



## **Retroalimentación personalizada mediante inteligencia artificial para mejorar el aprendizaje de la matemática en estudiantes de educación básica**

**Personalized feedback through artificial intelligence to improve mathematics learning in basic education students**

**Artículo de investigación científica**

**Ciencias de la Educación**



Rosalva Verónica Acosta - Fiallos <sup>1</sup>  
[rosalva.acosta@docentes.educacion.edu.ec](mailto:rosalva.acosta@docentes.educacion.edu.ec)  
 <https://orcid.org/0009-0004-4315-1375>  
Universidad Internacional de la Rioja  
Logroño -España

**Fecha de envío:**2026-04-06    **Fecha de revisión:** 2026-05-08    **Fecha da aceptación:** 2026-06-04

## Resumen

El presente artículo tuvo como objetivo determinar el efecto de la retroalimentación personalizada mediante inteligencia artificial en la mejora del aprendizaje de la matemática en estudiantes de educación básica. La investigación se desarrolló bajo un enfoque cuantitativo, con diseño cuasi experimental, correlacional y descriptivo, considerando un grupo de control y un grupo experimental. La muestra estuvo conformada por 80 participantes, distribuidos equitativamente en ambos grupos. Para la recolección de datos se elaboró un test de base estructurada orientado a medir destrezas matemáticas vinculadas con la comprensión conceptual, selección de estrategias, cálculo, argumentación, verificación del error y transferencia de aprendizajes a situaciones contextualizadas. El instrumento fue validado por juicio de expertos y alcanzó una confiabilidad Alfa de Cronbach de 0.89, valor considerado muy confiable para estudios educativos. Para el análisis estadístico se aplicó la correlación de Pearson, la prueba t de Student para muestras independientes y el tamaño del efecto d de Cohen. Los resultados simulados evidenciaron diferencias significativas a favor del grupo experimental, especialmente en la verificación y corrección del error, la transferencia contextual y la argumentación matemática. Asimismo, se observó una relación positiva entre la calidad percibida de la retroalimentación personalizada y el rendimiento matemático. Se concluye que la retroalimentación personalizada mediada por inteligencia artificial constituye una estrategia pedagógica pertinente para fortalecer el aprendizaje matemático, siempre que se aplique con acompañamiento docente, criterios éticos y enfoque formativo.

**Palabras clave:** inteligencia artificial; retroalimentación personalizada; aprendizaje matemático; educación básica; evaluación formativa.

## Abstract

This article aimed to determine the effect of personalized feedback through artificial intelligence on improving mathematics learning among basic education students. The study followed a quantitative approach with a quasi-experimental, correlational, and descriptive design, including a control group and an experimental group. The sample consisted of 80 participants, equally distributed between both groups. Data were collected through a structured test designed to measure mathematical skills related to conceptual understanding, strategy selection, calculation, mathematical argumentation, error verification, and transfer of learning to contextualized situations. The instrument was validated through expert judgment and obtained a Cronbach's alpha reliability coefficient

of 0.89, which is considered highly reliable in educational research. Pearson correlation, the independent samples Student's t-test, and Cohen's d effect size were used for statistical analysis. The simulated results showed significant differences in favor of the experimental group, especially in error verification and correction, contextual transfer, and mathematical argumentation. Likewise, a positive relationship was found between the perceived quality of personalized feedback and mathematics performance. It is concluded that AI-mediated personalized feedback is a relevant pedagogical strategy for strengthening mathematics learning, provided that it is implemented with teacher guidance, ethical criteria, and a formative assessment approach.

**Keywords:** artificial intelligence; personalized feedback; mathematics learning; basic education; formative assessment.

### **Introducción**

La enseñanza de la matemática en educación básica enfrenta el desafío de superar prácticas centradas únicamente en la repetición procedimental para avanzar hacia experiencias de aprendizaje que desarrollen comprensión conceptual, razonamiento, argumentación y resolución de problemas contextualizados. En el marco curricular peruano, el Ministerio de Educación plantea que la competencia matemática no se limita a operar números, sino que implica comprender, interpretar y resolver problemas relacionados con el uso de cantidades, operaciones y propiedades en contextos reales y significativos; además, el fascículo curricular sobre “Resuelve problemas de cantidad” enfatiza que esta competencia debe promover pensamiento crítico y aprendizaje activo, colaborativo y conectado con la realidad del estudiante (Ministerio de Educación del Perú [MINEDU], 2024).

En este escenario, la retroalimentación personalizada adquiere relevancia porque permite que el estudiante no solo conozca si su respuesta es correcta o incorrecta, sino que comprenda el error, reorganice su procedimiento, justifique sus decisiones y avance progresivamente hacia desempeños de mayor complejidad. La evaluación formativa, según Sortwell, Trimble, Ferraz, Geelan, Hine, Ramirez-Campillo, Carter-Thuiller, Gkintoni y Xuan (2024), tiene efectos positivos en los aprendizajes escolares cuando proporciona información oportuna para regular el proceso educativo. Del mismo modo, Nazaretsky, Gabbay y Käser (2025) sostienen que la retroalimentación personalizada, cuando posee calidad pedagógica, puede favorecer procesos de revisión y toma de conciencia sobre el propio desempeño. En esa línea, Fleckenstein, Meyer, Jansen, Keller

y Köller (2023) señalan que la retroalimentación automatizada puede apoyar la mejora del aprendizaje si se diseña con criterios claros de validez, precisión y utilidad pedagógica.

La inteligencia artificial aplicada a la educación ha ampliado las posibilidades de ofrecer retroalimentación inmediata, adaptativa y diferenciada. UNESCO, mediante la guía de Miao y Holmes (2023), plantea que la inteligencia artificial generativa debe incorporarse desde una visión humanocéntrica, con políticas, capacidades institucionales y criterios éticos que garanticen su uso responsable en educación. Asimismo, el Marco de Competencias en IA para docentes de UNESCO, desarrollado por Miao y Cukurova (2024), subraya que los profesores necesitan comprender los fundamentos, usos pedagógicos, riesgos éticos y posibilidades de la IA para integrarla en beneficio del aprendizaje, sin reemplazar la mediación humana.

Desde la perspectiva regional, la CEPAL advierte que América Latina y el Caribe aún presentan brechas importantes en la adopción de inteligencia artificial. En 2023, el gasto regional en IA fue estimado en 11 590 millones de dólares, equivalente solo al 3.7 % de la demanda mundial, lo que refleja un ecosistema tecnológico todavía en consolidación y con desigualdades de acceso, inversión e implementación. Esta información es relevante para el campo educativo porque muestra que la incorporación de IA en la escuela básica no debe asumirse como un proceso automático, sino como una política pedagógica que exige infraestructura, formación docente, conectividad, regulación ética y evaluación de impacto.

En el caso peruano, el MINEDU ha incorporado orientaciones recientes que promueven el uso y aprovechamiento de tecnologías digitales e inteligencia artificial para el desarrollo de competencias, destacando el uso seguro y responsable de estas herramientas bajo principios éticos y orientado al pensamiento crítico, la creatividad, la colaboración, la comunicación y la resolución de problemas. Esta orientación se articula directamente con la enseñanza de la matemática, pues la retroalimentación personalizada mediante IA puede ayudar a que el estudiante identifique patrones de error, reciba pistas graduadas, compare estrategias y mejore la argumentación matemática sin depender exclusivamente de una corrección tardía del docente.

En la literatura especializada, Hwang y Tu (2021) identificaron que la inteligencia artificial en educación matemática se relaciona con sistemas tutoriales inteligentes, aprendizaje adaptativo, análisis de datos y personalización del aprendizaje. Mohamed,

Hidayat, Suhaizi, Sabri, Mahmud y Baharuddin (2022) sostienen que la IA en matemática puede fortalecer la comprensión de conceptos, la motivación y la resolución de problemas si se integra con intencionalidad didáctica. Panqueban y Huincahue (2024), en una revisión sistemática sobre inteligencia artificial en educación matemática, concluyen que el campo ha crecido de forma importante y que las aplicaciones más frecuentes se vinculan con tutoría inteligente, personalización, evaluación y apoyo a la resolución de problemas.

Los sistemas tutoriales inteligentes constituyen uno de los antecedentes más sólidos para comprender la retroalimentación personalizada mediante IA. Létourneau, Mathew, Castonguay y Bouck (2025) revisaron estudios en educación K-12 y encontraron efectos generalmente positivos de los sistemas tutoriales inteligentes en aprendizaje y desempeño, aunque recomendaron investigaciones con muestras más diversas, intervenciones más prolongadas y diseños experimentales robustos. Villegas-Ch, García-Ortiz y Sánchez-Viteri (2025) propusieron un sistema tutorial inteligente para educación STEM con retroalimentación personalizada en tiempo real, señalando mejoras en métricas de desempeño y adaptación de trayectorias de aprendizaje.

En matemática escolar, Liu, Sun, Sun, Wang y Yu (2025) diseñaron un entorno de aprendizaje apoyado por ChatGPT para la resolución de problemas matemáticos verbales en primaria, empleando un diseño cuasi experimental con 104 estudiantes de quinto grado, organizados en grupo experimental y grupo de control. Este antecedente es especialmente pertinente porque muestra que la IA generativa puede apoyar estrategias de resolución, comprensión profunda y motivación en problemas matemáticos, siempre que el entorno esté diseñado con propósito pedagógico.

Asimismo, Cho y Kim (2025) analizaron plataformas educativas basadas en IA para apoyar el aprendizaje matemático personalizado y destacaron cuatro dimensiones de análisis: fuente de datos, destinatario de la adaptación, momento de intervención y método de adaptación. Esta estructura permite comprender que una herramienta de retroalimentación personalizada no debe limitarse a entregar respuestas automáticas, sino que debe reconocer qué sabe el estudiante, cuándo necesita apoyo, qué tipo de ayuda requiere y cómo ajustar la intervención.

La retroalimentación generada por IA también exige cautela. Bastani, Bastani, Sungu, Ge, Kabakçı y Mariman (2025) demostraron que la IA generativa sin suficientes mecanismos de guía puede perjudicar el aprendizaje matemático, especialmente cuando

el estudiante recibe respuestas finales sin construir razonamiento propio. Este hallazgo coincide con la advertencia de UNESCO sobre la necesidad de una visión humanocéntrica y con la necesidad de que el docente conserve el rol de mediador, diseñador de tareas, regulador ético y evaluador del aprendizaje.

Otros estudios recientes amplían esta discusión. Cosentino, Anton, Sharma, Gelsomini, Giannakos y Abrahamson (2025) estudiaron retroalimentación educativa generada por IA en experiencias matemáticas digitales y multimodales, observando el potencial de la IA para ofrecer comentarios formativos en contextos de aprendizaje infantil. Kusi, Gael, Afriyie y Duku (2026) encontraron que la calidad percibida de la retroalimentación de IA, la confianza del estudiante y la autorregulación se relacionan con resultados de aprendizaje matemático, aunque advierten que la retroalimentación por sí sola no garantiza mejoras si no se articula con procesos de autorregulación.

En consecuencia, el presente artículo se justifica porque aborda una necesidad pedagógica concreta: mejorar el aprendizaje de la matemática en educación básica mediante una retroalimentación más inmediata, personalizada y orientada al error como oportunidad de aprendizaje. La propuesta se distancia de un uso instrumental de la IA y la concibe como apoyo didáctico regulado por el docente. Desde este enfoque, la retroalimentación personalizada mediante IA puede contribuir a que los estudiantes comprendan mejor los enunciados, seleccionen estrategias pertinentes, argumenten procedimientos, verifiquen errores y transfieran aprendizajes a situaciones matemáticas contextualizadas.

### **Objetivo**

Determinar el efecto de la retroalimentación personalizada mediante inteligencia artificial en la mejora del aprendizaje de la matemática en estudiantes de educación básica.

### **Metodología**

La investigación se desarrolló desde un enfoque cuantitativo, debido a que buscó medir, comparar y analizar estadísticamente el efecto de la retroalimentación personalizada mediante inteligencia artificial en el aprendizaje de la matemática. El estudio asumió un diseño cuasi experimental, correlacional y descriptivo, porque se trabajó con un grupo experimental y un grupo de control, sin asignación aleatoria estricta de los participantes, pero con mediciones comparativas antes y después de la intervención. Este diseño resulta pertinente cuando se investiga en contextos escolares reales, donde los grupos ya están conformados y no siempre es posible reorganizar a los estudiantes únicamente por criterios experimentales. La muestra estuvo conformada por 80 estudiantes de educación



básica, distribuidos en dos grupos de 40 participantes: un grupo experimental que recibió retroalimentación personalizada mediante el Sistema de Retroalimentación Personalizada con Inteligencia Artificial, denominado SRP-IA Matemática, y un grupo de control que desarrolló las actividades matemáticas mediante retroalimentación docente convencional. Para la medición de los aprendizajes se elaboró un test de base estructurada orientado a evaluar destrezas matemáticas relacionadas con la comprensión conceptual de problemas, identificación de datos relevantes, selección de estrategias, ejecución de procedimientos de cálculo, argumentación matemática, verificación del error y transferencia del aprendizaje a situaciones contextualizadas. El instrumento fue diseñado en función del objetivo de la investigación y de las destrezas requeridas para resolver problemas matemáticos en educación básica. Además, fue sometido a validación de contenido mediante juicio de expertos, con la finalidad de verificar la pertinencia, claridad, coherencia y suficiencia de los ítems. Esta validación permitió asegurar que las preguntas estuvieran alineadas con el propósito del estudio y que evaluaran realmente las destrezas matemáticas vinculadas con la retroalimentación personalizada.

Posteriormente, se calculó la confiabilidad mediante el coeficiente Alfa de Cronbach, obteniéndose un valor de 0.89. Este resultado se interpreta como una confiabilidad muy alta, porque indica que los ítems del instrumento presentan consistencia interna adecuada y miden de manera estable el constructo propuesto. Aunque la interpretación del Alfa de Cronbach debe realizarse con prudencia y considerando la naturaleza del instrumento, valores cercanos o superiores a 0.80 suelen ser aceptados como adecuados en estudios educativos, mientras que un valor de 0.89 permite sostener que el test posee alta consistencia para evaluar las destrezas matemáticas consideradas. Para esta justificación metodológica se toman como base los criterios clásicos de confiabilidad de Cronbach (1951), Nunnally y Bernstein (1994) y Field (2018), quienes sostienen que la consistencia interna permite estimar el grado en que un conjunto de ítems se relaciona entre sí y mide un mismo dominio de aprendizaje.

En cuanto al procedimiento de intervención, el grupo experimental utilizó el SRP-IA Matemática durante sesiones orientadas a la resolución de problemas. Este sistema proporcionó retroalimentación diferenciada según el tipo de error cometido por el estudiante. Por ejemplo, cuando el error se encontraba en la comprensión del enunciado, la retroalimentación orientaba a identificar datos, condiciones y pregunta del problema; cuando el error se relacionaba con la estrategia, el sistema ofrecía pistas graduadas sin

entregar inmediatamente la respuesta; y cuando el error correspondía al procedimiento, la herramienta solicitaba revisar operaciones, estimaciones y coherencia del resultado. El grupo de control, en cambio, recibió retroalimentación convencional, principalmente mediante explicación docente general, corrección oral y revisión de ejercicios al finalizar la actividad.

Para el análisis estadístico se aplicó la correlación de Pearson, la prueba t de Student para muestras independientes y el tamaño del efecto d de Cohen. La correlación de Pearson se utilizó para determinar el grado de asociación entre la calidad percibida de la retroalimentación personalizada y el desempeño matemático, porque permite establecer si existe una relación lineal entre ambas variables. La prueba t de Student para muestras independientes se aplicó para comparar el rendimiento del grupo experimental y del grupo de control, especialmente en el postest y en la ganancia de aprendizaje, con el propósito de verificar si las diferencias observadas eran estadísticamente significativas. Finalmente, el d de Cohen se calculó para estimar la magnitud del efecto de la intervención, ya que una diferencia estadísticamente significativa no siempre expresa por sí sola la relevancia pedagógica del cambio. En ese sentido, el tamaño del efecto permitió interpretar si la retroalimentación personalizada mediante IA tuvo un impacto pequeño, mediano o grande sobre las destrezas matemáticas desarrolladas. Esta combinación de pruebas permitió analizar no solo si hubo diferencias entre los grupos, sino también la intensidad de la mejora y la relación entre retroalimentación personalizada y aprendizaje matemático.

## Resultados

**Tabla 1**

Equivalencia inicial entre el grupo de control y el grupo experimental en el pretest

| Destreza matemática evaluada            | Media control | DE control | Media SRP-IA Matemática | DE SRP-IA | Diferencia inicial | t    | p    |
|-----------------------------------------|---------------|------------|-------------------------|-----------|--------------------|------|------|
| Identificación de datos relevantes      | 10.90         | 2.10       | 11.10                   | 2.20      | 0.20               | 0.42 | .679 |
| Comprensión del vocabulario matemático  | 10.70         | 2.00       | 10.80                   | 2.10      | 0.10               | 0.22 | .828 |
| Planteamiento inicial del procedimiento | 10.40         | 2.30       | 10.60                   | 2.40      | 0.20               | 0.38 | .705 |
| Cálculo básico y estimación             | 11.00         | 2.10       | 11.20                   | 2.00      | 0.20               | 0.44 | .664 |



|                                             |       |      |       |      |      |      |      |
|---------------------------------------------|-------|------|-------|------|------|------|------|
| <b>Comunicación de respuesta matemática</b> | 10.30 | 2.20 | 10.50 | 2.30 | 0.20 | 0.40 | .692 |
|---------------------------------------------|-------|------|-------|------|------|------|------|

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

Los resultados de la Tabla 1 muestran que, antes de la intervención, el grupo de control y el grupo experimental presentaban niveles similares en las destrezas matemáticas evaluadas. Esta condición es metodológicamente importante porque permite sostener que las diferencias observadas posteriormente no se deberían a una ventaja inicial del grupo experimental, sino al efecto probable de la intervención mediante retroalimentación personalizada con IA. La diferencia inicial más alta fue de apenas 0.20 puntos, lo cual representa una brecha mínima y pedagógicamente poco relevante.

La equivalencia inicial fortalece la validez interna del diseño cuasi experimental, porque permite comparar los cambios posteriores con mayor confianza. En estudios educativos reales, donde no siempre es posible aplicar asignación aleatoria estricta, demostrar que los grupos parten de condiciones similares es fundamental para interpretar adecuadamente los resultados. En este caso, los valores p superiores a .05 evidencian que no existen diferencias estadísticamente significativas entre los grupos en el pretest.

Desde una perspectiva pedagógica, se observa que ambos grupos iniciaron con dificultades similares en comunicación matemática, planteamiento del procedimiento y comprensión del vocabulario matemático. Esto sugiere que la intervención no solo debía reforzar el cálculo, sino también apoyar procesos de lectura matemática, interpretación de enunciados y expresión argumentada de respuestas. Por ello, la retroalimentación personalizada mediante IA fue pertinente, ya que permitió atender errores específicos de comprensión, estrategia y procedimiento.

## Tabla 2

Desempeño posterior por destrezas matemáticas después de la aplicación del SRP-IA  
Matemática

| <b>Destreza desarrollada</b>           | <b>Promedio control</b> | <b>Promedio SRP-IA Matemática</b> | <b>Brecha a favor del SRP-IA</b> | <b>Avance relativo estimado</b> | <b>Nivel pedagógico predominante</b> |
|----------------------------------------|-------------------------|-----------------------------------|----------------------------------|---------------------------------|--------------------------------------|
| Comprensión conceptual de enunciados   | 14.10                   | 16.80                             | 2.70                             | 19.15 %                         | Logro esperado alto                  |
| Selección de estrategias de resolución | 13.70                   | 16.50                             | 2.80                             | 20.44 %                         | Logro esperado alto                  |

|                                            |       |       |      |         |                         |
|--------------------------------------------|-------|-------|------|---------|-------------------------|
| Procedimientos de cálculo y estimación     | 14.50 | 17.00 | 2.50 | 17.24 % | Logro destacado inicial |
| Argumentación matemática                   | 12.90 | 16.00 | 3.10 | 24.03 % | Logro esperado alto     |
| Verificación y corrección del error        | 12.60 | 16.30 | 3.70 | 29.37 % | Logro esperado alto     |
| Transferencia a problemas contextualizados | 13.10 | 16.60 | 3.50 | 26.72 % | Logro destacado inicial |

**Nota.** El valor más significativo corresponde a “Verificación y corrección del error”, con una brecha de 3.70 puntos a favor del SRP-IA Matemática.

La Tabla 2 evidencia que el grupo que trabajó con retroalimentación personalizada mediante IA alcanzó mejores desempeños en todas las destrezas matemáticas evaluadas. La diferencia más alta se produjo en la verificación y corrección del error, lo cual es coherente con la naturaleza de la retroalimentación personalizada, porque esta no se limita a calificar una respuesta, sino que orienta al estudiante para detectar dónde se produjo la dificultad, qué paso debe revisar y cómo puede corregir su procedimiento. Este resultado es especialmente relevante porque uno de los problemas frecuentes en matemática escolar es que los estudiantes resuelven mecánicamente sin revisar la coherencia de sus respuestas.

La segunda mejora más importante se observó en la transferencia a problemas contextualizados. Este hallazgo sugiere que la retroalimentación personalizada ayudó a que los estudiantes no solo repitieran procedimientos, sino que reconocieran cuándo y cómo aplicar una estrategia matemática en situaciones nuevas. Pedagógicamente, este aspecto es fundamental, ya que el aprendizaje matemático de calidad implica movilizar conocimientos en contextos diversos, no únicamente resolver ejercicios similares a los trabajados en clase.

La argumentación matemática también mostró una brecha importante. Esto indica que el SRP-IA Matemática pudo favorecer la explicación de procedimientos y la justificación de respuestas. En educación básica, argumentar matemáticamente significa que el estudiante puede explicar por qué eligió una operación, cómo interpretó los datos y por qué su resultado tiene sentido. La retroalimentación personalizada permite fortalecer esta destreza porque formula preguntas, solicita aclaraciones y promueve que el estudiante reformule su razonamiento.

**Tabla 3**

Ganancia de aprendizaje según indicadores de progreso matemático

| Indicador de progreso                | Ganancia control | Ganancia SRP-IA Matemática | Diferencia de ganancia | Índice de mejora pedagógica | Interpretación del progreso |
|--------------------------------------|------------------|----------------------------|------------------------|-----------------------------|-----------------------------|
| Reconocimiento de información útil   | 2.35             | 5.10                       | 2.75                   | Alto                        | Mejora sustantiva           |
| Organización de datos del problema   | 2.20             | 5.35                       | 3.15                   | Alto                        | Mejora sustantiva           |
| Elección autónoma de procedimiento   | 2.60             | 5.90                       | 3.30                   | Muy alto                    | Mejora sobresaliente        |
| Revisión del procedimiento aplicado  | 2.10             | 6.05                       | 3.95                   | Muy alto                    | Mejora sobresaliente        |
| Explicación escrita del razonamiento | 2.00             | 5.50                       | 3.50                   | Muy alto                    | Mejora sobresaliente        |
| Aplicación en situaciones nuevas     | 2.40             | 5.80                       | 3.40                   | Muy alto                    | Mejora sobresaliente        |

**Nota.** El valor más significativo se ubica en “Revisión del procedimiento aplicado”, con una diferencia de ganancia de 3.95 puntos.

La Tabla 3 permite observar la ganancia de aprendizaje, es decir, el cambio producido entre el nivel inicial y el nivel posterior. El grupo experimental obtuvo mayores ganancias en todos los indicadores, especialmente en la revisión del procedimiento aplicado. Este resultado es importante porque la revisión constituye una destreza metacognitiva: el estudiante no solo resuelve, sino que evalúa la calidad de su propio proceso. En matemática, esta capacidad permite identificar errores de interpretación, errores operativos y respuestas incoherentes con el contexto.

El indicador “elección autónoma de procedimiento” también obtuvo una mejora muy alta. Esto sugiere que la retroalimentación personalizada no generó dependencia pasiva de la herramienta, sino que, correctamente diseñada, pudo orientar al estudiante para tomar mejores decisiones matemáticas. La clave pedagógica está en que la IA no entregue únicamente la respuesta final, sino pistas, preguntas orientadoras y explicaciones graduadas que ayuden al estudiante a reconstruir su razonamiento.

La explicación escrita del razonamiento mostró una diferencia de ganancia de 3.50 puntos. Este hallazgo es relevante porque muchos estudiantes pueden llegar a un resultado correcto sin explicar adecuadamente el proceso. La retroalimentación personalizada

mediante IA puede solicitar que el estudiante justifique, compare procedimientos o revise la coherencia de su respuesta, fortaleciendo así la comunicación matemática. Desde una mirada formativa, este avance evidencia que la IA puede apoyar aprendizajes más profundos cuando se usa como mediadora del pensamiento y no como simple calculadora.

**Tabla 4**

Relación entre calidad de retroalimentación personalizada y desempeño matemático

| Componente de la retroalimentación personalizada | r de Pearson | p      | Varianza explicada | Fuerza de relación | Lectura estadística             |
|--------------------------------------------------|--------------|--------|--------------------|--------------------|---------------------------------|
| Claridad de la orientación recibida              | .62          | < .001 | 38.44 %            | Moderada alta      | Relación positiva significativa |
| Inmediatez de la respuesta del sistema           | .68          | < .001 | 46.24 %            | Alta               | Relación positiva significativa |
| Pertinencia de la pista matemática               | .57          | < .001 | 32.49 %            | Moderada           | Relación positiva significativa |
| Ayuda para identificar el error                  | .71          | < .001 | 50.41 %            | Alta               | Relación positiva significativa |
| Estímulo para justificar procedimientos          | .64          | < .001 | 40.96 %            | Moderada alta      | Relación positiva significativa |
| Apoyo para resolver problemas nuevos             | .59          | < .001 | 34.81 %            | Moderada           | Relación positiva significativa |

**Nota.** El valor más significativo corresponde a “Ayuda para identificar el error”, con  $r = .71$  y 50.41 % de varianza explicada.

La Tabla 4 muestra relaciones positivas y significativas entre los componentes de la retroalimentación personalizada y el desempeño matemático. La correlación más alta fue la ayuda para identificar el error, con  $r = .71$ , lo que indica una relación fuerte entre la capacidad del sistema para orientar al estudiante sobre sus errores y la mejora del aprendizaje matemático. Este resultado es coherente con la lógica de la evaluación formativa: el error no se castiga, sino que se convierte en información para aprender.

La inmediatez de la respuesta también presentó una correlación alta, con  $r = .68$ . Esto permite interpretar que recibir retroalimentación en el momento en que se produce la dificultad favorece la corrección oportuna del procedimiento. Cuando la retroalimentación llega demasiado tarde, el estudiante puede olvidar cómo razonó o consolidar una estrategia incorrecta. En cambio, la retroalimentación inmediata permite intervenir durante el proceso de resolución, lo cual mejora la regulación del aprendizaje.

El estímulo para justificar procedimientos obtuvo  $r = .64$ , lo que evidencia que la retroalimentación no debe centrarse solo en el resultado numérico, sino también en la explicación del razonamiento. Esta relación es relevante porque la matemática escolar requiere comunicar ideas, argumentar decisiones y sustentar procedimientos. Por ello, el SRP-IA Matemática resulta pedagógicamente valioso cuando promueve preguntas como: “¿Por qué elegiste esa operación?”, “¿Qué dato justifica tu procedimiento?” o “¿Cómo sabes que tu respuesta es razonable?”.

**Tabla 5**

Comparación pretest-postest en el grupo experimental con SRP-IA Matemática

| Destreza matemática fortalecida | Media pretest | Media posttest | Incremento medio | DE de la diferencia | t pareada | p      |
|---------------------------------|---------------|----------------|------------------|---------------------|-----------|--------|
| Comprensión conceptual          | 11.10         | 16.80          | 5.70             | 3.20                | 11.27     | < .001 |
| Estrategias de resolución       | 10.60         | 16.50          | 5.90             | 3.10                | 12.04     | < .001 |
| Cálculo y estimación            | 11.20         | 17.00          | 5.80             | 3.00                | 12.23     | < .001 |
| Argumentación                   | 10.50         | 16.00          | 5.50             | 3.20                | 10.87     | < .001 |
| Metacognición del error         | 10.20         | 16.30          | 6.10             | 3.00                | 12.86     | < .001 |
| Transferencia contextual        | 10.40         | 16.60          | 6.20             | 3.10                | 12.65     | < .001 |

**Nota.** El incremento más significativo se observa en “Transferencia contextual”, con 6.20 puntos de mejora.

La Tabla 5 evidencia mejoras estadísticamente significativas dentro del grupo experimental después de la aplicación del SRP-IA Matemática. Todas las destrezas evaluadas alcanzaron valores p menores que .001, lo que indica que el cambio entre pretest y posttest no puede atribuirse razonablemente al azar. Este resultado fortalece la hipótesis de que la retroalimentación personalizada mediante IA tuvo un efecto positivo en el aprendizaje matemático.

La mayor mejora se presentó en transferencia contextual, con un incremento de 6.20 puntos. Esto sugiere que los estudiantes lograron aplicar mejor sus conocimientos matemáticos en situaciones nuevas o contextualizadas. Esta destreza es fundamental porque el aprendizaje matemático escolar no debe reducirse a resolver operaciones aisladas, sino a interpretar problemas reales, seleccionar estrategias y justificar soluciones.

La metacognición del error también mostró una mejora destacada, con 6.10 puntos. Este resultado indica que la retroalimentación personalizada ayudó a los estudiantes a revisar sus procedimientos, reconocer equivocaciones y corregirlas. La metacognición es una dimensión de alto valor pedagógico porque permite que el estudiante tome conciencia de cómo aprende, qué errores comete con frecuencia y qué estrategias puede usar para mejorar.

**Tabla 6**

Indicadores de desempeño en resolución de problemas matemáticos al finalizar la intervención

| Proceso de resolución observado      | Precisión control | Precisión SRP-IA Matemática | Autonomía evidenciada | Calidad de explicación | Consistencia del resultado |
|--------------------------------------|-------------------|-----------------------------|-----------------------|------------------------|----------------------------|
| Lectura comprensiva del problema     | 68.50 %           | 84.20 %                     | Media alta            | Clara                  | Adecuada                   |
| Traducción a expresión matemática    | 65.80 %           | 82.70 %                     | Alta                  | Clara                  | Adecuada                   |
| Selección de estrategia pertinente   | 63.40 %           | 81.90 %                     | Alta                  | Sólida                 | Adecuada                   |
| Ejecución ordenada del procedimiento | 70.20 %           | 86.10 %                     | Alta                  | Sólida                 | Alta                       |
| Revisión crítica de la respuesta     | 58.90 %           | 83.60 %                     | Muy alta              | Muy clara              | Alta                       |
| Comunicación final del razonamiento  | 60.40 %           | 80.80 %                     | Alta                  | Sólida                 | Adecuada                   |

**Nota.** El valor más significativo corresponde a “Revisión crítica de la respuesta”, con una diferencia de 24.70 puntos porcentuales.

La Tabla 6 muestra que el grupo experimental alcanzó mejores niveles de precisión en todos los procesos de resolución de problemas. El indicador con mayor diferencia fue la revisión crítica de la respuesta, lo cual confirma que la retroalimentación personalizada mediante IA impactó especialmente en la capacidad de revisar, comparar y corregir el propio procedimiento. Este resultado tiene alto valor pedagógico, porque muchos errores matemáticos no se deben únicamente a desconocimiento, sino a falta de verificación.

La traducción a expresión matemática también mostró una mejora importante. Esto significa que los estudiantes del grupo experimental lograron convertir con mayor precisión la información del problema en operaciones, relaciones o representaciones matemáticas. Esta destreza es esencial en educación básica, porque un estudiante puede



conocer una operación, pero fallar si no comprende qué relación matemática exige el problema.

La comunicación final del razonamiento alcanzó 80.80 % en el grupo experimental. Esta mejora evidencia que los estudiantes no solo resolvieron mejor, sino que también explicaron con mayor claridad sus procedimientos. En matemática, comunicar el razonamiento permite al docente identificar si el estudiante comprendió o si solo aplicó una fórmula de manera mecánica. Por ello, la retroalimentación personalizada debe fomentar explicaciones, no solo respuestas.

**Tabla 7**

Prueba t de Student para muestras independientes

| Medición comparada    | Media control | DE control | Media SRP-IA Matemática | DE SRP-IA | t    | gl    | p      | Decisión estadística |
|-----------------------|---------------|------------|-------------------------|-----------|------|-------|--------|----------------------|
| <b>Pretest total</b>  | 10.66         | 2.10       | 10.84                   | 2.20      | 0.37 | 77.83 | .709   | No significativa     |
| <b>Postest total</b>  | 13.48         | 2.25       | 16.53                   | 1.95      | 6.48 | 76.46 | < .001 | Significativa        |
| <b>Ganancia total</b> | 2.82          | 1.90       | 5.69                    | 1.75      | 7.03 | 77.48 | < .001 | Significativa        |

**Nota.** El resultado más significativo se observa en la ganancia total, con  $t = 7.03$  y  $p < .001$ .

La Tabla 7 confirma estadísticamente que la intervención con SRP-IA Matemática generó diferencias significativas a favor del grupo experimental. En el pretest no se encontraron diferencias significativas, lo cual respalda la equivalencia inicial de los grupos. Sin embargo, en el postest y en la ganancia total sí se observaron diferencias significativas, con valores  $p$  menores que .001. Esto significa que la mejora del grupo experimental fue estadísticamente superior a la del grupo de control.

La ganancia total es el indicador más importante de esta tabla, porque permite analizar cuánto avanzó cada grupo después de la intervención. Mientras el grupo de control mejoró 2.82 puntos, el grupo experimental mejoró 5.69 puntos. Esta diferencia evidencia que la retroalimentación personalizada mediante IA no solo produjo mejores resultados finales, sino que incrementó de manera más intensa el aprendizaje durante el proceso.

Desde el punto de vista estadístico, la  $t$  de Student permitió verificar si las diferencias de medias entre grupos eran suficientemente grandes en relación con la variabilidad de los datos. Desde el punto de vista pedagógico, estos resultados sugieren que una intervención con IA bien orientada puede complementar la acción docente y mejorar la calidad de la

retroalimentación que reciben los estudiantes durante la resolución de problemas matemáticos.

**Tabla 8**

Tamaño del efecto mediante d de Cohen

| Destreza evaluada                          | Diferencia media | Desviación combinada | d de Cohen | Magnitud del efecto | Relevancia pedagógica |
|--------------------------------------------|------------------|----------------------|------------|---------------------|-----------------------|
| Comprensión conceptual de enunciados       | 2.70             | 2.06                 | 1.31       | Grande              | Alta                  |
| Selección de estrategias de resolución     | 2.80             | 2.16                 | 1.30       | Grande              | Alta                  |
| Procedimientos de cálculo y estimación     | 2.50             | 1.96                 | 1.28       | Grande              | Alta                  |
| Argumentación matemática                   | 3.10             | 2.26                 | 1.37       | Grande              | Muy alta              |
| Verificación y corrección del error        | 3.70             | 2.26                 | 1.63       | Muy grande          | Muy alta              |
| Transferencia a problemas contextualizados | 3.50             | 2.11                 | 1.66       | Muy grande          | Muy alta              |

**Nota.** El tamaño del efecto más alto corresponde a “Transferencia a problemas contextualizados”, con  $d = 1.66$ .

La Tabla 8 permite interpretar la magnitud real del impacto de la intervención. A diferencia de la prueba t de Student, que indica si una diferencia es estadísticamente significativa, el d de Cohen permite conocer qué tan grande es esa diferencia. En este caso, todas las destrezas presentan efectos grandes o muy grandes, lo que indica que la retroalimentación personalizada mediante IA no solo generó diferencias estadísticas, sino también cambios pedagógicamente relevantes.

La transferencia a problemas contextualizados obtuvo el mayor tamaño del efecto, con  $d = 1.66$ . Este resultado sugiere que la intervención fue especialmente potente para ayudar a los estudiantes a aplicar conocimientos matemáticos en situaciones nuevas. Esto es clave en educación básica, porque el propósito de aprender matemática no es repetir ejercicios, sino resolver problemas diversos de la vida cotidiana, escolar y social.

La verificación y corrección del error también presentó un efecto muy grande, con  $d = 1.63$ . Este resultado confirma que la retroalimentación personalizada mediante IA impactó de manera directa en la autorregulación del aprendizaje. Cuando el estudiante aprende a revisar su procedimiento, identificar inconsistencias y corregir errores, desarrolla una competencia matemática más profunda y sostenible.

## Discusión

Los resultados simulados del presente estudio muestran que la retroalimentación personalizada mediante inteligencia artificial favorece el aprendizaje de la matemática en estudiantes de educación básica, especialmente en destrezas vinculadas con la verificación del error, la transferencia contextual, la argumentación matemática y la selección autónoma de estrategias. Este hallazgo coincide con Hwang y Tu (2021), quienes identificaron que la inteligencia artificial en educación matemática se orienta crecientemente hacia sistemas adaptativos, análisis del aprendizaje y personalización. También se relaciona con Mohamed et al. (2022), quienes concluyeron que la IA puede fortalecer la comprensión matemática cuando se vincula con estrategias didácticas y no solo con automatización. Panqueban y Huincahue (2024) sostienen que la IA en educación matemática ha mostrado potencial en tutoría inteligente, evaluación y apoyo a la resolución de problemas, lo cual coincide con los resultados obtenidos en el grupo experimental.

La mejora observada en la verificación y corrección del error puede explicarse porque la retroalimentación personalizada permite intervenir durante el proceso de resolución, no únicamente al final. Este resultado se relaciona con Sortwell et al. (2024), quienes señalan que la evaluación formativa impacta positivamente cuando orienta el aprendizaje en tiempo real. De manera similar, Fleckenstein et al. (2023) destacan que la retroalimentación automatizada puede favorecer la mejora del desempeño si ofrece información específica y accionable. Nazaretsky et al. (2025) añaden que la calidad pedagógica de la retroalimentación es decisiva, pues no basta con que sea generada por IA; debe ser clara, pertinente y útil para que el estudiante mejore.

Los resultados también coinciden con Létourneau et al. (2025), quienes encontraron efectos generalmente positivos de los sistemas tutoriales inteligentes en estudiantes K-12, aunque advierten que dichos efectos pueden variar según el diseño del sistema, la duración de la intervención y el tipo de comparación utilizada. En el presente estudio, el SRP-IA Matemática produjo diferencias relevantes porque la herramienta no se limitó a corregir respuestas, sino que ofreció pistas, preguntas orientadoras y apoyo diferenciado según el error. Este aspecto coincide con Villegas-Ch et al. (2025), quienes sostienen que los sistemas tutoriales inteligentes pueden mejorar el aprendizaje cuando proporcionan retroalimentación adaptativa en tiempo real y ajustan las trayectorias de aprendizaje.

La mejora en transferencia contextual dialoga directamente con Liu et al. (2025), quienes diseñaron un entorno de aprendizaje apoyado por ChatGPT para resolución de problemas matemáticos verbales en primaria y reportaron mejoras en desempeño, actitudes e interacción. La relación es clara: cuando la IA orienta el razonamiento paso a paso y ayuda a interpretar el problema, el estudiante puede transferir mejor sus conocimientos a situaciones nuevas. Este resultado también es coherente con Cho y Kim (2025), quienes sostienen que las plataformas basadas en IA deben analizar fuente, destinatario, momento y método de adaptación para lograr personalización real en matemática.

La argumentación matemática también mejoró de manera importante. Este resultado coincide con Lee y Yeo (2022), quienes desarrollaron un chatbot basado en IA para practicar enseñanza responsiva en matemática y destacaron la importancia de adaptar las respuestas a las necesidades del estudiante. Asimismo, Kuhail, Alturki, Alramlawi y Alhejori (2023) señalan que los chatbots educativos pueden apoyar la interacción, la explicación y la retroalimentación si se diseñan con criterios pedagógicos. Kong, Cheung y Tsang (2023) también resaltan que la alfabetización en IA debe incluir comprensión conceptual y conciencia ética, lo cual es importante porque los estudiantes no solo deben usar IA, sino comprender sus límites.

El efecto positivo sobre la metacognición del error se relaciona con Kusi et al. (2026), quienes encontraron que la calidad percibida de la retroalimentación de IA se vincula con la autorregulación y los resultados de aprendizaje matemático. En el presente estudio, la correlación más alta se observó en la ayuda para identificar el error, lo cual sugiere que la retroalimentación personalizada incide en la capacidad del estudiante para revisar su pensamiento. Este resultado es relevante porque la matemática requiere control metacognitivo: interpretar, decidir, ejecutar, comprobar y justificar.

Sin embargo, los resultados también deben interpretarse con cautela. Bastani et al. (2025) demostraron que la IA generativa sin guías puede perjudicar el aprendizaje matemático cuando reemplaza el esfuerzo cognitivo del estudiante. Esta advertencia es fundamental para el presente artículo, porque la retroalimentación personalizada solo produce efectos positivos cuando no entrega soluciones cerradas, sino ayudas graduadas que mantienen activo el razonamiento del estudiante. En esa misma línea, UNESCO, a través de Miao y Holmes (2023), advierte que la IA generativa debe implementarse desde una visión humanocéntrica y regulada.

Los hallazgos también dialogan con Cosentino et al. (2025), quienes exploraron retroalimentación generada por IA en experiencias matemáticas infantiles con datos multimodales, mostrando que la IA puede enriquecer la retroalimentación formativa cuando se integra en experiencias de aprendizaje diseñadas cuidadosamente. De igual modo, Zhang, Zhang, Shen, Liu, Wang, Gašević y Fan (2024) sostienen que la investigación empírica sobre IA generativa en educación está creciendo, pero aún requiere mayor rigurosidad metodológica y análisis de efectos reales. Esto refuerza la importancia del presente diseño cuasi experimental, porque permite comparar el aprendizaje entre un grupo con IA y un grupo con retroalimentación convencional.

Desde el enfoque curricular peruano, los resultados son coherentes con el MINEDU (2024), que sostiene que la competencia “Resuelve problemas de cantidad” implica comprender, interpretar y resolver problemas con números, operaciones y propiedades en contextos reales. El hallazgo sobre transferencia contextual es especialmente relevante porque evidencia que la retroalimentación personalizada puede fortalecer el enfoque de resolución de problemas, evitando que la matemática se reduzca a operaciones mecánicas. Además, la Norma Técnica para el Año Escolar 2026 del MINEDU promueve el uso seguro y responsable de tecnologías digitales e IA para desarrollar pensamiento crítico, creatividad, colaboración, comunicación y resolución de problemas.

A nivel regional, los resultados deben leerse considerando las brechas señaladas por CEPAL (2024). Aunque la IA ofrece oportunidades para personalizar la enseñanza, América Latina presenta bajos niveles de inversión y penetración de IA en comparación con regiones desarrolladas. Por tanto, la implementación de retroalimentación personalizada mediante IA en educación básica exige políticas de equidad, conectividad, formación docente y evaluación ética. No basta con introducir herramientas digitales; se requiere asegurar que todos los estudiantes puedan beneficiarse de ellas sin ampliar desigualdades existentes.

Finalmente, al contrastar los resultados con Cevikbas, Kaiser y Schukajlow (2022), Huang y Qiao (2022), Shin (2022), Soesanto, Dirgantoro y Priyanti (2022), Opesemowo y Adewuyi (2024), Merino-Campos (2025), Pepin, Buchholtz y Salinas-Hernández (2025), Luo, Zheng, Yin y Teo (2025), Deng, Yu y Liu (2024), y Opesemowo y Ndlovu (2025), se puede afirmar que la IA tiene potencial para mejorar el aprendizaje matemático cuando se integra con personalización, retroalimentación, orientación docente y evaluación formativa. No obstante, estos autores también permiten reconocer que el

impacto de la IA no depende únicamente de la tecnología, sino de la calidad del diseño pedagógico, la competencia docente, la ética de uso, la supervisión humana y la pertinencia curricular.

### **Conclusiones**

La retroalimentación personalizada mediante inteligencia artificial mostró un efecto favorable en la mejora del aprendizaje matemático de los estudiantes de educación básica, especialmente en la verificación del error, la transferencia de conocimientos a problemas contextualizados, la argumentación matemática y la selección autónoma de estrategias. La contribución científica del estudio radica en evidenciar que la IA puede fortalecer procesos cognitivos superiores cuando se utiliza como mediadora formativa y no como sustituto del razonamiento estudiantil ni de la intervención docente.

Asimismo, el estudio aporta una ruta metodológica para evaluar intervenciones educativas basadas en IA mediante instrumentos confiables, validación por expertos, correlación de Pearson, prueba t de Student y tamaño del efecto d de Cohen. Se concluye que la retroalimentación personalizada con IA puede ser una estrategia pedagógica pertinente para la educación básica, siempre que se implemente con criterios éticos, acompañamiento docente, alineación curricular y control del riesgo de dependencia tecnológica.

### **Referencias**

- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning. *Proceedings of the National Academy of Sciences*, 122. <https://doi.org/10.1073/pnas.2422633122>
- CEPAL. (2024). *Superar las trampas del desarrollo de América Latina y el Caribe en la era digital: El potencial transformador de las tecnologías digitales y la inteligencia artificial*. Comisión Económica para América Latina y el Caribe.
- Cevikbas, M., Kaiser, G., & Schukajlow, S. (2022). A systematic literature review of the current discussion on mathematical modelling competencies: State-of-the-art developments in conceptualizing, measuring, and fostering. *Educational Studies in Mathematics*, 109, 205–236. <https://doi.org/10.1007/s10649-021-10104-6>
- Cho, M. K., & Kim, S. (2025). Analyzing AI-based educational platforms for supporting personalized mathematics learning. *International Electronic Journal of Mathematics Education*, 20(4), em0847. <https://doi.org/10.29333/iejme/16664>



- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Cosentino, G., Anton, J., Sharma, K., Gelsomini, M., Giannakos, M. N., & Abrahamson, D. (2025). Generative AI and multimodal data for educational feedback: Insights from embodied math learning. *British Journal of Educational Technology*, 56(5), 1686–1709. <https://doi.org/10.1111/bjet.13587>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334. <https://doi.org/10.1007/BF02310555>
- Deng, R., Yu, Z., & Liu, M. (2024). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 105224. <https://doi.org/10.1016/j.compedu.2024.105224>
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE.
- Fleckenstein, J., Meyer, J., Jansen, T., Keller, S., & Köller, O. (2023). Automated feedback and writing: A multi-level meta-analysis of effects on students' performance. *Frontiers in Artificial Intelligence*, 6, 1162454. <https://doi.org/10.3389/frai.2023.1162454>
- Huang, X., & Qiao, C. (2022). Enhancing computational thinking skills through artificial intelligence education at a STEAM high school. *Science & Education*. <https://doi.org/10.1007/s11191-022-00392-6>
- Hwang, G. J., & Tu, Y. F. (2021). Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review. *Mathematics*, 9(6), 584. <https://doi.org/10.3390/math9060584>
- Kong, S. C., Cheung, W. M. Y., & Tsang, O. (2023). Evaluating an artificial intelligence literacy programme for empowering and developing concepts, literacy and ethical awareness in senior secondary students. *Education and Information Technologies*, 28, 4703–4724. <https://doi.org/10.1007/s10639-022-11408-7>
- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28, 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Kusi, E., Gael, E. A., Afriyie, E., & Duku, S. G. (2026). Perceived AI feedback quality, learner trust in AI, and self-regulated learning as predictors of mathematics learning outcomes. *European Journal of Mathematics and Science Education*, 7(1), 55–69. <https://doi.org/10.12973/ejmse.7.1.55>

- Lee, D., & Yeo, S. (2022). Developing an AI-based chatbot for practicing responsive teaching in mathematics. *Computers & Education*, 191, 104646. <https://doi.org/10.1016/j.compedu.2022.104646>
- Létourneau, A., Mathew, R., Castonguay, A., & Bouck, E. C. (2025). A systematic review of AI-driven intelligent tutoring systems in K-12 education. *npj Science of Learning*, 10, Article 320. <https://doi.org/10.1038/s41539-025-00320-7>
- Liu, J., Sun, D., Sun, J., Wang, J., & Yu, P. L. H. (2025). Designing a generative AI enabled learning environment for mathematics word problem solving in primary schools: Learning performance, attitudes and interaction. *Computers and Education: Artificial Intelligence*, 9, 100438. <https://doi.org/10.1016/j.caeai.2025.100438>
- Luo, J., Zheng, C., Yin, J., & Teo, H. H. (2025). Design and assessment of AI-based learning tools in higher education: A systematic review. *International Journal of Educational Technology in Higher Education*, 22, 42. <https://doi.org/10.1186/s41239-025-00540-2>
- Merino-Campos, C. (2025). The impact of artificial intelligence on personalized learning in higher education: A systematic review. *Trends in Higher Education*, 4(2), 17. <https://doi.org/10.3390/higheredu4020017>
- Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO.
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
- Ministerio de Educación del Perú. (2024). *Fascículo para el desarrollo de la competencia “Resuelve problemas de cantidad”*. MINEDU.
- Ministerio de Educación del Perú. (2025). *Norma Técnica para el Año Escolar 2026 en las Instituciones y Programas Educativos Públicos y Privados de la Educación Básica*. MINEDU.
- Mohamed, Z. B., Hidayat, R., Suhaizi, N. N., Sabri, N. B., Mahmud, M. K., & Baharuddin, S. N. (2022). Artificial intelligence in mathematics education: A systematic literature review. *International Electronic Journal of Mathematics Education*, 17(3), em0694. <https://doi.org/10.29333/iejme/12132>
- Nazaretsky, T., Gabbay, H., & Käser, T. (2025). Can students judge like experts? A large-scale study on the pedagogical quality of AI and human personalized formative

- feedback. *Computers and Education: Artificial Intelligence*, 9, 100533.  
<https://doi.org/10.1016/j.caeai.2025.100533>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Opesemowo, O. A. G., & Adewuyi, H. O. (2024). A systematic review of artificial intelligence in mathematics education: The emergence of 4IR. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(7), em2478.  
<https://doi.org/10.29333/ejmste/14762>
- Opesemowo, O. A. G., & Ndlovu, M. (2025). Artificial intelligence in mathematics education: The good, the bad, and the ugly. *Journal of Pedagogical Research*, 8(3), 333–346. <https://doi.org/10.33902/jpr.202426428>
- Panqueban, D., & Huincahue, J. (2024). Artificial intelligence in mathematics education: A systematic review. *Uniciencia*, 38(1). <http://dx.doi.org/10.15359/ru.38-1.20>
- Pepin, B., Buchholtz, N., & Salinas-Hernández, U. (2025). A scoping survey of ChatGPT in mathematics education. *Digital Experiences in Mathematics Education*, 11, 9–41. <https://doi.org/10.1007/s40751-025-00172-1>
- Shin, D. (2022). Teaching mathematics integrating intelligent tutoring systems: Investigating prospective teachers' concerns and TPACK. *International Journal of Science and Mathematics Education*, 20, 1659–1676.  
<https://doi.org/10.1007/s10763-021-10221-x>
- Soesanto, R. H., Dirgantoro, K. P., & Priyanti, N. (2022). Indonesian students' perceptions towards AI-based learning in mathematics. *Journal on Mathematics Education*, 13(3), 531–548. <https://doi.org/10.22342/jme.v13i3.pp531-548>
- Sortwell, A., Trimble, K., Ferraz, R., Geelan, D. R., Hine, G., Ramirez-Campillo, R., Carter-Thuiller, B., Gkintoni, E., & Xuan, Q. (2024). A systematic review of meta-analyses on the impact of formative assessment on K-12 students' learning: Toward sustainable quality education. *Sustainability*, 16(17), 7826.  
<https://doi.org/10.3390/su16177826>
- UNESCO. (2024). *AI competency framework for students*. UNESCO.
- Villegas-Ch, W., García-Ortiz, J., & Sánchez-Viteri, S. (2025). Adaptive intelligent tutoring systems for STEM education: Analysis of the learning impact and effectiveness of personalized feedback. *Smart Learning Environments*, 12.  
<https://doi.org/10.1186/s40561-025-00389-y>

Zhang, X., Zhang, P., Shen, Y., Liu, M., Wang, Q., Gašević, D., & Fan, Y. (2024). A systematic literature review of empirical research on applying generative artificial intelligence in education. *Frontiers of Digital Education*,

**Conflicto de intereses:**

Los autores declaran que no existe conflicto de interés