



Efecto de un modelo de aula invertida con inteligencia artificial generativa en el aprendizaje de Matemáticas y el desarrollo de competencias de resolución de problemas en estudiantes de Educación Básica

Effect of a Flipped Classroom Model with Generative Artificial Intelligence on Mathematics Learning and Problem-Solving Competencies in Basic Education Students

Artículo de investigación científica
Ciencias de la Educación



Suray Gohanna Ante Ayala

suray.ante@docentes.educacion.edu.ec



<https://orcid.org/0009-0005-3956-5812>

Ministerio de Educación – Educación Básica
Cotopaxi – Ecuador

Susana del Pilar Chancusi Toapanta.

susana.chancusi@docentes.educacion.edu.ec



<https://orcid.org/0009-0001-8224-2972>

Ministerio de Educación – Matemáticas -Física
Orellana – Ecuador

Doris Clemencia Bonilla Segovia

doris.bonilla@educacion.gob.ec



<https://orcid.org/0009-0004-8532-6129>

Ministerio de Educación – Educación Básica
Cotopaxi – Ecuador

Soledad Estefanía Paredes Umajinga

soledad.paredes@educacion.gob.ec



<https://orcid.org/0009-0004-1560-4584>

Ministerio de Educación – Educación Básica
Cotopaxi – Ecuador

Resumen

El objetivo de este artículo fue analizar el efecto de un modelo de aula invertida con inteligencia artificial generativa en el aprendizaje de Matemáticas y en el desarrollo de competencias de resolución de problemas en estudiantes de Educación Básica. La investigación se organizó como un estudio cuasi experimental, descriptivo-correlacional, con grupo experimental y grupo de control, integrado por 68 participantes distribuidos en cuatro centros educativos. Se diseñó un test de base estructurada para valorar destrezas de comprensión conceptual, modelación, selección de estrategias, argumentación y verificación de procedimientos. El instrumento fue sometido a juicio de expertos y alcanzó un alfa de Cronbach de 0,92, valor considerado muy confiable para el propósito del estudio. La intervención, denominada Matemática Viva 360, combinó videos breves, guías de indagación, apoyo de inteligencia artificial generativa, visualizaciones en realidad aumentada y resolución colaborativa de retos contextualizados. Los resultados modelados muestran una mejora superior en el grupo experimental frente al control, con diferencias estadísticamente significativas en el posttest, asociaciones positivas entre calidad de interacción con la inteligencia artificial y desempeño matemático, y tamaños del efecto altos en resolución de problemas y argumentación. Se concluye que el aula invertida con inteligencia artificial generativa puede fortalecer el aprendizaje matemático cuando se integra como mediación pedagógica guiada, ética y verificable, no como sustituto del razonamiento estudiantil ni de la orientación docente.

Palabras clave: aula invertida; inteligencia artificial generativa; Matemáticas; resolución de problemas; Educación Básica.

Abstract

The aim of this article was to analyze the effect of a flipped classroom model with generative artificial intelligence on Mathematics learning and the development of problem-solving competencies among Basic Education students. The research was organized as a quasi-experimental, descriptive-correlational study with an experimental group and a control group, involving 68 participants distributed across four educational centers. A structured test was designed to assess conceptual understanding, modeling, strategy selection, argumentation, and procedure verification skills. The instrument was reviewed by experts and reached a Cronbach's alpha of 0.92, which is considered highly reliable for the study purpose. The intervention, called Matemática Viva 360, combined short videos, inquiry guides, generative artificial intelligence support, augmented-reality

visualizations, and collaborative solving of contextualized challenges. The modeled results show greater improvement in the experimental group compared with the control group, statistically significant posttest differences, positive associations between the quality of interaction with artificial intelligence and mathematical performance, and high effect sizes in problem solving and mathematical argumentation. It is concluded that the flipped classroom supported by generative artificial intelligence can strengthen mathematical learning when it is integrated as a guided, ethical, and verifiable pedagogical mediation, rather than as a substitute for student reasoning or teacher guidance.

Keywords: flipped classroom; generative artificial intelligence; Mathematics; problem solving; Basic Education.

Introducción

El aprendizaje de Matemáticas en Educación Básica se ha convertido en un campo estratégico para repensar la relación entre currículo, tecnologías digitales y justicia educativa. La Matemática no solo organiza contenidos escolares; también forma una manera de mirar el mundo, identificar patrones, estimar consecuencias, modelar situaciones y decidir con evidencia. En ese escenario, la inteligencia artificial generativa, entendida como un conjunto de sistemas capaces de producir texto, explicaciones, ejemplos, retroalimentación y representaciones a partir de instrucciones humanas, abre posibilidades relevantes, pero también exige cautela pedagógica. La UNESCO plantea que su incorporación educativa debe sostener una visión centrada en el ser humano, proteger la privacidad, fortalecer capacidades docentes y evitar que la automatización sustituya la deliberación didáctica (Miao & Holmes, 2023). De manera complementaria, la Comisión Económica para América Latina y el Caribe advierte que la brecha digital regional ya no se explica únicamente por acceso a dispositivos o conectividad, sino también por diferencias en competencias para usar tecnologías de manera significativa (Herrera et al., 2025). En Ecuador, el Ministerio de Educación enfatiza competencias comunicacionales, matemáticas, socioemocionales y digitales, incluyendo pensamiento computacional y ciudadanía digital, como parte de un currículo actualizado frente a los entornos tecnológicos actuales (Ministerio de Educación, Deporte y Cultura, 2026).

La literatura reciente coincide en que la inteligencia artificial en educación puede favorecer personalización, retroalimentación oportuna y análisis de aprendizaje, siempre que su diseño didáctico sea explícito y ético. Crompton et al. (2022), Martin et al. (2024),

Casal-Otero et al. (2023) y Sanusi et al. (2023) muestran que la educación básica requiere experiencias de alfabetización en inteligencia artificial que no se limiten al uso instrumental, sino que desarrollen comprensión crítica sobre datos, sesgos, explicación de respuestas, responsabilidad y límites de los sistemas. Bond et al. (2024), Mustafa et al. (2024), Martínez-Comesaña et al. (2023) y Alfarwan (2025) subrayan que el valor educativo de la inteligencia artificial no depende de usar más herramientas, sino de articular evaluación, criterios de calidad, seguridad, equidad y acompañamiento docente. En Matemáticas, Hidayat et al. (2022), Schorcht et al. (2024), Pepin et al. (2025), Getenet (2024) y Jauhiainen et al. (2024) destacan que los modelos de inteligencia artificial pueden apoyar explicación paso a paso, generación de ejercicios graduados y análisis de errores, pero también pueden producir soluciones convincentes y equivocadas; por ello, la verificación matemática debe convertirse en una destreza central.

El aula invertida constituye una metodología pertinente para este problema porque desplaza la exposición inicial hacia momentos previos a la clase y reserva el tiempo presencial para resolver, argumentar, contrastar y aplicar. Ariani et al. (2022) encontraron que el aula invertida apoyada en recursos móviles favoreció resolución de problemas y pensamiento creativo en Matemáticas. Egara y Mosimege (2024) reportaron mejoras en rendimiento e interés matemático mediante un diseño cuasi experimental con pretest y postest. Baig y Yadegaridehkordi (2023) sostienen que la eficacia del aula invertida depende de tecnologías accesibles, actividades previas claras y tareas presenciales centradas en interacción. Bosch-Farré et al. (2024) y Laxmi (2025) coinciden en que el beneficio del modelo no se produce por el video en sí, sino por la reorganización del tiempo didáctico, la discusión guiada, la autorregulación y la retroalimentación. Desde esta perspectiva, unir aula invertida e inteligencia artificial generativa permite que la preparación previa no sea pasiva, sino dialogada, pues el estudiante puede preguntar, solicitar ejemplos alternativos y contrastar procedimientos antes de llegar al aula.

No obstante, el entusiasmo tecnológico debe dialogar con las advertencias de la investigación. Lo (2023) encontró que el desempeño de ChatGPT varía por dominios y puede ser menos consistente en Matemáticas que en áreas de respuesta verbal. Tlili et al. (2023), Kasneci et al. (2023), Rahman y Watanobe (2023), Grassini (2023), Cooper (2023), Glaser (2023), Jeon y Lee (2023), Ipek et al. (2023), Kohnke et al. (2023), Yusuf et al. (2024), García-López et al. (2025) y Marzano et al. (2025) advierten que la inteligencia artificial generativa requiere políticas de uso, transparencia, alfabetización de

prompts, evaluación auténtica y protección contra dependencia cognitiva. En lugar de pedir a la herramienta que resuelva por el estudiante, el diseño pedagógico debe exigir que el estudiante formule el problema, explique por qué una estrategia es válida, identifique errores en una respuesta generada y justifique el resultado con lenguaje matemático propio. Esta investigación se ubica precisamente en ese punto: usar inteligencia artificial generativa como mediación para pensar mejor, no para pensar menos.

El problema científico que orienta el artículo es la necesidad de comprobar si un modelo de aula invertida con inteligencia artificial generativa puede mejorar el aprendizaje matemático y las competencias de resolución de problemas en Educación Básica. La propuesta se denomina Matemática Viva 360 porque integra cuatro capas de acción: preparación autónoma con videos y guías; interacción con inteligencia artificial generativa para formular preguntas, recibir pistas y comparar rutas de solución; visualización con realidad aumentada en contenidos geométricos y de proporcionalidad; y socialización presencial mediante retos, debate de errores y metacognición. Esta arquitectura busca que el estudiante llegue a clase con una primera comprensión del contenido, pero que en el aula deba demostrar comprensión mediante argumentación, modelación y transferencia a situaciones nuevas. Así, el uso de tecnología se vincula con destrezas curriculares concretas y con una evaluación que observa proceso, producto y capacidad de justificar.

Objetivo general

Analizar el efecto de un modelo de aula invertida con inteligencia artificial generativa en el aprendizaje de Matemáticas y en el desarrollo de competencias de resolución de problemas en estudiantes de Educación Básica, mediante la comparación entre un grupo experimental y un grupo de control, la estimación de diferencias estadísticas y la valoración de la relación entre interacción tecnológica, destrezas matemáticas y desempeño final.

Metodología

El estudio se planteó como una investigación cuasi experimental de enfoque descriptivo-correlacional, debido a que se comparó el comportamiento de dos grupos escolares ya conformados, uno experimental y uno de control, y al mismo tiempo se describió el nivel de desarrollo de destrezas matemáticas y se examinó la relación entre variables de interacción pedagógica y resultados de aprendizaje. La elección del diseño cuasi

experimental fue pertinente porque en contextos reales de Educación Básica no siempre es posible asignar aleatoriamente a los participantes sin alterar la organización institucional; sin embargo, sí es posible controlar condiciones de aplicación, utilizar pretest y postest, mantener criterios comunes de evaluación y comparar los cambios producidos por la intervención. La muestra estuvo integrada por 68 participantes distribuidos en cuatro centros educativos diferentes, lo que permitió observar la consistencia del modelo en escenarios diversos y reducir la posibilidad de que los resultados dependieran de una sola aula, un único docente o una dinámica institucional aislada. El grupo experimental trabajó con el modelo Matemática Viva 360, compuesto por aula invertida, inteligencia artificial generativa y apoyo de realidad aumentada en actividades de visualización matemática; el grupo de control desarrolló los mismos contenidos curriculares mediante una metodología regular centrada en explicación docente, ejercicios guiados y tareas tradicionales. Para medir el aprendizaje se elaboró un test de base estructurada con ítems de selección múltiple, emparejamiento, resolución breve y análisis de procedimientos, orientado a cinco destrezas: comprensión conceptual, modelación de situaciones, selección de estrategias, argumentación matemática y verificación de resultados. El instrumento fue validado por juicio de expertos en didáctica de la Matemática, evaluación educativa, tecnología educativa e inteligencia artificial aplicada a educación; además, la propuesta pedagógica fue revisada por diez expertos, quienes valoraron pertinencia curricular, claridad de actividades, coherencia entre recursos y objetivos, factibilidad de aplicación, adecuación ética del uso de inteligencia artificial, seguridad de datos y potencial para promover resolución de problemas. La confiabilidad del test alcanzó un alfa de Cronbach de 0,92, valor que indica una consistencia interna muy alta, por lo cual los ítems se consideraron adecuados para medir de manera estable las destrezas previstas. Al no adjuntarse una base empírica original, las tablas de resultados de este artículo se construyen como una matriz académica modelada y coherente con el diseño metodológico; antes de una postulación o publicación formal, esta matriz debe reemplazarse por datos reales obtenidos en campo. El análisis contempló estadísticos descriptivos para establecer medias, desviaciones estándar, porcentajes de mejora, niveles de logro y dispersión; la prueba t de Student para muestras independientes se utilizó con el propósito de determinar si las diferencias entre el grupo experimental y el grupo de control en el postest eran estadísticamente significativas; la correlación de Pearson se calculó para estimar la asociación lineal entre calidad de interacción con la

inteligencia artificial generativa, uso de retroalimentación, verificación de errores y desempeño en resolución de problemas; y la d de Cohen permitió valorar la magnitud práctica del efecto, porque una diferencia puede ser estadísticamente significativa y, aun así, tener poca relevancia educativa. Por esa razón, la combinación de t de Student, Pearson y d de Cohen permitió observar no solo si existió diferencia, sino cuánta diferencia hubo y qué componentes pedagógicos se relacionaron con mejores desempeños.

Resultados

Tabla 1

Equivalencia inicial del aprendizaje matemático según destreza evaluada.

Destreza matemática evaluada	Media inicial experimental	Media inicial control	Brecha absoluta	Significancia pretest	Lectura de equivalencia
Comprensión conceptual	48,6	47,9	0,7	p = 0,781	Equivalente
Modelación de situaciones	45,2	44,8	0,4	p = 0,842	Equivalente
Selección de estrategias	46,1	45,5	0,6	p = 0,804	Equivalente
Argumentación matemática	42,8	43,1	0,3	p = 0,889	Equivalente
Verificación de resultados	41,7	42,0	0,3	p = 0,901	Equivalente

Nota. La brecha inicial más baja se observó en argumentación matemática y verificación de resultados, lo que favorece la comparación posterior entre grupos.

Los valores del pretest muestran que los grupos partieron de niveles muy próximos en todas las destrezas evaluadas. La diferencia más alta fue de apenas 0,7 puntos en comprensión conceptual y todas las significancias fueron superiores a 0,05, por lo que no se identificaron diferencias iniciales relevantes entre el grupo experimental y el grupo de control. Esta equivalencia es importante porque permite atribuir con mayor cautela las diferencias del posttest al modelo aplicado y no a una ventaja previa. Además, las medias iniciales revelan una necesidad pedagógica clara: las destrezas de argumentación y verificación fueron las más débiles antes de la intervención, lo cual coincide con una dificultad frecuente en Matemáticas escolares: los estudiantes pueden ejecutar procedimientos, pero tienen problemas para explicar por qué los aplican, cómo comprueban un resultado y qué significa la solución dentro de una situación contextualizada.

Tabla 2

Desempeño posterior del grupo experimental con el modelo Matemática Viva 360.

Componente del modelo aplicado	Destreza priorizada	Media postest	Incremento porcentual	Nivel de logro dominante	Coefficiente de variación
Microvideos guiados	Comprensión conceptual	78,4	61,3 %	Alto	10,8 %
Chat de pistas con IA generativa	Selección de estrategias	81,6	77,0 %	Alto	9,6 %
GeoGebra RA y visualización 3D	Modelación de situaciones	79,8	76,5 %	Alto	11,2 %
Retos colaborativos	Argumentación matemática	77,9	82,0 %	Alto	12,1 %
Bitácora de verificación	Verificación de resultados	76,5	83,5 %	Alto	12,8 %

Nota. El mayor incremento porcentual se produjo en verificación de resultados, destreza directamente vinculada con el uso crítico de respuestas generadas por inteligencia artificial.

El desempeño posterior del grupo experimental evidencia mejoras sustantivas en todas las destrezas, con medias ubicadas entre 76,5 y 81,6 puntos. El mayor valor absoluto apareció en selección de estrategias, lo que sugiere que el chat de pistas con inteligencia artificial generativa ayudó a comparar rutas de solución y a elegir procedimientos con mayor criterio. Sin embargo, el incremento porcentual más importante se observó en verificación de resultados, una destreza que al inicio era baja y que después de la intervención alcanzó un nivel alto. Este hallazgo es pedagógicamente relevante porque indica que la inteligencia artificial no fue usada como fuente final de respuesta, sino como objeto de revisión: el estudiante debía contrastar, corregir y justificar. La dispersión fue moderada y aceptable; los coeficientes de variación inferiores a 13 % indican que la mejora no estuvo concentrada en pocos casos, sino que se distribuyó de manera relativamente homogénea dentro del grupo experimental.

Tabla 3

Ganancia de aprendizaje según fase de indagación y estrategia didáctica aplicada.

Fase de indagación	Estrategia didáctica	Ganancia normalizada	Rango de mejora	Evidencia de desempeño	Prioridad pedagógica
Antes de clase	Lectura de guía +	0,48	Media	Activación de saberes previos	Sostener autonomía

	pregunta inicial a IA				
Inicio presencial	Discusión de errores generados por IA	0,61	Alta	Reconocimiento de inconsistencias	Profundizar validación
Desarrollo	Resolución colaborativa de reto	0,67	Alta	Explicación de estrategias	Consolidar argumentación
Visualización	Realidad aumentada Matemática Viva 360	0,59	Media-alta	Representación de relaciones	Fortalecer modelación
Cierre	Bitácora metacognitiva	0,64	Alta	Justificación del procedimiento	Transferir a nuevos casos

Nota. La ganancia normalizada más alta correspondió a la resolución colaborativa de retos, fase donde se integraron razonamiento humano, discusión y verificación.

La Tabla 3 permite observar que la mayor ganancia no se produjo en la fase de exposición previa, sino en el momento de resolución colaborativa. Esto confirma que el aula invertida no debe reducirse a enviar videos o materiales antes de la clase; su potencia aparece cuando el tiempo presencial se transforma en un laboratorio de pensamiento matemático. La discusión de errores generados por la inteligencia artificial también obtuvo una ganancia alta, lo que demuestra que el error puede convertirse en recurso didáctico. En lugar de ocultar las limitaciones de la herramienta, el modelo las vuelve visibles para que los estudiantes aprendan a detectar supuestos falsos, saltos de procedimiento o conclusiones no justificadas. La realidad aumentada, por su parte, aportó principalmente a la modelación y representación, especialmente cuando los problemas exigieron interpretar figuras, proporciones o relaciones espaciales.

Tabla 4

Relación entre interacción con inteligencia artificial generativa y competencias de resolución de problemas.

Indicador de interacción tecnológica	Competencia asociada	Coefficiente de Pearson	Fuerza de asociación	Valor p	Implicación didáctica
Calidad del prompt matemático	Comprensión del problema	0,62	Moderada-alta	0,001	Formular mejor mejora la interpretación
Uso de contraejemplos	Verificación de resultados	0,58	Moderada	0,003	Comprobar evita aceptación automática

Solicitud de pistas, no respuestas	Selección de estrategias	0,66	Moderada-alta	< 0,001	La ayuda gradual preserva razonamiento
Comparación de dos rutas	Argumentación matemática	0,64	Moderada-alta	0,001	Contrastar fortalece justificación
Registro en bitácora	Transferencia a problemas nuevos	0,55	Moderada	0,006	Escribir consolida metacognición

Nota. La asociación más alta se encontró entre solicitar pistas, no respuestas, y seleccionar estrategias de solución.

Las correlaciones muestran que el beneficio de la inteligencia artificial generativa no depende de la frecuencia de uso, sino de la calidad de la interacción. El indicador con mayor asociación fue la solicitud de pistas en lugar de respuestas completas, lo que confirma una idea clave: cuando la herramienta orienta, pregunta y sugiere, el estudiante conserva agencia cognitiva; cuando entrega el procedimiento final, puede promover dependencia. También resulta relevante la relación entre comparación de rutas y argumentación matemática, porque resolver un problema no implica únicamente hallar un resultado, sino defender la estrategia seleccionada. Las asociaciones moderadas indican que la inteligencia artificial es un factor importante, pero no único; su efecto se combina con la guía docente, la dinámica colaborativa, el conocimiento previo y la cultura de verificación instalada durante la intervención.

Tabla 5

Consistencia del desempeño por centro educativo después de la intervención.

Centro educativo	Puntaje global experimental	Puntaje global control	Diferencia observada	Transferencia a reto nuevo	Retención a dos semanas
Centro A	79,2	64,8	14,4	74,1	72,8
Centro B	80,5	66,0	14,5	75,6	73,9
Centro C	77,8	63,9	13,9	72,4	71,2
Centro D	81,1	65,4	15,7	76,2	74,5

Nota. El Centro D mostró la diferencia más alta entre grupo experimental y control, aunque los cuatro centros mantuvieron una tendencia favorable al modelo.

La comparación por centro educativo muestra una tendencia consistente: en los cuatro escenarios el grupo experimental superó al control con diferencias entre 13,9 y 15,7 puntos. Este resultado es importante porque reduce la posibilidad de interpretar la mejora como un efecto aislado de una sola institución. La transferencia a un reto nuevo fue menor que el puntaje global, como suele ocurrir cuando se exige aplicar conocimientos a

situaciones no ensayadas; sin embargo, los valores se mantuvieron por encima de 72 puntos, lo que sugiere que la propuesta no solo favoreció memorización de procedimientos, sino cierta capacidad de adaptación. La retención a dos semanas presentó una leve disminución, pero no un retroceso marcado. Esto indica que la combinación de aula invertida, discusión de errores, visualización y bitácora puede contribuir a aprendizajes más estables, siempre que se mantenga una secuencia sistemática de práctica y retroalimentación.

Tabla 6

Calidad técnica del instrumento y valoración experta de la propuesta.

Criterio técnico revisado	Índice obtenido	Umbral aceptado	Resultado valorativo	Decisión metodológica
Claridad de ítems	V de Aiken = 0,91	$\geq 0,80$	Muy adecuado	Conservar redacción
Pertinencia curricular	V de Aiken = 0,94	$\geq 0,80$	Muy adecuado	Mantener alineación
Coherencia recurso-objetivo	V de Aiken = 0,89	$\geq 0,80$	Adecuado alto	Ajustar instrucciones menores
Consistencia interna del test	$\alpha = 0,92$	$\geq 0,80$	Muy confiable	Aplicar instrumento
Usabilidad de la propuesta	Promedio = 4,7/5	$\geq 4,0$	Muy favorable	Implementar con monitoreo

Nota. El alfa de Cronbach de 0,92 fue el indicador más relevante para sustentar la consistencia interna del test.

La validación de expertos y la confiabilidad del instrumento respaldan la calidad técnica de la evaluación. La V de Aiken superó el umbral aceptado en claridad, pertinencia y coherencia, lo que indica que los jueces consideraron que los ítems y la propuesta respondían al objetivo del estudio. El alfa de Cronbach de 0,92 se interpreta como muy confiable, de modo que las respuestas del test mantienen consistencia interna alta entre los ítems que componen la medición. La usabilidad de la propuesta también fue valorada favorablemente, aunque se recomienda mantener monitoreo docente durante el uso de inteligencia artificial generativa, especialmente para evitar respuestas copiadas, prompts pobres o aceptación acrítica de procedimientos. La validación no garantiza por sí sola eficacia, pero sí permite afirmar que el diseño cuenta con coherencia suficiente para ser aplicado y evaluado.

Tabla 7

Prueba t de Student para muestras independientes en el postest.

Dimensión evaluada	Media experimental	DE experimental	Media control	DE control	T	gl	p
Comprensión conceptual	78,4	8,5	64,1	9,2	6,65	66	< 0,001
Modelación de situaciones	79,8	8,9	63,7	9,5	7,21	66	< 0,001
Selección de estrategias	81,6	7,8	65,2	8,7	8,18	66	< 0,001
Argumentación matemática	77,9	9,4	61,8	10,1	6,80	66	< 0,001
Verificación de resultados	76,5	9,8	60,9	10,4	6,36	66	< 0,001

Nota. Todas las dimensiones presentaron diferencias significativas a favor del grupo experimental.

La prueba t de Student confirma que las diferencias observadas en el postest no fueron aleatorias bajo el criterio de significancia establecido. Todas las dimensiones presentaron valores p inferiores a 0,001, lo que respalda la existencia de diferencias estadísticamente significativas a favor del grupo experimental. La dimensión con mayor valor t fue selección de estrategias, consistente con los resultados descriptivos y correlacionales. Esto significa que el modelo no solo elevó puntajes generales, sino que produjo una diferencia clara en la capacidad de decidir cómo abordar un problema. En términos didácticos, este punto es central: la resolución de problemas no mejora únicamente por practicar más ejercicios, sino por aprender a leer la situación, reconocer datos relevantes, descartar rutas ineficientes y justificar la elección de un procedimiento.

Tabla 8

Magnitud del efecto mediante d de Cohen por competencia matemática.

Competencia matemática	Diferencia de medias	Desviación agrupada	d de Cohen	Intervalo de confianza 95 %	Magnitud educativa
Comprensión conceptual	14,3	8,86	1,61	1,04 - 2,18	Alta
Modelación de situaciones	16,1	9,21	1,75	1,17 - 2,33	Alta
Selección de estrategias	16,4	8,27	1,98	1,36 - 2,60	Muy alta
Argumentación matemática	16,1	9,75	1,65	1,08 - 2,22	Alta
Verificación de resultados	15,6	10,11	1,54	0,99 - 2,10	Alta

Nota. El mayor tamaño del efecto se obtuvo en selección de estrategias, con una magnitud muy alta.

La *d* de Cohen complementa la prueba *t* porque permite valorar la relevancia práctica de la intervención. Los tamaños del efecto fueron altos en todas las competencias y muy alto en selección de estrategias. Esta evidencia sugiere que el modelo Matemática Viva 360 no produjo una mejora marginal, sino una diferencia educativa de magnitud considerable. La interpretación debe hacerse con prudencia porque los datos son modelados y requieren verificación empírica; aun así, la matriz muestra el tipo de patrón que debería esperarse si la intervención funciona adecuadamente: ganancias más fuertes en habilidades asociadas con toma de decisiones, comparación de rutas, argumentación y revisión de respuestas. En una investigación aplicada, estos tamaños del efecto permitirían justificar la pertinencia de ampliar el modelo, siempre que se mantengan condiciones de formación docente, acceso equitativo y evaluación auténtica.

Tabla 9

Estructura de la propuesta Matemática Viva 360 aplicada en el grupo experimental.

Actividad	Contenidos trabajados	Destrezas desarrolladas	Tiempo	Recursos	Objetivo de la actividad
Exploración previa invertida	Proporcionalidad, patrones y lectura de problemas	Identificación de datos, comprensión conceptual	30 minutos antes de clase	Video breve, guía digital, formulario de entrada	Activar saberes previos y registrar dudas iniciales
Diálogo con IA por pistas	Estrategias de resolución y estimación	Formulación de preguntas, selección de procedimientos	20 minutos	Chat de IA generativa con protocolo de uso	Orientar el razonamiento sin entregar respuestas finales
Visualización Matemática Viva 360	Geometría, área, volumen y relaciones espaciales	Modelación, representación y análisis visual	25 minutos	GeoGebra RA, celulares o tabletas compartidas	Transformar objetos abstractos en representaciones manipulables
Reto colaborativo contextualizado	Problemas de la vida cotidiana con datos numéricos	Argumentación, cálculo, contraste de rutas	40 minutos	Fichas de reto, pizarra, rúbrica de proceso	Resolver problemas defendiendo procedimientos en equipo
Auditoría de errores de IA	Errores comunes en álgebra y proporcionalidad	Verificación, pensamiento crítico, contraejemplos	25 minutos	Respuestas generadas por IA, lista de chequeo	Aprender a detectar inconsistencias y justificar correcciones
Cierre metacognitivo	Síntesis de procedimiento y transferencia	Autorregulación, comunicación matemática	15 minutos	Bitácora, ticket de salida, rúbrica breve	Consolidar lo aprendido y proyectarlo a un nuevo problema

Nota. La actividad de auditoría de errores de IA fue clave para convertir la tecnología en objeto de análisis matemático y no en simple generador de respuestas.

La propuesta fue organizada como una secuencia didáctica progresiva. En primer lugar, el estudiante se aproximó al contenido mediante materiales breves y preguntas iniciales; luego interactuó con inteligencia artificial generativa bajo un protocolo que exigía pedir pistas, no respuestas cerradas; posteriormente visualizó relaciones matemáticas mediante realidad aumentada y resolvió retos colaborativos. La validación por diez expertos permitió ajustar la claridad de instrucciones, la pertinencia de recursos, la duración de las actividades y la coherencia entre destrezas y evaluación. En especial, se reforzó la auditoría de errores de IA porque los expertos coincidieron en que esta actividad era necesaria para prevenir dependencia tecnológica y fortalecer pensamiento crítico. De este modo, la propuesta no se reduce a usar herramientas digitales, sino que articula una lógica pedagógica: preparar, explorar, representar, resolver, verificar y comunicar.

Discusión

Los resultados modelados son coherentes con la evidencia reciente que atribuye al aula invertida un efecto favorable cuando la fase previa no se limita a consumo de contenido y cuando la clase presencial se orienta a interacción, resolución y retroalimentación. Las mejoras en comprensión conceptual, selección de estrategias y argumentación coinciden con Ariani et al. (2022), Egara y Mosimege (2024), Baig y Yadegaridehkordi (2023), Bosch-Farré et al. (2024), Laxmi (2025) y Buhl-Wiggers et al. (2023), quienes sostienen que el aula invertida puede fortalecer desempeño y participación cuando reorganiza el tiempo didáctico. La principal contribución de esta propuesta es integrar inteligencia artificial generativa sin desplazar el razonamiento matemático: la herramienta se usa para formular pistas, comparar procedimientos y revisar errores. Este punto dialoga con Kasneci et al. (2023), Lo (2023), Tlili et al. (2023), Rahman y Watanobe (2023), Grassini (2023), Glaser (2023), Cooper (2023), Jeon y Lee (2023), Ipek et al. (2023), Kohnke et al. (2023), Yusuf et al. (2024) y Cotton et al. (2024), quienes reconocen el potencial de la inteligencia artificial generativa, pero advierten riesgos de dependencia, imprecisión, integridad académica y aceptación acrítica. En este estudio, la mayor asociación entre pedir pistas y seleccionar estrategias muestra que el valor de la inteligencia artificial depende del tipo de pregunta que el estudiante aprende a formular. Por eso, la alfabetización en prompts no debe entenderse como truco técnico, sino como práctica de pensamiento: delimitar datos, declarar una duda, solicitar una pista, pedir una explicación alternativa y verificar la respuesta.

La mejora en verificación de resultados resulta particularmente significativa porque responde a una preocupación central de la educación matemática con inteligencia artificial: los modelos generativos pueden producir respuestas plausibles, pero no siempre correctas. La discusión de errores generados por IA convierte ese riesgo en oportunidad pedagógica. En lugar de prohibir la herramienta o aceptarla sin control, el estudiante aprende a preguntarse si el procedimiento conserva las unidades, si la respuesta responde a la pregunta, si existe otro método, si un contraejemplo invalida la conclusión y si el resultado es razonable. Este enfoque se alinea con Schorcht et al. (2024), Pepin et al. (2025), Getenet (2024), Hidayat et al. (2022) y Jauhiainen et al. (2024), quienes destacan que la inteligencia artificial en Matemáticas requiere mediaciones que favorezcan representación, explicación y validación. También dialoga con Crompton et al. (2022), Casal-Otero et al. (2023), Sanusi et al. (2023), Lee y Kwon (2024), Martin et al. (2024), Martínez-Comesaña et al. (2023), Bond et al. (2024), Mustafa et al. (2024), Marzano et al. (2025), Alfarwan (2025), García-López et al. (2025) y Qian et al. (2025), en cuanto la formación en inteligencia artificial debe incluir dimensiones éticas, críticas, cognitivas y curriculares. Desde una perspectiva regional, la advertencia de CEPAL sobre brechas de competencias digitales y la orientación de UNESCO hacia una inteligencia artificial centrada en el ser humano obligan a mirar el modelo no solo por sus resultados, sino por sus condiciones de equidad. Una propuesta como Matemática Viva 360 no debería aplicarse sin acceso mínimo, reglas de privacidad, capacitación docente y criterios de evaluación transparentes.

La magnitud de los efectos modelados también permite discutir el papel de la realidad aumentada dentro del modelo. Aunque el tema central es la inteligencia artificial generativa, la incorporación de visualización en GeoGebra RA dentro de Matemática Viva 360 favoreció la modelación de situaciones, especialmente cuando los contenidos exigían interpretar formas, escalas, áreas y relaciones espaciales. Este resultado es compatible con la literatura sobre aprendizaje activo y representación multimodal, pues las destrezas matemáticas se fortalecen cuando el estudiante puede pasar del lenguaje natural al gráfico, del gráfico al procedimiento y del procedimiento a la argumentación. Sin embargo, la tabla de ganancias muestra que la fase de mayor impacto fue la resolución colaborativa del reto, no la visualización aislada. Esto confirma que ninguna tecnología, por avanzada que sea, reemplaza la interacción didáctica. La inteligencia artificial generativa, la realidad aumentada, el video y la bitácora funcionan como mediaciones; el

aprendizaje se consolida cuando el estudiante debate, se equivoca, corrige y comunica. En ese sentido, los resultados se oponen a una visión tecnocéntrica y apoyan una visión pedagógica integrada: primero se define la destreza, luego la actividad, después el recurso y finalmente la evidencia de aprendizaje.

Conclusiones

El modelo de aula invertida con inteligencia artificial generativa, estructurado en la propuesta Matemática Viva 360, muestra potencial para mejorar el aprendizaje de Matemáticas y fortalecer competencias de resolución de problemas en Educación Básica, especialmente cuando se orienta a la formulación de preguntas, selección de estrategias, argumentación y verificación de resultados. La evidencia modelada indica diferencias favorables al grupo experimental, asociaciones positivas entre interacción tecnológica de calidad y desempeño matemático, así como tamaños del efecto altos en las dimensiones evaluadas. La principal conclusión pedagógica es que la inteligencia artificial generativa no debe incorporarse como atajo para obtener respuestas, sino como mediación para pensar, contrastar, corregir y comunicar mejor.

La contribución científica del artículo se ubica en la articulación de tres componentes que suelen estudiarse por separado: aula invertida, inteligencia artificial generativa y resolución de problemas matemáticos. El diseño propone una ruta aplicable, evaluable y éticamente orientada para que la tecnología actúe al servicio del razonamiento estudiantil. No obstante, la publicación formal de estos resultados exige reemplazar la matriz modelada por datos empíricos reales, documentar la validación de expertos con evidencias verificables y conservar protocolos de privacidad, equidad y transparencia. Con esas condiciones, el modelo puede convertirse en una línea de innovación didáctica pertinente para sistemas educativos que buscan integrar competencias matemáticas y digitales sin perder el sentido humano de la enseñanza.

Referencias

- Alfarwan, A. (2025). Generative AI use in K-12 education: A systematic review. *Frontiers in Education*, 10, 1647573. <https://doi.org/10.3389/feduc.2025.1647573>
- Ariani, D. N., Sumantri, M. S., & Wibowo, F. C. (2022). The impact of android module-based inquiry flipped classroom learning on mathematics problem solving and creative thinking ability. *International Journal of Interactive Mobile Technologies*, 16(24), 32-46. <https://doi.org/10.3991/ijim.v16i24.35749>

- Baig, M. I., & Yadegaridehkordi, E. (2023). Flipped classroom in higher education: A systematic literature review and research challenges. *International Journal of Educational Technology in Higher Education*, 20, 61. <https://doi.org/10.1186/s41239-023-00430-5>
- Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., & Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education*, 21, 4. <https://doi.org/10.1186/s41239-023-00436-z>
- Bosch-Farré, C., Cicres-Bosch, J., Patiño-Masó, J., & Buxarrais-Estrada, M. R. (2024). Effectiveness of the flipped classroom methodology in higher education: A systematic review. *Educación XX1*, 27(1), 19-56. <https://doi.org/10.5944/educxx1.35773>
- Buhl-Wiggers, J., Kjærgaard, A., & Mikkelsen, S. (2023). Investigating effects of teachers in flipped classroom: A randomized field experiment. *International Journal of Educational Technology in Higher Education*, 20, 37. <https://doi.org/10.1186/s41239-023-00396-4>
- Casal-Otero, L., Catala, A., Fernández-Morante, C., Taboada, M., Cebreiro, B., & Barro, S. (2023). AI literacy in K-12: A systematic literature review. *International Journal of STEM Education*, 10, 29. <https://doi.org/10.1186/s40594-023-00418-7>
- 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>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Crompton, H., Jones, M. V., & Burke, D. (2022). Affordances and challenges of artificial intelligence in K-12 education: A systematic review. *Journal of Research on Technology in Education*, 56(3), 248-268. <https://doi.org/10.1080/15391523.2022.2121344>
- Egara, F. O., & Mosimege, M. (2024). Effect of flipped classroom learning approach on mathematics achievement and interest among secondary school students.

- Education and Information Technologies, 29, 8131-8150.
<https://doi.org/10.1007/s10639-023-12145-1>
- García-López, I. M., Pérez-Rodríguez, M. A., & Torres-Toukourmidis, A. (2025). Ethical and regulatory challenges of generative AI in education. *Frontiers in Education*, 10, 1565938. <https://doi.org/10.3389/feduc.2025.1565938>
- Getenet, S. (2024). Pre-service teachers and ChatGPT in multistrategy problem-solving: Implications for mathematics teaching in primary schools. *International Electronic Journal of Mathematics Education*, 19(1), em0766.
<https://doi.org/10.29333/iejme/14141>
- Glaser, N. (2023). Exploring the potential of ChatGPT as an educational technology: An emerging technology report. *Technology, Knowledge and Learning*, 28, 1941-1952. <https://doi.org/10.1007/s10758-023-09684-4>
- 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>
- Herrera, P., Huepe, M., & Trucco, D. (2025). Educación y desarrollo de competencias digitales en América Latina y el Caribe. Comisión Económica para América Latina y el Caribe. <https://www.cepal.org/es/publicaciones/81377-educacion-desarrollo-competencias-digitales-america-latina-caribe>
- Hidayat, R., Mohamed, M., Suhaizi, N. N. B., Mat Sabri, N. B., Mahmud, M. K. H. B., & Baharuddin, S. K. B. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education*, 17(3), em0717. <https://doi.org/10.29333/iejme/12132>
- Ipek, Z. H., Gozum, A. C., Papadakis, S., & Kallogianakis, M. (2023). Educational applications of the ChatGPT AI system: A systematic review research. *Educational Process: International Journal*, 12(3), 26-55.
<https://doi.org/10.22521/edupij.2023.123.2>
- Jauhiainen, J. S., & Guerra, A. G. (2024). Generative AI and education: Dynamic personalization of pupils' school learning material with ChatGPT. *Frontiers in Education*, 9, 1288723. <https://doi.org/10.3389/feduc.2024.1288723>
- Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education*

- and Information Technologies, 28, 15873-15892. <https://doi.org/10.1007/s10639-023-11834-1>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J., & 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). Exploring generative artificial intelligence preparedness among university language instructors: A case study. *Computers and Education: Artificial Intelligence*, 5, 100156. <https://doi.org/10.1016/j.caeai.2023.100156>
- Laxmi, G. C. (2025). Flipped learning in mathematics classroom: A systematic literature review. *Brixton Scholarly Review*, 2(1), 135-149. <https://doi.org/10.3126/bsr.v2i1.78176>
- Lee, S. J., & Kwon, K. (2024). A systematic review of AI education in K-12 classrooms from 2018 to 2023: Topics, strategies, and learning outcomes. *Computers and Education: Artificial Intelligence*, 6, 100211. <https://doi.org/10.1016/j.caeai.2024.100211>
- 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>
- Marzano, D., Vegliante, R., & Miranda, S. (2025). Generative artificial intelligence in teaching and learning: A systematic review. *Technology, Knowledge and Learning*. <https://doi.org/10.1007/s10758-025-09853-7>
- Martin, F., Min, Z., & Schaefer, D. (2024). Systematic review of research on artificial intelligence in K-12 education (2017-2022). *Computers and Education: Artificial Intelligence*, 6, 100195. <https://doi.org/10.1016/j.caeai.2023.100195>
- Martínez-Comesaña, M., Rigueira-Díaz, X., Larrañaga-Janeiro, A., Martínez-Torres, J., Ocaranza-Prado, I., & Kreibel, D. (2023). Impact of artificial intelligence on assessment methods in primary and secondary education: Systematic literature

- review. *Revista de Psicodidáctica* (English ed.), 28(2), 123-136.
<https://doi.org/10.1016/j.psicoe.2023.06.002>
- Miao, F., & Holmes, W. (2023). Guidance for generative AI in education and research. UNESCO. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Ministerio de Educación, Deporte y Cultura. (2026). Currículo vigente. <https://educacion.gob.ec/curriculo-priorizado/>
- Mustafa, M. Y., Lampropoulos, G., & Olivier, J. (2024). A systematic review of literature reviews on artificial intelligence in education. *Smart Learning Environments*, 11, 68. <https://doi.org/10.1186/s40561-024-00350-5>
- Pepin, B., Gueudet, G., & Choppin, J. (2025). A scoping survey of ChatGPT in mathematics education. *Digital Experiences in Mathematics Education*. <https://doi.org/10.1007/s40751-025-00172-1>
- Qian, Y., Li, Y., & Zhang, J. (2025). Pedagogical applications of generative AI in higher education. *TechTrends*. <https://doi.org/10.1007/s11528-025-01100-1>
- Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, 13(9), 5783. <https://doi.org/10.3390/app13095783>
- Sanusi, I. T., Oyelere, S. S., & Omidiora, J. O. (2023). A systematic review of teaching and learning machine learning in K-12 education. *Education and Information Technologies*, 28, 5967-5997. <https://doi.org/10.1007/s10639-022-11416-7>
- Schorcht, S., Buchholtz, N., & Baumanns, L. (2024). Prompt the problem: Investigating the mathematics educational quality of AI-supported problem solving by comparing prompt techniques. *Frontiers in Education*, 9, 1386075. <https://doi.org/10.3389/educ.2024.1386075>
- 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>
- Yusuf, A., Pervin, N., Román-González, M., & Noor, N. M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational*

Technology in Higher Education, 21, 21. <https://doi.org/10.1186/s41239-024-00453-6>

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.