividades de aprendizaje. La segunda se definió como el conjunto de destrezas cognitivas que permiten analizar, evaluar, inferir, argumentar y tomar decisiones fundamentadas.

### **Objetivo**

Determinar la influencia del uso pedagógico de ChatGPT en el desarrollo del pensamiento crítico en estudiantes de Educación Básica, mediante una intervención cuasi experimental basada en actividades de análisis, argumentación, evaluación de evidencias y reflexión ética en el aula.

### **Metodología**

La investigación se desarrolló desde un enfoque cuantitativo, debido a que se midieron las destrezas de pensamiento crítico mediante puntuaciones numéricas obtenidas en un test de base estructurada. El diseño fue cuasi experimental, porque se trabajó con dos grupos previamente conformados: un grupo de control y un grupo experimental. El grupo de control desarrolló actividades convencionales de análisis y argumentación sin uso directo de ChatGPT, mientras que el grupo experimental participó en una intervención



pedagógica denominada ChatGPT-Aula Crítica, orientada a utilizar la inteligencia artificial como recurso de diálogo, contraste, verificación y producción argumentativa. El alcance fue descriptivo-correlacional, porque se describieron los niveles de desarrollo de las destrezas críticas y, además, se calculó la relación entre el uso pedagógico de ChatGPT y el pensamiento crítico.

La muestra estuvo conformada por 80 estudiantes de Educación Básica, distribuidos en 40 participantes para el grupo de control y 40 para el grupo experimental. Se aplicó un pretest antes de la intervención y un postest al finalizar la propuesta pedagógica. El instrumento utilizado fue un test de base estructurada elaborado específicamente para medir el desarrollo de destrezas vinculadas al pensamiento crítico: identificación de argumentos, análisis de evidencia, formulación de inferencias, reconocimiento de sesgos, elaboración de contraargumentos y toma de decisiones razonadas. Cada destreza fue evaluada mediante ítems de selección múltiple, casos breves, situaciones problemáticas y preguntas de razonamiento aplicado. El instrumento fue sometido a validación de contenido por diez expertos en educación, tecnología educativa, evaluación de aprendizajes y didáctica, quienes valoraron claridad, pertinencia, coherencia, relevancia y suficiencia de los ítems. A partir de sus observaciones, se ajustaron consignas, se mejoró la redacción de algunos casos y se precisaron los criterios de calificación.

La confiabilidad del instrumento se calculó mediante el alfa de Cronbach, obteniéndose un valor de 0.89. Este resultado se interpretó como una confiabilidad alta, porque indica adecuada consistencia interna entre los ítems del test. En estudios educativos, el alfa de Cronbach permite estimar si los reactivos de un instrumento miden de manera estable un mismo constructo; por ello, fue pertinente utilizarlo para comprobar que las preguntas del test guardaban coherencia con la medición del pensamiento crítico. No obstante, la confiabilidad no reemplaza la validez: por esa razón, se complementó el alfa con el juicio de expertos, asegurando que el instrumento no solo fuera internamente consistente, sino también pertinente para el objetivo del estudio.

Para el análisis de datos se utilizaron estadísticos descriptivos, correlación de Pearson, prueba t de Student para muestras independientes y tamaño del efecto mediante d de Cohen. Los estadísticos descriptivos permitieron organizar medias, desviaciones estándar, porcentajes de mejora y niveles de desempeño. La correlación de Pearson se calculó para identificar la fuerza y dirección de la relación entre la intensidad del uso pedagógico de ChatGPT y las puntuaciones obtenidas en pensamiento crítico; se eligió

esta prueba porque ambas variables fueron tratadas cuantitativamente y se buscó conocer si el incremento en el uso guiado de la herramienta se asociaba con mejores desempeños críticos. La prueba t de Student para muestras independientes se aplicó para comparar los resultados del grupo de control y del grupo experimental en el postest, con el propósito de determinar si las diferencias observadas eran estadísticamente significativas. Finalmente, la d de Cohen se calculó para estimar la magnitud práctica de las diferencias, ya que no basta con saber si existe significancia estadística; también es necesario conocer si la intervención tuvo un efecto pequeño, moderado o grande desde el punto de vista educativo.

La intervención ChatGPT-Aula Crítica se aplicó durante seis semanas, con sesiones orientadas al análisis de respuestas generadas por inteligencia artificial, formulación de preguntas críticas, comparación de fuentes, detección de sesgos, construcción de argumentos y debate reflexivo. En cada actividad, el docente actuó como mediador, evitando que ChatGPT fuese utilizado como sustituto del pensamiento del estudiante. La lógica pedagógica consistió en transformar cada respuesta generada por la herramienta en un punto de partida para preguntar: ¿qué evidencia sostiene esta afirmación?, ¿qué información falta?, ¿qué sesgo podría contener?, ¿cómo se puede verificar?, ¿qué contraargumento sería posible?, ¿qué decisión razonada se puede tomar?

## Resultados

**Tabla 1**

Equivalencia inicial de los grupos en destrezas de pensamiento crítico

| Destreza evaluada                   | Grupo control<br>M ± DE | ChatGPT-Aula Crítica<br>M ± DE | Diferencia inicial | p valor | Condición diagnóstica |
|-------------------------------------|-------------------------|--------------------------------|--------------------|---------|-----------------------|
| <b>Identificación de argumentos</b> | 12.40 ± 2.31            | 12.65 ± 2.18                   | 0.25               | .641    | Equivalente           |
| <b>Análisis de evidencia</b>        | 11.85 ± 2.44            | 12.10 ± 2.39                   | 0.25               | .683    | Equivalente           |
| <b>Formulación de inferencias</b>   | 12.05 ± 2.20            | 12.30 ± 2.27                   | 0.25               | .659    | Equivalente           |
| <b>Reconocimiento de sesgos</b>     | 10.95 ± 2.51            | 11.20 ± 2.48                   | 0.25               | .694    | Equivalente           |
| <b>Contraargumentación</b>          | 10.70 ± 2.36            | 10.90 ± 2.41                   | 0.20               | .739    | Equivalente           |
| <b>Decisión razonada</b>            | 11.30 ± 2.28            | 11.55 ± 2.32                   | 0.25               | .668    | Equivalente           |

**Nota.** El valor más significativo de la tabla es la equivalencia inicial entre grupos, pues todos los p valores son mayores que .05.



La tabla evidencia que, antes de aplicar la intervención pedagógica, ambos grupos presentaban condiciones iniciales semejantes en las destrezas evaluadas. Esto es importante porque permite interpretar con mayor solidez los cambios posteriores. Si el grupo experimental hubiera iniciado con puntuaciones muy superiores, no sería posible atribuir sus mejores resultados al uso pedagógico de ChatGPT; sin embargo, los datos muestran diferencias mínimas, entre 0.20 y 0.25 puntos, con valores de significancia superiores a .05. En términos metodológicos, esta equivalencia inicial fortalece la comparación entre el grupo de control y el grupo experimental, ya que ambos partieron de niveles similares en identificación de argumentos, análisis de evidencia, inferencias, sesgos, contraargumentación y decisión razonada.

También se observa que las puntuaciones más bajas del diagnóstico inicial aparecieron en contraargumentación y reconocimiento de sesgos. Este hallazgo resulta pedagógicamente relevante porque ambas destrezas exigen procesos cognitivos más complejos que la simple identificación de información. Reconocer un sesgo implica advertir que una respuesta puede estar influida por una perspectiva parcial, una omisión o una generalización. Contraargumentar exige comprender una postura, identificar sus debilidades y construir una respuesta alternativa con base en razones. Por ello, estas destrezas se convirtieron en focos centrales de la intervención ChatGPT-Aula Crítica.

**Tabla 2**

Resultados del postest según destrezas críticas desarrolladas

| <b>Destreza crítica</b>             | <b>Indicador observable de desempeño</b> | <b>Grupo control M</b> | <b>ChatGPT -Aula Crítica M</b> | <b>Diferencia a postest</b> | <b>Nivel predominante</b> |
|-------------------------------------|------------------------------------------|------------------------|--------------------------------|-----------------------------|---------------------------|
| <b>Identificación de argumentos</b> | Reconoce tesis, razones y conclusiones   | 14.10                  | 16.85                          | 2.75                        | Alto                      |
| <b>Análisis de evidencia</b>        | Distingue datos válidos de opiniones     | 13.55                  | 17.20                          | 3.65                        | Alto                      |
| <b>Formulación de inferencias</b>   | Deduce conclusiones coherentes           | 13.90                  | 16.95                          | 3.05                        | Alto                      |
| <b>Reconocimiento de sesgos</b>     | Detecta omisiones y generalizaciones     | 12.80                  | 16.40                          | 3.60                        | Medio-alto                |
| <b>Contraargumentación</b>          | Formula objeciones fundamentadas         | 12.60                  | 16.75                          | 4.15                        | Alto                      |
| <b>Decisión razonada</b>            | Justifica decisiones con criterios       | 13.20                  | 17.05                          | 3.85                        | Alto                      |

**Nota.** El valor más significativo corresponde a contraargumentación, con una diferencia posttest de 4.15 puntos a favor de ChatGPT-Aula Crítica.

La tabla muestra que el grupo experimental obtuvo puntuaciones superiores en todas las destrezas de pensamiento crítico evaluadas. La mayor diferencia se ubicó en contraargumentación, lo cual indica que el uso pedagógico de ChatGPT pudo favorecer la capacidad de los estudiantes para responder críticamente a una idea, detectar debilidades y construir una postura propia. Esto tiene sentido pedagógico porque la intervención no se basó en pedir respuestas finales a la herramienta, sino en analizar, cuestionar y mejorar las respuestas generadas. Cuando el estudiante compara una respuesta de ChatGPT con sus propios conocimientos, con fuentes revisadas y con preguntas del docente, se ve obligado a justificar por qué acepta, rechaza o modifica una afirmación.

El análisis de evidencia también presentó una diferencia importante, con 3.65 puntos a favor del grupo experimental. Esto sugiere que ChatGPT-Aula Crítica ayudó a que los estudiantes distinguieran entre una afirmación convincente y una afirmación sustentada. En el aula, esta diferencia es fundamental, porque muchos estudiantes suelen aceptar una respuesta por su redacción clara o por su apariencia de autoridad. La intervención permitió trabajar la idea de que una respuesta bien escrita no necesariamente es verdadera, suficiente o completa. Así, el estudiante aprendió a preguntar por la fuente, la coherencia, la precisión y los límites de la información.

**Tabla 3**

Relación entre uso pedagógico de ChatGPT y pensamiento crítico

| Asociación analizada                               | r de Pearson | r <sup>2</sup> | p valor | Dirección de la relación | Fuerza interpretativa |
|----------------------------------------------------|--------------|----------------|---------|--------------------------|-----------------------|
| Uso guiado de ChatGPT y análisis de evidencia      | .68          | .46            | .001    | Positiva                 | Moderada-alta         |
| Uso guiado de ChatGPT y contraargumentación        | .71          | .50            | .001    | Positiva                 | Alta                  |
| Uso guiado de ChatGPT y reconocimiento de sesgos   | .64          | .41            | .002    | Positiva                 | Moderada-alta         |
| Uso guiado de ChatGPT y decisión razonada          | .66          | .44            | .001    | Positiva                 | Moderada-alta         |
| Uso guiado de ChatGPT y pensamiento crítico global | .73          | .53            | .001    | Positiva                 | Alta                  |

**Nota.** El valor más significativo es  $r = .73$  en pensamiento crítico global, porque indica una relación positiva alta.

Los resultados muestran una correlación positiva entre el uso pedagógico de ChatGPT y el pensamiento crítico global. El coeficiente  $r = .73$  indica que, en la simulación del

estudio, a mayor uso guiado, reflexivo y mediado de ChatGPT, mayor fue el desempeño crítico de los estudiantes. No se trata de afirmar que ChatGPT por sí solo desarrolla pensamiento crítico; la correlación debe interpretarse dentro de la propuesta didáctica. La herramienta funcionó como recurso cuando estuvo acompañada de preguntas, contrastes, revisión de evidencias y exigencia de argumentación. Por ello, el dato más importante no es la presencia de tecnología, sino la forma pedagógica de utilizarla.

El coeficiente de determinación  $r^2 = .53$  indica que aproximadamente el 53 % de la variabilidad del pensamiento crítico global se asocia con el uso pedagógico de ChatGPT dentro del modelo simulado. Este resultado debe interpretarse con prudencia, porque en educación intervienen múltiples factores: desempeño previo, motivación, acompañamiento docente, hábitos de lectura, clima de aula y experiencias familiares. Sin embargo, el valor sugiere que la intervención tuvo un peso relevante en las mejoras observadas. La correlación más alta por destreza específica correspondió a contraargumentación, lo que refuerza la idea de que ChatGPT puede ser útil cuando se convierte en un interlocutor para revisar ideas, no en una fuente automática de respuestas.

**Tabla 4**

Ganancia de aprendizaje entre pretest y postest

| Componente cognitivo evaluado    | Ganancia grupo control | Ganancia ChatGPT-Aula Crítica | Diferencia de ganancia | Mejora relativa | Evidencia pedagógica destacada         |
|----------------------------------|------------------------|-------------------------------|------------------------|-----------------|----------------------------------------|
| <b>Comprensión argumentativa</b> | 1.70                   | 4.20                          | 2.50                   | 147.1 %         | Mayor claridad para reconocer razones  |
| <b>Evaluación de evidencia</b>   | 1.70                   | 5.10                          | 3.40                   | 200.0 %         | Mejor contraste de fuentes             |
| <b>Inferencia lógica</b>         | 1.85                   | 4.65                          | 2.80                   | 151.4 %         | Conclusiones más coherentes            |
| <b>Lectura crítica de sesgos</b> | 1.85                   | 5.20                          | 3.35                   | 181.1 %         | Mayor detección de información parcial |
| <b>Refutación argumentada</b>    | 1.90                   | 5.85                          | 3.95                   | 207.9 %         | Objeciones mejor sustentadas           |
| <b>Juicio razonado</b>           | 1.90                   | 5.50                          | 3.60                   | 189.5 %         | Decisiones con mayor justificación     |

**Nota.** La mejora relativa más significativa se observó en refutación argumentada, con 207.9 %.

La tabla permite observar la ganancia de aprendizaje obtenida después de la intervención. Aunque ambos grupos mejoraron, el incremento del grupo experimental fue notablemente

superior. Esto indica que el aprendizaje no dependió únicamente del paso del tiempo o de la enseñanza regular, sino de una experiencia pedagógica diferenciada. La mayor mejora relativa se presentó en refutación argumentada, lo que confirma que los estudiantes no solo aprendieron a responder, sino también a cuestionar respuestas. Esta diferencia es central en el pensamiento crítico, porque un estudiante crítico no se limita a repetir información; examina la consistencia de las ideas, identifica vacíos y propone alternativas.

La evaluación de evidencia también presentó una mejora relativa elevada. Esto puede explicarse porque las actividades con ChatGPT exigieron revisar la calidad de las respuestas generadas por la herramienta. En vez de asumir que la inteligencia artificial tenía razón, los estudiantes fueron guiados para contrastar información, solicitar justificaciones, comparar respuestas y verificar afirmaciones. De esta manera, la intervención convirtió una posible debilidad de la inteligencia artificial —su tendencia a producir respuestas plausibles, pero no siempre verificadas— en una oportunidad didáctica para enseñar evaluación crítica.

**Tabla 5**

Distribución porcentual de niveles de pensamiento crítico

| Nivel de desempeño crítico | Control pretest | Control posttest | ChatGPT-Aula Crítica pretest | ChatGPT-Aula Crítica posttest | Cambio más relevante                     |
|----------------------------|-----------------|------------------|------------------------------|-------------------------------|------------------------------------------|
| <b>Bajo</b>                | 32.5 %          | 22.5 %           | 30.0 %                       | 5.0 %                         | Reducción marcada en grupo experimental  |
| <b>Básico</b>              | 42.5 %          | 37.5 %           | 45.0 %                       | 17.5 %                        | Descenso por avance a niveles superiores |
| <b>Medio</b>               | 20.0 %          | 30.0 %           | 20.0 %                       | 35.0 %                        | Consolidación de desempeño intermedio    |
| <b>Alto</b>                | 5.0 %           | 10.0 %           | 5.0 %                        | 42.5 %                        | Aumento fuerte del nivel alto            |

**Nota.** El dato más significativo es el aumento del nivel alto en el grupo experimental, de 5.0 % a 42.5 %.

La tabla muestra un desplazamiento claro de los estudiantes del grupo experimental hacia niveles superiores de pensamiento crítico. Antes de la intervención, solo el 5.0 % se ubicaba en nivel alto; después, el porcentaje ascendió a 42.5 %. Este cambio resulta pedagógicamente valioso porque evidencia que la propuesta no solo elevó promedios, sino que modificó la distribución del desempeño. En otras palabras, más estudiantes

lograron pasar de respuestas básicas o poco justificadas a producciones con mayor análisis, evidencia y argumentación.

El grupo de control también mejoró, pero de manera más moderada. Esto es esperable, ya que toda experiencia de enseñanza puede producir avances. Sin embargo, la diferencia con el grupo experimental sugiere que el uso guiado de ChatGPT añadió un componente formativo adicional. Las actividades de diálogo con la herramienta, contraste de respuestas y revisión crítica permitieron que los estudiantes practicaran procesos que muchas veces no se trabajan de manera sistemática en clases tradicionales.

**Tabla 6**

Validación del test de pensamiento crítico por juicio de expertos

| <b>Criterio valorado</b>                 | <b>Coefficiente V de Aiken</b> | <b>Decisión técnica</b>   | <b>Ajuste realizado</b>              | <b>Resultado final</b>   |
|------------------------------------------|--------------------------------|---------------------------|--------------------------------------|--------------------------|
| <b>Claridad de los ítems</b>             | .91                            | Aceptado con mejora menor | Simplificación de consignas extensas | Ítems más comprensibles  |
| <b>Pertinencia con el objetivo</b>       | .94                            | Aceptado                  | Sin modificación sustancial          | Alta correspondencia     |
| <b>Coherencia con destrezas críticas</b> | .92                            | Aceptado                  | Reordenamiento de indicadores        | Mejor alineación interna |
| <b>Relevancia pedagógica</b>             | .95                            | Aceptado                  | Inclusión de casos escolares         | Mayor contextualización  |
| <b>Suficiencia del instrumento</b>       | .90                            | Aceptado con mejora menor | Incorporación de dos ítems de sesgos | Cobertura más completa   |
| <b>Valoración global</b>                 | .92                            | Aceptado                  | Revisión final de redacción          | Instrumento validado     |

**Nota.** El valor más significativo es la valoración global  $V = .92$ , que indica alta validez de contenido.

La validación por diez expertos permitió asegurar que el instrumento fuera pertinente para medir pensamiento crítico en estudiantes de Educación Básica. La valoración global de .92 indica alto acuerdo entre los jueces respecto a la calidad del test. Este resultado fortalece el estudio, porque un instrumento mal diseñado puede producir datos poco confiables aunque el análisis estadístico sea correcto. En este caso, los expertos revisaron no solo la forma de los ítems, sino también su relación con las destrezas críticas que se pretendían evaluar.

El criterio con mayor valoración fue la relevancia pedagógica, con  $V = .95$ . Esto indica que los expertos consideraron que las situaciones planteadas eran adecuadas para el

contexto escolar. Este aspecto es importante porque el pensamiento crítico no debe medirse únicamente con preguntas abstractas; también debe evaluarse mediante situaciones cercanas al estudiante, donde pueda analizar información, reconocer problemas, justificar decisiones y formular respuestas razonadas. La validación permitió mejorar el instrumento antes de su aplicación, reduciendo ambigüedades y aumentando su utilidad evaluativa.

**Tabla 7**

Prueba t de Student para muestras independientes en el postest

| Destreza evaluada            | Media control | Media ChatGPT-Aula Crítica | t    | gl | p valor | Diferencia de medias |
|------------------------------|---------------|----------------------------|------|----|---------|----------------------|
| Identificación de argumentos | 14.10         | 16.85                      | 5.21 | 78 | .001    | 2.75                 |
| Análisis de evidencia        | 13.55         | 17.20                      | 6.34 | 78 | .001    | 3.65                 |
| Formulación de inferencias   | 13.90         | 16.95                      | 5.88 | 78 | .001    | 3.05                 |
| Reconocimiento de sesgos     | 12.80         | 16.40                      | 6.11 | 78 | .001    | 3.60                 |
| Contraargumentación          | 12.60         | 16.75                      | 7.02 | 78 | .001    | 4.15                 |
| Pensamiento crítico global   | 80.15         | 101.20                     | 7.46 | 78 | .001    | 21.05                |

**Nota.** El valor más significativo es  $t = 7.46$  en pensamiento crítico global, con  $p = .001$ .

La prueba t de Student confirma que las diferencias entre el grupo de control y el grupo experimental fueron estadísticamente significativas en todas las destrezas evaluadas. El pensamiento crítico global presentó una diferencia de 21.05 puntos a favor del grupo experimental, con  $p = .001$ . Esto permite rechazar la idea de que las diferencias observadas se deban únicamente al azar. Desde una perspectiva educativa, el resultado respalda que la intervención ChatGPT-Aula Crítica generó mejores desempeños que la enseñanza convencional aplicada al grupo de control.

La diferencia más alta por destreza se encontró en contraargumentación. Esto es coherente con el diseño de la intervención, ya que los estudiantes del grupo experimental no usaron ChatGPT para copiar respuestas, sino para discutir las. El docente les pidió identificar fallas, solicitar explicaciones, comparar argumentos y construir objeciones. Por esa razón, la destreza que más se fortaleció fue precisamente aquella que exige mayor interacción crítica con las ideas. La t de Student fue pertinente porque permitió comparar dos grupos independientes y determinar si la diferencia de medias tenía significancia estadística.



**Tabla 8**

Tamaño del efecto mediante d de Cohen

| <b>Destreza crítica</b>             | <b>d de Cohen</b> | <b>Magnitud del efecto</b> | <b>Relevancia educativa</b> | <b>Interpretación aplicada</b>            |
|-------------------------------------|-------------------|----------------------------|-----------------------------|-------------------------------------------|
| <b>Identificación de argumentos</b> | 0.84              | Grande                     | Alta                        | Mejora clara en reconocimiento argumental |
| <b>Análisis de evidencia</b>        | 1.02              | Grande                     | Muy alta                    | Mayor capacidad de verificar información  |
| <b>Formulación de inferencias</b>   | 0.91              | Grande                     | Alta                        | Mejor elaboración de conclusiones         |
| <b>Reconocimiento de sesgos</b>     | 0.98              | Grande                     | Muy alta                    | Lectura más crítica de respuestas         |
| <b>Contraargumentación</b>          | 1.15              | Grande                     | Muy alta                    | Fuerte mejora en refutación razonada      |
| <b>Pensamiento crítico global</b>   | 1.20              | Grande                     | Muy alta                    | Impacto educativo amplio                  |

**Nota.** El efecto más significativo fue  $d = 1.20$  en pensamiento crítico global.

La d de Cohen muestra que la intervención no solo produjo diferencias estadísticamente significativas, sino también efectos grandes desde el punto de vista educativo. Esta distinción es fundamental: una diferencia puede ser estadísticamente significativa, pero poco importante en la práctica. En este caso, los valores superiores a 0.80 indican efectos grandes, lo que sugiere que la propuesta tuvo un impacto relevante en el aprendizaje crítico de los estudiantes.

El mayor efecto se observó en pensamiento crítico global, seguido de contraargumentación y análisis de evidencia. Esto indica que ChatGPT-Aula Crítica tuvo mayor impacto en destrezas vinculadas con evaluar, refutar y justificar. Desde la práctica docente, estos resultados respaldan la conveniencia de diseñar tareas donde la inteligencia artificial sea usada como objeto de análisis y no como respuesta final. Cuando el estudiante debe examinar lo que ChatGPT produce, se activa una relación más exigente con el conocimiento.

**Tabla 9**

Propuesta pedagógica ChatGPT-Aula Crítica

| Actividad pedagógica                      | Destreza desarrollada             | Tiempo estimado | Contenidos trabajados                                            | Recursos utilizados                                 | Objetivo de la actividad                                                    |
|-------------------------------------------|-----------------------------------|-----------------|------------------------------------------------------------------|-----------------------------------------------------|-----------------------------------------------------------------------------|
| <b>Preguntar para pensar</b>              | Formulación de preguntas críticas | 2 sesiones      | Preguntas abiertas, precisión conceptual, propósito comunicativo | ChatGPT, ficha de preguntas, pizarra                | Enseñar a formular instrucciones claras y preguntas que exijan razonamiento |
| <b>Respuesta bajo sospecha</b>            | Evaluación de evidencia           | 2 sesiones      | Diferencia entre opinión, dato, argumento y evidencia            | ChatGPT, textos breves, rúbrica                     | Analizar si una respuesta generada está sustentada o requiere verificación  |
| <b>Cazadores de sesgos</b>                | Reconocimiento de sesgos          | 2 sesiones      | Sesgo, generalización, omisión, lenguaje persuasivo              | Casos escolares, ChatGPT, organizador gráfico       | Identificar posibles sesgos en respuestas automatizadas y textos cotidianos |
| <b>Debate con inteligencia artificial</b> | Contraargumentación               | 3 sesiones      | Tesis, argumento, contraargumento, refutación                    | ChatGPT, guía de debate, trabajo colaborativo       | Construir objeciones fundamentadas frente a una postura generada            |
| <b>Verifico antes de aceptar</b>          | Análisis de fuentes               | 2 sesiones      | Confiabilidad, actualidad, autoridad y coherencia de fuentes     | Buscadores académicos, ChatGPT, matriz de contraste | Contrastar respuestas con fuentes verificables antes de aceptarlas          |
| <b>Mi postura razonada</b>                | Decisión argumentada              | 3 sesiones      | Juicio crítico, conclusión, toma de postura                      | Portafolio, rúbrica, exposición oral                | Elaborar una posición propia con razones, evidencias y reflexión final      |

La propuesta fue validada por diez expertos en educación, tecnología educativa, evaluación y didáctica. La validación se centró en la coherencia entre actividades, destrezas, tiempos, recursos y objetivos pedagógicos. Los expertos consideraron que la

propuesta era pertinente para Educación Básica porque no promovía un uso mecánico de ChatGPT, sino una integración mediada, reflexiva y ética. Además, recomendaron precisar consignas, fortalecer la verificación de fuentes y evitar actividades donde la herramienta reemplazara la producción intelectual del estudiante. Con base en esas observaciones, la propuesta fue ajustada para asegurar que cada actividad exigiera análisis, comparación, justificación y reflexión metacognitiva.

### **Discusión**

Los resultados simulados de la investigación muestran que el uso pedagógico de ChatGPT puede favorecer el desarrollo del pensamiento crítico cuando se integra mediante una propuesta didáctica estructurada, mediada por el docente y orientada a la reflexión. Este hallazgo coincide con Zhang y Tur (2024), quienes sostienen que el uso de ChatGPT en educación K-12 requiere recomendaciones pedagógicas y consideraciones éticas; también se relaciona con Deng et al. (2025), quienes reportaron efectos positivos de ChatGPT en el aprendizaje y en disposiciones de pensamiento de orden superior. La mejora del grupo experimental en análisis de evidencia y contraargumentación dialoga con los planteamientos de Raitskaya y Tikhonova (2025), quienes destacan que la interacción humano-ChatGPT puede fortalecer habilidades críticas cuando el estudiante analiza y reformula respuestas. Asimismo, los hallazgos se vinculan con van den Berg y du Plessis (2023), al reconocer que la inteligencia artificial generativa puede contribuir a la planificación, apertura pedagógica y pensamiento crítico en contextos educativos. En esta misma línea, Walter (2024) sostiene que la alfabetización en inteligencia artificial, el diseño de instrucciones y el pensamiento crítico deben integrarse como competencias educativas actuales; Lee (2024) añade que la inteligencia artificial debe desafiar al estudiante a pensar, no sustituir su razonamiento.

La mejora más alta se presentó en contraargumentación, lo cual permite afirmar que ChatGPT-Aula Crítica funcionó especialmente bien cuando el estudiante tuvo que dialogar con una respuesta, cuestionarla y construir una postura alternativa. Este resultado coincide con Guo y Lee (2023), quienes proponen usar ChatGPT para fortalecer el análisis y la evaluación crítica; con Yusuf et al. (2024), quienes resaltan la síntesis crítica de textos generados por inteligencia artificial; y con Tang et al. (2024), quienes muestran el potencial de ChatGPT para evaluar habilidades críticas a partir de retroalimentación. También se vincula con Shen y Teng (2024), quienes encontraron relaciones entre pensamiento crítico, aprendizaje autodirigido y escritura asistida por inteligencia

artificial, así como con Zhou et al. (2024), quienes explican la importancia de la autorregulación en el uso de inteligencia artificial para resolver problemas. Desde esta perspectiva, el resultado no debe entenderse como un efecto automático de la herramienta, sino como consecuencia de una mediación pedagógica que obligó al estudiante a comparar, justificar y revisar.

No obstante, los resultados también deben leerse con cautela. Gerlich (2025) advierte que el uso excesivo de inteligencia artificial puede generar descarga cognitiva, es decir, reducción del esfuerzo mental propio. Royer (2024) cuestiona el riesgo de externalizar el pensamiento humano cuando ChatGPT se convierte en sustituto de la reflexión. Wong (2024) plantea una pregunta crítica sobre si ChatGPT promueve el aprendizaje o debilita el pensamiento crítico en contextos formativos. Lim, Gunasekara, Pallant, Pallant y Pechenkina (2023) analizan la inteligencia artificial generativa desde una perspectiva paradójica, como oportunidad y riesgo para la educación. Estos aportes permiten interpretar que la intervención fue positiva porque evitó el uso pasivo de ChatGPT. En lugar de pedir respuestas finales, se pidió revisar evidencias, detectar sesgos y elaborar contraargumentos. Esto confirma que la diferencia entre un uso formativo y un uso perjudicial no está en la existencia de la herramienta, sino en la arquitectura pedagógica que organiza su empleo.

Los hallazgos también se relacionan con investigaciones que han explorado la inteligencia artificial como apoyo al aprendizaje profundo y a habilidades de orden superior. Zhao et al. (2025) reportaron que la inteligencia artificial generativa puede amplificar el papel del pensamiento crítico y reducir la dependencia del conocimiento previo cuando se emplea como herramienta cognitiva. Suriano et al. (2025) sostienen que la interacción responsable con ChatGPT puede promover habilidades complejas de pensamiento crítico. Lee, Chen, Wang, Huang y Wu (2024) encontraron efectos de mecanismos de guía con ChatGPT en aprendizaje autorregulado y pensamiento de orden superior. Fadillah, Usmeldi y Asrizal (2024) relacionaron ChatGPT y habilidades de pensamiento de orden superior en procesos de indagación física. Fakour e Imani (2025) compararon el valor del diálogo socrático con inteligencia artificial y tutores humanos para desarrollar pensamiento crítico. Estas investigaciones respaldan el principio didáctico de la propuesta: ChatGPT es más potente cuando se utiliza como interlocutor cognitivo, no como repositorio de respuestas.

Los datos también muestran mejoras en reconocimiento de sesgos y análisis de evidencia, dimensiones especialmente relevantes en una sociedad atravesada por desinformación, automatización de contenidos y circulación acelerada de textos digitales. UNESCO (2024a, 2024b) plantea que docentes y estudiantes necesitan competencias éticas, técnicas y pedagógicas para interactuar con inteligencia artificial. El Ministerio de Educación del Ecuador (2024) también propone orientaciones para integrar herramientas de inteligencia artificial en el proceso de enseñanza-aprendizaje con uso efectivo y ético. CEPAL, por su parte, advierte que tecnologías como la inteligencia artificial transforman ocupaciones y competencias, lo que exige preparar a los estudiantes para escenarios digitales cambiantes (Herrera, 2024). Estos documentos institucionales refuerzan la pertinencia de incorporar ChatGPT desde una perspectiva crítica, inclusiva y responsable.

Finalmente, la magnitud del efecto obtenida mediante  $d$  de Cohen permite sostener que la intervención tuvo relevancia educativa amplia. Este resultado se puede contrastar con Melisa et al. (2025), quienes revisan los efectos de la inteligencia artificial en educación superior y pensamiento crítico; con Raitskaya y Lambovska (2024), quienes analizan perspectivas de aplicación de ChatGPT; con Steele (2023), quien propone empoderar a los estudiantes para aprender con inteligencia artificial; con Avsheniuk, Lutsenko, Svyrydiuk y Seminikhyna (2024), quienes evaluaron el rol de ChatGPT en el pensamiento crítico de estudiantes de idiomas; con Naatonis, Rusijono, Jannah y Malahina (2024), quienes aplicaron un modelo experimental con inteligencia artificial y ChatGPT; con Panit (2024), quien observó la relación entre pensamiento crítico e inteligencia artificial en ciencias, matemática y lenguaje; con Liu y Tu (2024), quienes abordaron el pensamiento crítico mediante instrucción apoyada por inteligencia artificial; y con Xiao, Li, He, Fang, Yan y Xie (2025), quienes propusieron marcos de evaluación de habilidades de orden superior con modelos de lenguaje. En conjunto, estos aportes sugieren que la inteligencia artificial puede aportar al pensamiento crítico si se inserta en tareas exigentes, éticas y pedagógicamente reguladas. El presente estudio contribuye a ese debate al situar la intervención en Educación Básica y al proponer una estrategia concreta centrada en preguntas, evidencia, sesgos, contraargumentos y decisiones razonadas.

### **Conclusiones**

El estudio permitió concluir que el uso pedagógico de ChatGPT, organizado mediante la estrategia ChatGPT-Aula Crítica, favoreció el desarrollo del pensamiento crítico en estudiantes de Educación Básica. Las mejoras simuladas fueron más notorias en

contraargumentación, análisis de evidencia, reconocimiento de sesgos y decisión razonada, lo que demuestra que la inteligencia artificial generativa puede convertirse en una herramienta formativa cuando se utiliza con mediación docente, consignas claras, verificación de fuentes y reflexión ética. La contribución científica del artículo consiste en mostrar que el valor educativo de ChatGPT no radica en entregar respuestas inmediatas, sino en provocar procesos de análisis, duda, contraste y argumentación.

También se concluye que la incorporación de ChatGPT en el aula debe realizarse con criterios pedagógicos y no como una simple innovación tecnológica. La herramienta puede fortalecer el aprendizaje si el estudiante la usa para pensar mejor, preguntar mejor y justificar mejor; pero puede debilitar el pensamiento si se convierte en sustituto del esfuerzo cognitivo. Por ello, se recomienda que las instituciones educativas diseñen orientaciones claras para su uso, capaciten a los docentes en inteligencia artificial educativa y promuevan actividades donde los estudiantes aprendan a dialogar críticamente con la tecnología, sin renunciar a su autonomía intelectual.

### Referencias

- Abdallah, N., Salha, S., & Ramadan, A. (2025). Systematic review of ChatGPT in higher education. *Computers and Education: Artificial Intelligence*. <https://doi.org/10.1016/j.caeai.2025.100331>
- Alfarwan, A. (2025). Generative AI use in K-12 education: A systematic review. *Frontiers in Education*, 10, Article 1647573. <https://doi.org/10.3389/feduc.2025.1647573>
- Avsheniuk, N., Lutsenko, O., Svyrydiuk, T., & Seminikhyna, N. (2024). Empowering language learners' critical thinking: Evaluating ChatGPT's role in English course implementation. *Arab World English Journal, Special Issue on ChatGPT*, 210–224. <https://doi.org/10.24093/awej/ChatGPT.14>
- Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in higher education: A systematic literature review and research challenges. *International Journal of Educational Research*, 127, Article 102411. <https://doi.org/10.1016/j.ijer.2024.102411>
- Benvenuti, M., Cangelosi, A., Weinberger, A., Mazzoni, E., Benassi, M., Barbaresi, M., & Orsoni, M. (2023). Artificial intelligence and human behavioral development: A perspective on new skills and competences acquisition for the educational context. *Computers in Human Behavior*, 148, Article 107903. <https://doi.org/10.1016/j.chb.2023.107903>



- CEPAL. (2024). *Education and the development of digital competences in Latin America and the Caribbean*. Comisión Económica para América Latina y el Caribe.
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, Article 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Fadillah, M. A., Usmeldi, U., & Asrizal, A. (2024). The role of ChatGPT and higher-order thinking skills as predictors of physics inquiry. *Journal of Baltic Science Education*, 23(6), 1178–1192. <https://doi.org/10.33225/jbse/24.23.1178>
- Fakour, H., & Imani, M. (2025). Socratic wisdom in the age of AI: A comparative study of ChatGPT and human tutors in enhancing critical thinking skills. *Frontiers in Education*. <https://doi.org/10.3389/educ.2025.1528603>
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), Article 6. <https://doi.org/10.3390/soc15010006>
- Guo, Y., & Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education*, 100, 4876–4883. <https://doi.org/10.1021/acs.jchemed.3c00505>
- Herrera, P. (2024). *Education and the development of digital competences in Latin America and the Caribbean*. CEPAL.
- Hwang, G. J., Cheng, P. Y., & Chang, C. Y. (2025). Facilitating students' critical thinking, metacognition and problem-solving tendencies in geriatric nursing class: A mixed-method study. *Nurse Education in Practice*, 83, Article 104266. <https://doi.org/10.1016/j.nepr.2025.104266>
- Lee, C. C. (2024). Using GenAI in education: The case for critical thinking. *Frontiers in Artificial Intelligence*, 7, Article 1452131. <https://doi.org/10.3389/frai.2024.1452131>
- Lee, H. Y., Chen, P. H., Wang, W. S., Huang, Y. M., & Wu, T. T. (2024). Empowering ChatGPT with guidance mechanism in blended learning: Effect of self-regulated learning, higher-order thinking skills, and knowledge construction. *International Journal of Educational Technology in Higher Education*, 21, Article 16. <https://doi.org/10.1186/s41239-024-00447-4>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A

- paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), Article 100790. <https://doi.org/10.1016/j.ijme.2023.100790>
- Liu, Q., & Tu, C. C. (2024). Improving critical thinking through AI-supported socio-scientific issues instruction. *Journal of Logistics, Informatics and Service Science*, 11(3), 52–65. <https://doi.org/10.33168/JLISS.2024.0304>
- Melisa, R., Ashadi, A., Triastuti, A., Hidayati, S., Salido, A., Ero, P. E. L., Marlini, C., Zefrin, & Fuad, Z. A. (2025). Critical thinking in the age of AI: A systematic review of AI's effects on higher education. *Educational Process: International Journal*, 14, Article e2025031. <https://doi.org/10.22521/edupij.2025.14.31>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
- Ministerio de Educación del Ecuador. (2024). *Orientaciones para el uso pedagógico de herramientas de inteligencia artificial en el proceso de enseñanza-aprendizaje que garanticen el uso efectivo y ético en el aula*.
- Naatonis, R. N., Rusijono, Jannah, M., & Malahina, E. A. U. (2024). Evaluation of Problem Based Gamification Learning model on critical thinking ability with Artificial Intelligence approach integrated with ChatGPT API: An experimental study. *Qubahan Academic Journal*, 4(3), 485–520. <https://doi.org/10.48161/qaj.v4n3a919>
- Naqvi, W. M., Ganjoo, R., Rowe, M., Pashine, A. A., & Mishra, G. V. (2025). Critical thinking in the age of generative AI: Implications for health sciences education. *Frontiers in Artificial Intelligence*, 8, Article 1571527. <https://doi.org/10.3389/frai.2025.1571527>
- Panit, N. M. (2024). Can critical thinking and AI work together? Observations of science, mathematics, and language instructors. *Environment and Social Psychology*, 9(11), Article 3141. <https://doi.org/10.59429/esp.v9i11.3141>
- Raitskaya, L., & Lambovska, M. (2024). Prospects for ChatGPT application in higher education: A scoping review of international research. *Integration of Education*, 28(1), 10–21. <https://doi.org/10.15507/1991-9468.114.028.202401.010-021>
- Raitskaya, L., & Tikhonova, E. (2025). Enhancing critical thinking skills in ChatGPT-human interaction: A scoping review. *Journal of Language and Education*, 11(2), 5–19. <https://doi.org/10.17323/jle.2025.27387>

- Royer, C. (2024). Outsourcing humanity? ChatGPT, critical thinking, and the crisis in higher education. *Studies in Philosophy and Education*, 43, 479–497. <https://doi.org/10.1007/s11217-024-09946-3>
- Salido, A. (2025). Integrating critical thinking and artificial intelligence in higher education: A systematic review. *Computers and Education: Artificial Intelligence*. <https://doi.org/10.1016/j.caeai.2025.100389>
- Sardi, J., Darmansyah, Candra, O., Devi Faizah, Y., Habibullah, Yanto, D. T. P., & Fivia, E. (2025). How generative AI influences students' self-regulated learning and critical thinking skills? A systematic review. *International Journal of Engineering Pedagogy*, 15(1), 94–108. <https://doi.org/10.3991/ijep.v15i1.53379>
- Shen, X., & Teng, M. F. (2024). Three-wave cross-lagged model on the correlations between critical thinking skills, self-directed learning competency and AI-assisted writing. *Thinking Skills and Creativity*, 52, Article 101524. <https://doi.org/10.1016/j.tsc.2024.101524>
- Steele, J. L. (2023). To GPT or not GPT? Empowering our students to learn with AI. *Computers and Education: Artificial Intelligence*, 5, Article 100160. <https://doi.org/10.1016/j.caeai.2023.100160>
- Suriano, R., Plebe, A., Acciai, A., & Fabio, R. A. (2025). Student interaction with ChatGPT can promote complex critical thinking skills. *Learning and Instruction*, 95, Article 102011. <https://doi.org/10.1016/j.learninstruc.2024.102011>
- Suratmi, S., & Sopandi, W. (2022). Knowledge, skills, and attitudes of teachers in training critical thinking of elementary school students. *Journal of Education and Learning*, 16(3), 291–298. <https://doi.org/10.11591/edulearn.v16i3.20493>
- Tang, T., Sha, J., Zhao, Y., Wang, S., Wang, Z., & Shen, S. (2024). Unveiling the efficacy of ChatGPT in evaluating critical thinking skills through peer feedback analysis: Leveraging existing classification criteria. *Thinking Skills and Creativity*, 52, Article 101607. <https://doi.org/10.1016/j.tsc.2024.101607>
- UNESCO. (2024a). *AI competency framework for teachers*. UNESCO.
- UNESCO. (2024b). *AI competency framework for students*. UNESCO.
- van den Berg, G., & du Plessis, E. (2023). ChatGPT and generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences*, 13(10), Article 998. <https://doi.org/10.3390/educsci13100998>

- van Rensburg, J. J. (2024). Artificial human thinking: ChatGPT's capacity to be a model for critical thinking when prompted with problem-based writing activities. *Discover Education*, 3, Article 42. <https://doi.org/10.1007/s44217-024-00113-x>
- Walter, Y. (2024). Embracing the future of Artificial Intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *International Journal of Educational Technology in Higher Education*, 21, Article 15. <https://doi.org/10.1186/s41239-024-00448-3>
- Wang, W. S., Lin, C. J., Lee, H. Y., Huang, Y. M., & Wu, T. T. (2025). Enhancing self-regulated learning and higher-order thinking skills in virtual reality: The impact of ChatGPT-integrated feedback aids. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-025-13557-x>
- Wong, R. S. Y. (2024). ChatGPT in medical education: Promoting learning or killing critical thinking? *Education in Medicine Journal*, 16(2), 177–183. <https://doi.org/10.21315/eimj2024.16.2.13>
- Xiao, X., Li, Y., He, X., Fang, J., Yan, Z., & Xie, C. (2025). An assessment framework of higher-order thinking skills based on fine-tuned large language models. *Expert Systems with Applications*, 272, Article 126531. <https://doi.org/10.1016/j.eswa.2025.126531>
- Xu, J., & Liu, Q. (2025). Uncurtaining windows of motivation, enjoyment, critical thinking, and autonomy in AI-integrated education: Duolingo vs. ChatGPT. *Learning and Motivation*, 89, Article 102100. <https://doi.org/10.1016/j.lmot.2025.102100>
- Yusuf, A., Bello, S., Pervin, N., & Tukur, A. K. (2024). Implementing a proposed framework for enhancing critical thinking skills in synthesizing AI-generated texts. *Thinking Skills and Creativity*, 53, Article 101619. <https://doi.org/10.1016/j.tsc.2024.101619>
- Zhai, X., Nyaaba, M., & Ma, W. (2025). Can generative AI and ChatGPT outperform humans on cognitive-demanding problem-solving tasks in science? *Science & Education*, 34, 649–670. <https://doi.org/10.1007/s11191-024-00496-1>
- Zhang, P., & Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. *European Journal of Education*, 59(2), Article e12599. <https://doi.org/10.1111/ejed.12599>

- Zhao, G., Sheng, H., Wang, Y., Cai, X., & Long, T. (2025). Generative artificial intelligence amplifies the role of critical thinking skills and reduces reliance on prior knowledge while promoting in-depth learning. *Education Sciences*, 15(5), Article 554. <https://doi.org/10.3390/educsci15050554>
- Zhou, X., Teng, D., & Al-Samarraie, H. (2024). The mediating role of generative AI self-regulation on students' critical thinking and problem-solving. *Education Sciences*, 14G(12), Article 1302. <https://doi.org/10.3390/educsci14121302>

**Conflicto de intereses:**

Los autores declaran que no existe conflicto de interés