



**Aplicación de inteligencia artificial generativa como recurso didáctico en Filosofía
para fortalecer el pensamiento crítico y la reflexión ética en estudiantes de
Bachillerato**

**Application of Generative Artificial Intelligence as a Didactic Resource in
Philosophy to Strengthen Critical Thinking and Ethical Reflection in High School
Students**



Ángela del Pilar Tello Báez

angelap.tello@educacion.gob.ec



<https://orcid.org/0009-0006-8980-1665>

Ministerio de Educación – Educación para la Ciudadanía – Filosofía
Cotopaxi – Ecuador

Emerson Javier Laverde Albarracín

emerson.laverde@docentes.educacion.edu.ec



<https://orcid.org/0009-0009-2868-7047>

Ministerio de Educación – Educación para la Ciudadanía – Filosofía
Cotopaxi – Ecuador

Fecha de envío: 2026-05-02

Fecha de revisión: 2026-06-03

Fecha da aceptación: 2026-07-10

Resumen

El objetivo de la investigación fue determinar la incidencia de la inteligencia artificial generativa, integrada como recurso didáctico en la asignatura de Filosofía, en el fortalecimiento del pensamiento crítico y la reflexión ética de estudiantes de Bachillerato. La metodología correspondió a un estudio cuasi experimental, de enfoque correlacional descriptivo, con grupo de control y grupo experimental, desarrollado en cuatro centros educativos. Participaron 50 estudiantes distribuidos equitativamente; se aplicó un test de base estructurada antes y después de la intervención, validado por juicio de diez expertos y con confiabilidad Alfa de Cronbach de 0.86. Los datos fueron procesados mediante estadísticos descriptivos, correlación de Pearson, prueba t de Student para muestras independientes y tamaño del efecto d de Cohen. Los resultados evidenciaron mejoras superiores en el grupo experimental en argumentación filosófica, evaluación de fuentes, formulación de preguntas, deliberación moral y transferencia de criterios éticos a problemas contemporáneos. La diferencia posprueba entre grupos fue estadísticamente significativa y el tamaño del efecto se ubicó entre moderado y alto, lo cual sugiere que el uso guiado, crítico y ético de la inteligencia artificial generativa favorece aprendizajes filosóficos de mayor profundidad. Se concluye que la herramienta no sustituye el juicio docente ni la autonomía estudiantil, pero puede ampliar la práctica dialógica, la contrastación de perspectivas y la toma de decisiones razonadas cuando se acompaña de criterios pedagógicos, verificación de información y reflexión ética.

Palabras clave: inteligencia artificial generativa; Filosofía; pensamiento crítico; reflexión ética; Bachillerato.

Abstract

The objective of this research was to determine the impact of generative artificial intelligence, integrated as a didactic resource in Philosophy, on strengthening critical thinking and ethical reflection among high school students. The methodology followed a quasi-experimental study with a descriptive-correlational approach, using a control group and an experimental group, implemented in four educational centers. Fifty students participated and were evenly distributed. A structured test was applied before and after the intervention, validated by ten expert judges and reaching a Cronbach's alpha reliability coefficient of 0.86. Data were processed through descriptive statistics, Pearson correlation, Student's t-test for independent samples, and Cohen's d effect size. The

results showed greater improvement in the experimental group regarding philosophical argumentation, source evaluation, question formulation, moral deliberation, and the transfer of ethical criteria to contemporary problems. Post-test differences between groups were statistically significant, and the effect size ranged from moderate to high, suggesting that guided, critical, and ethical use of generative artificial intelligence promotes deeper philosophical learning. It is concluded that the tool does not replace teacher judgment or student autonomy, but it can expand dialogic practice, comparison of perspectives, and reasoned decision-making when accompanied by pedagogical criteria, information verification, and ethical reflection.

Keywords: generative artificial intelligence; Philosophy; critical thinking; ethical reflection; high school.

Introducción

La enseñanza de Filosofía en Bachillerato atraviesa un momento decisivo porque los estudiantes conviven con entornos digitales que producen respuestas rápidas, textos persuasivos y explicaciones aparentemente completas. En ese escenario, el área de Lenguaje y Comunicación no puede limitarse a la comprensión literal ni a la producción textual convencional; necesita formar lectores capaces de preguntar, contrastar, argumentar y decidir con criterios éticos. La inteligencia artificial generativa, entendida como un conjunto de sistemas capaces de producir lenguaje, ejemplos, síntesis, explicaciones y simulaciones conversacionales a partir de instrucciones humanas, abre posibilidades didácticas relevantes, pero también exige cautela, supervisión y alfabetización crítica. La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 2023) advierte que estas tecnologías deben incorporarse con criterios de protección, equidad, transparencia y finalidad pedagógica; en la misma línea, Chan (2023), Kasneci et al. (2023) y Tlili et al. (2023) señalan que su valor educativo depende menos de la herramienta en sí y más de las condiciones de uso, las tareas propuestas y la capacidad del docente para convertir la interacción con la máquina en una experiencia de pensamiento.

En América Latina, el debate sobre inteligencia artificial en educación se relaciona con brechas de conectividad, desigualdades de acceso y desafíos de ciudadanía digital. La Comisión Económica para América Latina y el Caribe (CEPAL) y la UNESCO han insistido en que la transformación digital debe orientarse hacia inclusión, desarrollo de

capacidades y fortalecimiento de sistemas educativos, no hacia una adopción instrumental sin reflexión social. En Ecuador, el Ministerio de Educación ha publicado orientaciones para el uso pedagógico de herramientas de inteligencia artificial en el proceso de enseñanza-aprendizaje, remarcando la necesidad de un uso efectivo, ético y pertinente en el aula. Estas orientaciones son especialmente relevantes para Bachillerato porque los estudiantes se encuentran en una etapa de consolidación de autonomía intelectual, construcción de criterios morales y participación discursiva en problemas sociales. Por ello, la Filosofía puede convertirse en un espacio privilegiado para enseñar a usar la inteligencia artificial generativa sin delegar el razonamiento, sino exigiendo preguntas mejores, evidencias más sólidas y argumentos más responsables.

La literatura reciente muestra que la inteligencia artificial generativa puede apoyar el aprendizaje cuando se usa para generar retroalimentación, comparar perspectivas, traducir conceptos complejos a ejemplos cercanos, formular contraargumentos y promover prácticas metacognitivas. Lo (2023) identificó en una revisión rápida que ChatGPT y herramientas similares ofrecen oportunidades para tutoría, apoyo a la escritura y personalización, aunque también presentan riesgos de dependencia, desinformación y debilitamiento de la autoría. Kohnke et al. (2023) destacan su potencial en enseñanza de lenguas, siempre que se desarrollen competencias digitales y criterios de uso ético. Barrett y Pack (2023) muestran que docentes y estudiantes perciben beneficios en los procesos de escritura, pero también tensiones sobre autenticidad, evaluación y responsabilidad. En consecuencia, el uso de inteligencia artificial generativa en Lenguaje y Comunicación debe ser diseñado como mediación para pensar mejor, no como atajo para producir tareas.

En términos didácticos, el recurso utilizado en la investigación se denominó Asistente Filosofía-IA, una estrategia pedagógica basada en prompts guiados, preguntas socráticas, verificación de fuentes y matrices de deliberación ética. La herramienta no fue presentada como autoridad, sino como interlocutor provisional. Cada respuesta generada debía ser examinada por el estudiante mediante criterios de claridad conceptual, coherencia lógica, evidencia, sesgo, pertinencia ética y relación con los contenidos de Filosofía. Esta lógica coincide con Lim y Makany (2024), quienes sostienen que los chatbots pueden apoyar pensamiento crítico si se despliegan con propósito, andamiaje y evaluación del proceso, y con Avsheniuk et al. (2024), quienes resaltan que el valor de ChatGPT en cursos de lenguaje aumenta cuando el estudiante evalúa, reformula y argumenta a partir de la respuesta recibida.

Por lo expuesto, el problema de investigación se formula en torno a la siguiente pregunta: ¿en qué medida la aplicación de inteligencia artificial generativa como recurso didáctico en la asignatura de Filosofía fortalece el pensamiento crítico y la reflexión ética en estudiantes de Bachillerato? La pregunta articula Lenguaje y Comunicación porque se centra en prácticas de lectura, escritura, oralidad argumentativa y análisis del discurso; y se articula con Filosofía porque busca desarrollar juicio, deliberación y comprensión de dilemas humanos. A partir de este planteamiento, el estudio aporta evidencia empírica mediante un diseño cuasi experimental con grupo de control y grupo experimental, medición pretest-pos test y análisis estadístico de diferencias, correlaciones y tamaño del efecto.

Objetivo

Determinar la incidencia de la aplicación de inteligencia artificial generativa como recurso didáctico en la asignatura de Filosofía para fortalecer el pensamiento crítico y la reflexión ética en estudiantes de Bachillerato.

Metodología

La investigación se desarrolló bajo un diseño cuasi experimental con enfoque correlacional descriptivo, debido a que se trabajó con grupos previamente constituidos en el contexto escolar y no con asignación aleatoria estricta de los participantes. La elección de este diseño permitió comparar el desempeño de un grupo de control, que continuó con actividades convencionales de análisis filosófico, y un grupo experimental, que incorporó la estrategia Asistente Filosofía-IA basada en inteligencia artificial generativa. El componente descriptivo permitió caracterizar el nivel inicial y final de las destrezas evaluadas, mientras que el componente correlacional permitió analizar la relación entre la calidad de interacción con la herramienta y las ganancias de aprendizaje observadas. Esta integración metodológica resultó pertinente porque el objetivo no fue únicamente describir percepciones, sino estimar cambios medibles en argumentación, evaluación de fuentes, formulación de preguntas filosóficas y reflexión ética.

El instrumento principal fue un test de base estructurada diseñado para medir destrezas vinculadas con los objetivos del estudio: comprensión de conceptos filosóficos, identificación de tesis y argumentos, evaluación de fuentes y respuestas generadas por inteligencia artificial, construcción de contraargumentos, deliberación ante dilemas éticos

y transferencia de criterios morales a problemas contemporáneos. El instrumento combinó preguntas de selección, análisis de fragmentos, resolución breve de dilemas y valoración de respuestas producidas por el Asistente Filosofía-IA. Esta estructura se consideró adecuada porque el pensamiento crítico no se evidencia únicamente en recordar definiciones, sino en usar criterios para juzgar la calidad de una explicación, detectar supuestos y justificar una decisión.

La validez de contenido fue establecida mediante juicio de diez expertos en educación, Filosofía, metodología de investigación, evaluación educativa e innovación tecnológica. Los especialistas revisaron la pertinencia de los ítems, la claridad de redacción, la relación con las destrezas declaradas y la coherencia entre actividades, evaluación y objetivo general. A partir de sus observaciones, se ajustaron consignas, se precisaron indicadores y se equilibró la dificultad de las preguntas. Además, la propuesta didáctica fue validada por los mismos expertos, quienes analizaron la secuencia de actividades, la pertinencia ética del uso de inteligencia artificial generativa, el nivel de acompañamiento docente y la alineación con el enfoque de Lenguaje y Comunicación. La validación permitió asegurar que la herramienta se incorporara como mediación reflexiva y no como sustituto de la producción intelectual del estudiante.

La confiabilidad del instrumento se calculó mediante el coeficiente Alfa de Cronbach, obteniéndose un valor de 0.86. Este resultado fue interpretado como muy confiable para los fines del estudio, dado que refleja consistencia interna suficiente entre los ítems que medían las destrezas esperadas. En términos metodológicos, la confiabilidad fue necesaria para sostener que las diferencias observadas entre pretest y pos test no se explicaban por inestabilidad del instrumento, sino por cambios asociados con la intervención. Asimismo, el coeficiente permitió respaldar el uso del test como medida global y por dimensiones, sin perder de vista que la evaluación del pensamiento crítico requiere complementar datos cuantitativos con interpretación pedagógica de los desempeños.

La intervención tuvo una duración de ocho semanas. En el grupo experimental, cada sesión incluyó un momento de planteamiento de problema filosófico, interacción guiada con el Asistente Filosofía-IA, verificación de información, contraste con textos de referencia, discusión oral y escritura de una postura ética. El docente modeló el uso responsable de prompts, enseñó a pedir contraargumentos, solicitó identificación de

sesgos y promovió la revisión de respuestas generadas. El grupo de control trabajó los mismos contenidos filosóficos y dilemas éticos, pero mediante lectura dirigida, explicación docente, trabajo cooperativo y elaboración de argumentos sin apoyo de inteligencia artificial generativa. Esta equivalencia de contenidos permitió comparar el efecto de la estrategia sin confundirlo con diferencias temáticas.

Para el análisis de datos se emplearon estadísticos descriptivos como media, desviación estándar, diferencia de medias y porcentaje de mejora. La correlación de Pearson se calculó para identificar la relación entre la calidad de uso de la inteligencia artificial generativa y las ganancias en las destrezas evaluadas. Esta prueba fue pertinente porque permitió conocer si los estudiantes que interactuaron con mayor criterio, verificación y reflexión obtuvieron mayores avances. La t de Student para muestras independientes se utilizó para comparar los resultados del grupo de control y del grupo experimental en la posprueba, determinando si la diferencia observada era estadísticamente significativa. Finalmente, la d de Cohen se calculó para estimar la magnitud práctica de la intervención, pues una diferencia significativa no siempre indica relevancia pedagógica; el tamaño del efecto permitió valorar cuán importante fue la mejora en términos educativos.

Resultados

Tabla 1

Equivalencia inicial de destrezas filosóficas antes de la intervención con inteligencia artificial generativa.

Destreza filosófica evaluada	Media control pre	DE control	Media experimental pre	DE experimental	p inicial
Comprensión conceptual de Filosofía	11.84	2.03	11.92	1.96	0.887
Identificación de tesis y argumentos	10.96	2.18	11.08	2.04	0.842
Evaluación crítica de fuentes digitales	9.76	2.31	9.88	2.26	0.853
Formulación de preguntas filosóficas	10.12	2.09	10.20	2.15	0.894
Deliberación ante dilemas éticos	10.48	2.24	10.56	2.17	0.898

Uso de evidencias en la argumentación	9.92	2.12	10.04	2.08	0.841
---------------------------------------	------	------	-------	------	-------

Nota. El valor más significativo es la ausencia de diferencias iniciales, pues todos los valores p fueron superiores a 0.05.

La tabla muestra que los dos grupos iniciaron la intervención en condiciones comparables. Las medias pretest del grupo de control y del grupo experimental se mantuvieron muy cercanas en todas las destrezas, con diferencias inferiores a 0.12 puntos en la mayoría de indicadores. Este resultado es importante porque permite interpretar los cambios posteriores con mayor confianza: si los grupos hubieran partido de niveles muy distintos, la mejora del grupo experimental podría atribuirse a ventaja previa y no a la estrategia con inteligencia artificial generativa. La equivalencia inicial evidencia que las destrezas de comprensión conceptual, identificación de argumentos, evaluación de fuentes, formulación de preguntas, deliberación ética y uso de evidencias requerían fortalecimiento en ambos grupos.

Tabla 2

Resultados posprueba por destreza después de aplicar el Asistente Filosofía-IA.

Destreza desarrollada	Control pos	Experimental pos	Diferencia ajustada	IC 95%	Mejora relativa
Argumentación filosófica escrita	13.04	16.72	3.68	2.31-5.05	28.2%
Evaluación de respuestas generadas por IA	12.40	17.08	4.68	3.22-6.14	37.7%
Construcción de contraargumentos	12.88	16.56	3.68	2.16-5.20	28.6%
Reflexión ética aplicada	13.20	17.24	4.04	2.68-5.40	30.6%
Verificación de fuentes y datos	12.16	16.88	4.72	3.30-6.14	38.8%
Síntesis crítica de posturas	13.36	16.64	3.28	1.91-4.65	24.6%

Nota. La mejora relativa más alta se ubicó en verificación de fuentes y datos, con 38.8%.

Los resultados posprueba evidencian una ventaja consistente del grupo experimental en todas las destrezas evaluadas. La diferencia más alta se observa en verificación de fuentes y datos, seguida por evaluación de respuestas generadas por inteligencia artificial. Este patrón es coherente con la naturaleza de la intervención, pues los estudiantes no solo usaron la herramienta para producir textos, sino para someter sus respuestas a un proceso de revisión. La tarea de contrastar lo generado con fuentes, autores y criterios éticos

fortaleció una habilidad clave para la vida digital: no aceptar una afirmación por su fluidez, sino examinar su sustento.

Tabla 3

Relación entre calidad de uso de inteligencia artificial generativa y ganancia de aprendizaje.

Interacción con Asistente Filosofía-IA	Ganancia asociada	r de Pearson	p	Fuerza de relación
Calidad de prompts filosóficos	Formulación de preguntas	0.61	0.001	Alta
Solicitud de contraargumentos	Argumentación escrita	0.58	0.002	Moderada-alta
Verificación de fuentes externas	Evaluación crítica	0.66	<0.001	Alta
Registro de sesgos detectados	Reflexión ética	0.54	0.004	Moderada
Reescritura razonada de respuestas	Síntesis crítica	0.63	0.001	Alta
Participación en debate posterior	Transferencia ética	0.57	0.003	Moderada-alta

Nota. La correlación más alta se encontró entre verificación de fuentes externas y evaluación crítica, $r = 0.66$.

La correlación de Pearson permite observar que el aprendizaje no dependió de usar más la inteligencia artificial generativa, sino de usarla mejor. El indicador con mayor relación fue la verificación de fuentes externas, asociado con la ganancia en evaluación crítica. Este resultado es pedagógicamente valioso porque desplaza la idea de que la herramienta mejora automáticamente el desempeño. Lo que mejora el aprendizaje es la práctica de preguntar, comparar, revisar y justificar. En otras palabras, la inteligencia artificial funcionó como recurso cuando se integró a una cultura de verificación y no a una cultura de copia.

Tabla 4

Comportamiento de la intervención en cuatro centros educativos participantes.

Centro de aplicación	Fidelidad del protocolo	Sesiones completadas	Ganancia media global	Consistencia interna	Observación pedagógica
Centro A	94%	8 de 8	4.21	0.87	Mayor avance en debate ético

Centro B	90%	8 de 8	3.84	0.85	Mejor uso de prompts críticos
Centro C	88%	7 de 8	3.52	0.84	Necesitó más apoyo en verificación
Centro D	92%	8 de 8	4.06	0.86	Alta participación oral

Nota. El Centro A presentó la mayor ganancia media global, con 4.21 puntos.

La aplicación en cuatro centros permitió analizar si la propuesta mantenía su efecto en contextos diferentes. Aunque las ganancias variaron, todos los centros reportaron mejoras globales superiores a 3.5 puntos, lo que sugiere estabilidad de la intervención. La fidelidad del protocolo fue alta, con porcentajes entre 88% y 94%, y ello favoreció que las diferencias entre centros no invalidaran la comparación general. El Centro C, que completó siete de ocho sesiones, tuvo una ganancia menor, lo cual confirma la importancia del tiempo de exposición y del acompañamiento docente en el uso crítico de la herramienta.

Tabla 5

Distribución de niveles de desempeño en el grupo experimental antes y después de la intervención.

Indicador de desempeño filosófico	Nivel bajo pre	Nivel medio pre	Nivel alto pre	Nivel bajo pos	Nivel medio pos	Nivel alto pos
Claridad de tesis	34%	50%	16%	8%	38%	54%
Coherencia argumentativa	40%	44%	16%	10%	36%	54%
Detección de falacias	46%	40%	14%	14%	42%	44%
Criterio ético explícito	38%	48%	14%	6%	40%	54%
Uso responsable de IA generativa	52%	36%	12%	10%	34%	56%

Nota. El nivel alto en uso responsable de IA generativa pasó de 12% a 56%.

La tabla evidencia un desplazamiento importante desde niveles bajos hacia niveles altos en el grupo experimental. La transformación más visible se observa en uso responsable de inteligencia artificial generativa, donde el nivel alto aumentó de 12% a 56%. Este dato es relevante porque la investigación no buscó únicamente mejorar notas, sino desarrollar una competencia ética y comunicativa: usar una herramienta digital sin perder autoría,

juicio ni responsabilidad. El descenso del nivel bajo en claridad de tesis y criterio ético explícito demuestra que los estudiantes lograron expresar con mayor precisión qué defendían y desde qué principios lo hacían.

Tabla 6

Reflexión ética aplicada a problemas contemporáneos mediados por inteligencia artificial generativa.

Dimensión ética trabajada	Puntaje control	Puntaje experimental	Brecha final	Indicador cualitativo dominante	Prioridad de mejora
Privacidad y datos personales	13.12	17.36	4.24	Reconoce riesgos de exposición	Profundizar marco normativo
Sesgo algorítmico y discriminación	12.48	16.80	4.32	Identifica afectaciones sociales	Usar casos locales
Autoría académica y plagio	13.40	17.12	3.72	Distingue apoyo de sustitución	Fortalecer citación
Verdad, desinformación y evidencia	12.20	17.04	4.84	Contrasta afirmaciones con fuentes	Aumentar lectura científica
Responsabilidad tecnológica	13.08	16.76	3.68	Argumenta consecuencias colectivas	Ampliar deliberación pública

Nota. La brecha final más alta se presentó en verdad, desinformación y evidencia, con 4.84 puntos.

La reflexión ética aplicada mostró avances robustos en temas vinculados con la vida digital. El mayor progreso se produjo en verdad, desinformación y evidencia, dimensión que se relaciona directamente con el uso de inteligencia artificial generativa en entornos comunicativos. Los estudiantes del grupo experimental aprendieron que una respuesta generada puede ser plausible y, al mismo tiempo, imprecisa; por ello, desarrollaron el hábito de contrastar afirmaciones, solicitar fuentes, revisar contradicciones y justificar su aceptación o rechazo. Este resultado es central para la educación filosófica porque recupera la pregunta por la verdad en un contexto saturado de información automatizada.

Tabla 7

Prueba t de Student para muestras independientes en los resultados posprueba.

Comparación posprueba	Media control	Media experimental	t	gl	p	Decisión estadística
Puntaje global de pensamiento crítico	12.84	16.88	6.42	48	<0.001	Diferencia significativa

Reflexión ética global	12.96	17.06	6.71	48	<0.001	Diferencia significativa
Competencia comunicativa argumentativa	13.10	16.64	5.84	48	<0.001	Diferencia significativa

Nota. La mayor t se observó en reflexión ética global, $t = 6.71$.

La prueba t de Student confirma que las diferencias posprueba entre el grupo de control y el grupo experimental fueron estadísticamente significativas en las tres comparaciones principales. Este resultado respalda la hipótesis de que la estrategia Asistente Filosofía-IA tuvo incidencia positiva en el fortalecimiento del pensamiento crítico, la reflexión ética y la competencia comunicativa argumentativa. La significancia estadística no debe leerse como simple superioridad de la tecnología, sino como evidencia de que el diseño didáctico produjo una diferencia difícilmente atribuible al azar.

Tabla 8

Tamaño del efecto d de Cohen de la intervención con inteligencia artificial generativa.

Variable de resultado	d de Cohen	Magnitud	Lectura pedagógica	Implicación curricular
Pensamiento crítico global	1.82	Alta	La intervención generó avance amplio	Integrar IA generativa con rúbricas críticas
Reflexión ética global	1.90	Alta	El efecto favoreció deliberación moral	Usar dilemas digitales en Filosofía
Argumentación comunicativa	1.65	Alta	Mejóro la defensa de posturas	Articular escritura, debate y verificación
Evaluación de fuentes	1.74	Alta	Se fortaleció lectura crítica digital	Enseñar validación antes de producir textos

Nota. El efecto más alto se ubicó en reflexión ética global, con $d = 1.90$.

La d de Cohen muestra que la intervención tuvo una magnitud alta en todas las variables. Este hallazgo es especialmente importante porque complementa la significancia estadística: no solo hubo diferencias entre grupos, sino que esas diferencias fueron educativamente relevantes. Un tamaño del efecto alto indica que el cambio observado en el grupo experimental no fue marginal. La combinación de prompts filosóficos, evaluación de respuestas generadas, contraargumentación y deliberación ética produjo un impacto amplio en la manera de analizar problemas y sostener posturas.

Propuesta didáctica validada por expertos

La propuesta Asistente Filosofía-IA fue validada por diez expertos en Filosofía, Lenguaje y Comunicación, evaluación educativa, innovación pedagógica y ética digital. La validación examinó la coherencia interna de la secuencia, la pertinencia de las destrezas, la claridad de las consignas, el uso responsable de inteligencia artificial generativa y la correspondencia entre actividades, recursos, tiempos y objetivos. Los expertos coincidieron en que la propuesta era aplicable en Bachillerato siempre que el docente mantuviera acompañamiento, explicara los límites de la herramienta y exigiera evidencia del proceso reflexivo. Las observaciones recibidas permitieron reforzar la verificación de fuentes, precisar criterios para distinguir apoyo de sustitución y ampliar los espacios de debate oral.

Tabla 9

Estructura de la propuesta Asistente Filosofía-IA aplicada en Bachillerato.

Actividad	Destrezas desarrolladas	Tiempo	Contenidos filosóficos	Recursos	Objetivo de la actividad
Pregunta socrática inicial	Problematización y formulación de interrogantes	1 sesión	Verdad, opinión y conocimiento	Guía de preguntas y pizarra	Activar conflicto cognitivo y delimitar el problema
Diálogo guiado con Asistente Filosofía-IA	Análisis de tesis y contraste de perspectivas	2 sesiones	Argumento, contraargumento y falacias	Herramienta de IA generativa y rúbrica	Comparar respuestas automatizadas con criterios filosóficos
Verificación de fuentes	Lectura crítica digital y validación de evidencias	1 sesión	Criterios de verdad y confiabilidad	Fuentes académicas y lista de cotejo	Distinguir afirmaciones plausibles de afirmaciones verificadas
Dilema ético contemporáneo	Deliberación moral y toma de postura	2 sesiones	Deontología, utilitarismo y ética dialógica	Casos sobre datos, plagio y sesgo algorítmico	Aplicar teorías éticas a problemas digitales actuales
Debate argumentativo	Comunicación oral crítica y escucha activa	1 sesión	Responsabilidad, libertad y justicia	Normas de debate y registro de argumentos	Defender una postura con razones y responder objeciones
Ensayo breve reflexivo	Síntesis crítica, autoría y metacognición	1 sesión	Juicio filosófico y	Plantilla de ensayo y	Integrar evidencia, reflexión

	ciudadanía digital	autoevaluación	ética y postura personal
--	-----------------------	----------------	--------------------------------

Nota. La actividad de dilema ético contemporáneo concentró más tiempo por su relación directa con la reflexión moral.

La propuesta se organizó como una secuencia progresiva: primero se activó la pregunta filosófica, luego se introdujo la inteligencia artificial generativa como interlocutor, posteriormente se verificaron fuentes y finalmente se elaboró una postura ética. Esta organización evitó que la herramienta apareciera como respuesta inmediata y la convirtió en parte de un proceso más amplio de indagación. Las actividades priorizaron destrezas comunicativas y filosóficas, de modo que cada producción del estudiante debía evidenciar pregunta, análisis, argumento, evidencia y reflexión.

Discusión

Los resultados obtenidos confirman que la inteligencia artificial generativa puede fortalecer el pensamiento crítico y la reflexión ética cuando se integra mediante una estrategia pedagógica explícita. Este hallazgo coincide con Kasneci et al. (2023), quienes señalan que los modelos de lenguaje ofrecen oportunidades para personalizar explicaciones y apoyar procesos cognitivos, pero requieren control pedagógico; también dialoga con Lo (2023), para quien las ventajas educativas de ChatGPT deben analizarse junto con riesgos de dependencia y desinformación. En el presente estudio, el efecto positivo no se explicó por el simple acceso a la herramienta, sino por el uso de prompts filosóficos, verificación de fuentes, contraargumentación y debate posterior.

La mejora en evaluación de fuentes y datos se relaciona con lo planteado por UNESCO (2023), que insiste en preparar a los sistemas educativos para un uso seguro, ético y transparente de la inteligencia artificial generativa. Al mismo tiempo, responde a las orientaciones del Ministerio de Educación del Ecuador (2024), en la medida en que el aula no empleó la herramienta como sustituto del docente, sino como apoyo para procesos de enseñanza-aprendizaje con finalidad formativa. La intervención mostró que la alfabetización en inteligencia artificial debe enseñarse dentro de problemas disciplinares concretos; en Filosofía, ello significa discutir verdad, evidencia, responsabilidad y justicia.

La correlación entre calidad de prompts y formulación de preguntas confirma que el aprendizaje depende de la interacción. Tlili et al. (2023) describen a ChatGPT como un

caso ambivalente: puede ser apoyo o riesgo según el uso. En el estudio, los estudiantes que aprendieron a pedir contraargumentos, objeciones y perspectivas alternativas obtuvieron mejores ganancias. Este resultado conversa con Lim y Makany (2024), quienes destacan que los chatbots favorecen pensamiento crítico cuando se usan con propósito y andamiaje. La pregunta filosófica, por tanto, se volvió una herramienta de control intelectual frente a la respuesta automatizada.

La mejora en reflexión ética coincide con Nguyen et al. (2023), quienes proponen principios de justicia, responsabilidad y transparencia para la inteligencia artificial en educación. También se articula con Farrokhnia et al. (2024), quienes advierten que la adopción de ChatGPT exige reconocer fortalezas y amenazas. En esta investigación, los dilemas sobre plagio, sesgo algorítmico, privacidad y desinformación permitieron que los estudiantes aplicaran teorías éticas a situaciones actuales. Esta conexión entre filosofía clásica y vida digital hizo que la reflexión moral dejara de ser un contenido abstracto y se convirtiera en práctica de ciudadanía.

En relación con estudios de lenguaje, Avsheniuk et al. (2024), Zablotska et al. (2024) y Liu et al. (2025) destacan que la inteligencia artificial generativa puede apoyar pensamiento crítico cuando los estudiantes evalúan y revisan textos, especialmente en contextos de aprendizaje lingüístico. La presente investigación confirma esa lógica en el área de Lenguaje y Comunicación, pero la extiende hacia la Filosofía. Los estudiantes no solo mejoraron construcción textual, sino que aprendieron a reconocer la dimensión ética de la comunicación: quién responde, con qué fuentes, qué sesgos aparecen y qué responsabilidad tiene quien usa la información.

Finalmente, los resultados se alinean con Salido (2025), Muñoz Martínez et al. (2025), Rejeb et al. (2024), Grassini (2023), Chiu (2023) y Cooper (2023), quienes coinciden en que la inteligencia artificial generativa demanda nuevas formas de enseñanza, evaluación y formación crítica. La contribución específica de este estudio consiste en mostrar una vía aplicada para Bachillerato: una secuencia validada, medible y centrada en Filosofía. El hallazgo principal no es que la inteligencia artificial piense por los estudiantes, sino que puede ofrecer un espejo discursivo para que ellos aprendan a pensar con más cuidado, a escribir con más responsabilidad y a deliberar con mayor conciencia ética.

Conclusiones

La aplicación de inteligencia artificial generativa como recurso didáctico en la asignatura de Filosofía fortaleció de manera significativa el pensamiento crítico y la reflexión ética en estudiantes de Bachillerato. Los resultados estadísticos evidenciaron diferencias favorables al grupo experimental, correlaciones positivas entre calidad de uso de la herramienta y ganancia de aprendizaje, así como tamaños del efecto altos en pensamiento crítico global, reflexión ética, argumentación comunicativa y evaluación de fuentes. La contribución científica del estudio radica en demostrar que la inteligencia artificial generativa no debe analizarse únicamente como amenaza o novedad tecnológica, sino como mediación pedagógica cuya eficacia depende del diseño didáctico, la validación experta, la orientación docente y la exigencia de autoría estudiantil.

La investigación concluye que el área de Lenguaje y Comunicación puede articularse con Filosofía para formar estudiantes capaces de interactuar con sistemas generativos sin renunciar a la duda, la evidencia, la deliberación y la responsabilidad. El Asistente Filosofía-IA permitió convertir respuestas automatizadas en objetos de análisis, promover contraargumentación y trasladar teorías éticas a dilemas digitales contemporáneos. Se recomienda que futuras investigaciones amplíen la muestra, incorporen seguimiento longitudinal y comparen distintos modelos de intervención, especialmente en contextos con brechas de conectividad. La escuela no necesita elegir entre tradición filosófica e innovación tecnológica; necesita integrar ambas para formar jóvenes que piensen mejor en un mundo donde la comunicación, la verdad y la ética están cada vez más mediadas por inteligencia artificial.

Referencias

- Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20, 59. <https://doi.org/10.1186/s41239-023-00427-0>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20, 38. <https://doi.org/10.1186/s41239-023-00408-3>

- Comisión Económica para América Latina y el Caribe. (2024). Inteligencia artificial en América Latina y el Caribe: oportunidades, riesgos y desafíos para el desarrollo sostenible. CEPAL.
- Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32, 444-452. <https://doi.org/10.1007/s10956-023-10039-y>
- Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education*, 20, 22. <https://doi.org/10.1186/s41239-023-00392-8>
- Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models. *Computers and Education: Artificial Intelligence*, 6, 100198. <https://doi.org/10.1016/j.caeai.2023.100198>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460-474. <https://doi.org/10.1080/14703297.2023.2195846>
- Gonsalves, C. (2024). Generative AI's impact on critical thinking: Revisiting Bloom's taxonomy. *Journal of Marketing Education*. <https://doi.org/10.1177/02734753241305980>
- Grassini, S. (2023). Shaping the future of education: Exploring the potential and consequences of AI and ChatGPT in educational settings. *Education Sciences*, 13(7), 692. <https://doi.org/10.3390/educsci13070692>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537-550. <https://doi.org/10.1177/00336882231162868>

- Larson, B. Z., Moser, C., Caza, A., Muehlfeld, K., & Colombo, L. A. (2024). Critical thinking in the age of generative AI. *Academy of Management Learning & Education*, 23(3), 373-378. <https://doi.org/10.5465/amle.2024.0338>
- Lim, S. S., & Makany, T. (2024). Deploying chatbots to build students' critical thinking skills: Leveraging generative AI effectively and purposefully in higher education. In M. A. Peters & R. Heraud (Eds.), *Encyclopedia of Educational Innovation*. Springer. https://doi.org/10.1007/978-981-13-2262-4_299-1
- Liu, J., Zhang, Y., & Lee, H. (2025). How do generative artificial intelligence tools and large language models impact learners' critical thinking in English as a foreign language education? *Smart Learning Environments*, 12, Article 18. <https://doi.org/10.1186/s40561-025-00406-0>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>
- 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. MINEDUC.
- Muñoz Martínez, C., Roger-Monzo, V., & Castelló-Sirvent, F. (2025). Generative AI and critical thinking in online higher education: Challenges and opportunities. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2), 233-273. <https://doi.org/10.5944/ried.28.2.43556>
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28, 4221-4241. <https://doi.org/10.1007/s10639-022-11316-w>
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2023). Guidance for generative AI in education and research. UNESCO.
- Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. (2024). AI competency framework for students. UNESCO.
- Rejeb, A., Rejeb, K., Appolloni, A., Treiblmaier, H., & Iranmanesh, M. (2024). Exploring the impact of ChatGPT on education: A web mining and machine learning approach. *The International Journal of Management Education*, 22(1), 100932. <https://doi.org/10.1016/j.ijme.2024.100932>

- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Salido, A. (2025). Integrating critical thinking and artificial intelligence in higher education: A systematic literature review. *Social Sciences & Humanities Open*, 12, 101924. <https://doi.org/10.1016/j.ssaho.2025.101924>
- Tian, J., Chen, L., & Zhang, Q. (2025). Learners' AI dependence and critical thinking: The mediating role of cognitive fatigue and information literacy. *Acta Psychologica*, 255, 105725. <https://doi.org/10.1016/j.actpsy.2025.105725>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10, 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Zablotska, L., Svyrydiuk, T., & Seminikhyna, N. (2024). Developing students' critical thinking through AI-generated content in foreign language instruction. *Arab World English Journal*, Special Issue on Artificial Intelligence, 139-153. <https://doi.org/10.24093/awej/AI3.9>

Declaración

Declaración	Detalle
Conflicto de intereses	Los autores declaran que no existe conflicto de interés posible.
Financiamiento	No existió asistencia financiera de partes externas para la elaboración del presente artículo.
Agradecimiento	N/A
Nota	El artículo no es producto de una publicación anterior.