

Marlon Andrés Suscal Orellana; Jorge Luis Armijos Carrión

<https://doi.org/10.5281/zenodo.17560999>

ARTÍCULO DE INVESTIGACIÓN CIENCIAS DE LA EDUCACIÓN

**Eficacia de herramientas de inteligencia artificial en el desarrollo de
habilidades matemáticas en estudiantes de básica media**

***Effectiveness of artificial intelligence tools in developing mathematical skills in
secondary school students***

Marlon Andrés Suscal Orellana

marlonsuscalore@gmail.com

Departamento de Posgrado, Universidad Técnica de Machala, Machala, El Oro,
Ecuador

<https://orcid.org/0000-0002-4499-0662>

Jorge Luis Armijos Carrión

jlarmijos@utmachala.edu.ec

Departamento de Posgrado, Universidad Técnica de Machala, Machala, El Oro,
Ecuador

<https://orcid.org/0000-0003-0312-786X>

Recibido: 21/09/2025

Revisado: 27/09/2025

Aprobado: 01/10/2025

Publicado: 31/10/2025



RESUMEN

Se realizó un análisis sobre el empleo de las herramientas de inteligencia artificial (IA) en la enseñanza de matemáticas y su relevancia en las instituciones públicas; con este propósito se planteó el objetivo de explorar las evidencias sobre la eficacia de la IA en el desarrollo de las habilidades matemáticas en los estudiantes de básica media entre el periodo 2020-2025. En cuanto a la metodología aplicada, se utilizó una revisión sistemática de base de datos como Web of Science, Scopus, ERIC, SciELO. Posteriormente, se realizó la depuración de registros en Rayyan y una lectura crítica de los estudios, conformado por 35 artículos. Los resultados que se obtuvieron describen en conjunto tutorías inteligentes, asistentes conversacionales con modelos de lenguaje de gran escala (LLM), además, se incluyen herramientas de realidad aumentada y de reconocimiento manuscrito, otros estudios demostraron mejoras en el aprendizaje cuando se incluyeron apoyos generados por modelos de IA.

Descriptores: Inteligencia artificial, enseñanza de las matemáticas, educación secundaria.

ABSTRACT

An analysis was conducted on the use of artificial intelligence (AI) tools in mathematics teaching and their relevance in public institutions. To this end, the objective was to explore the evidence on the effectiveness of AI in developing mathematical skills in elementary school students between 2020 and 2025. Regarding the methodology applied, a systematic review of databases such as Web of Science, Scopus, ERIC, and SciELO was used. Subsequently, records were purged in Rayyan and the studies, consisting of 35 articles, were critically read. The results obtained jointly describe intelligent tutoring, conversational assistants with large-scale language models (LLM), and augmented reality and handwriting recognition tools. Other studies demonstrated improvements in learning when supports generated by AI models were included.

Descriptors: Artificial intelligence, mathematics teaching, secondary education.

INTRODUCCIÓN

Según la Organización para la Cooperación y el Desarrollo Económico (OECD, 2023), hay un descenso en el desempeño matemático de adolescentes, hecho que sitúa el fortalecimiento de habilidades numéricas y algebraicas como foco de interés en el ámbito educativo; en este contexto, el Programa para la Evaluación Internacional de Estudiantes (PISA) registró una caída inédita del rendimiento en matemáticas y un aumento de estudiantes con niveles bajos de este dominio. A la vez, la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO, 2023) documentó problemas en la integración tecnológica y, para ello, precisó guías prácticas, recomendaciones y orientaciones que indican cómo debe utilizarse la tecnología en la enseñanza para que se asegure un uso pedagógico responsable y seguro.

En la disciplina de educación, las herramientas de IA son entendidas como sistemas que automatizan procesos de representación del conocimiento, por medio de tutorías adaptativas, retroalimentación y análisis de aprendizaje (Miao y Holmes, 2023), mientras que el desarrollo de habilidades matemáticas comprende la capacidad de razonar, formular, emplear e interpretar matemáticas en contextos diversos, con niveles de competencia definidos (OECD, 2022). En este caso, este estudio se focaliza en estudiantes de básica media por su ubicación curricular como puente entre el dominio de habilidades numéricas iniciales y la preparación para la matemática más avanzada, de modo que representa una etapa receptiva al apoyo de herramientas tecnológicas.

En el plano mundial, PISA evidenció que el promedio en matemáticas fue de 472 puntos o nivel 2, considerado el nivel mínimo básico; de ellos, un cuarto de alumnos se situó en niveles bajos por debajo de 357 puntos, menor al nivel 1 donde hay dificultades para resolver incluso los problemas más simples. En los sistemas educativos asiáticos se evidenciaron mejores resultados, Singapur promedió 575 puntos o nivel 4, donde los niños son capaces de trabajar con modelos e interpretar resultados (OECD, 2022). Asimismo, la UNESCO (2022) mostró desafíos en el logro en matemática y diferencias entre países en los niveles de tercero a sexto grado. De este modo, la evidencia confirma brechas persistentes y, en ciertos sistemas, alertas



técnicas sobre estimación en matemáticas de 8.º grado por concentraciones en el segmento inferior (IEA, 2023b).

Como intervención para mejorar estos puntajes, investigaciones como Gbolade y Omoponle (2024) y Martin et al. (2024) describieron efectos positivos y variados de tutores inteligentes y plataformas adaptativas en logros matemáticos. Además, el Estudio Internacional de Tendencias en Matemáticas y Ciencias (TIMSS por sus siglas en inglés) de la Asociación Internacional para la Evaluación del Logro Educativo (IEA, 2023a) permitió observar tendencias positivas en 4º y 8º grado, con un centro de escala en 500 puntos y variaciones amplias entre regiones.

A nivel regional, el Instituto Mexicano para la Competitividad (IMCO, 2023), señala que 2 de cada 3 estudiantes no logran alcanzar un nivel básico en el área de matemáticas. De esta manera, en Perú, Ponce et al. (2024), registró mediante su estudio, puntajes entre 420 y 482, los cuales estaban por debajo del promedio de la OECD, con predominio de niveles de desempeño inicial. Asimismo, en países hispanohablantes, diferentes estudios como Corica et al. (2025), Sanabria y Regil (2024) evidenciaron que varias experiencias tuvieron resultados favorables con la utilización de chatbots y sistemas de recomendación sustentados en la inteligencia artificial, mostrando efectos variados sobre el razonamiento matemático. Respecto a esto, Bolaño y Duarte (2024) informaron que existe variabilidad metodológica sobre los estudios de IA en la educación general básica.

De acuerdo con el Instituto Nacional de Evaluación Educativa de Ecuador (INEVAL, 2023) en el subnivel Básica Medio, el promedio nacional en Matemática fue 675 puntos en 2022-2023 y 680 en 2023-2024, en escala de 0 a 1000. Además, en bachillerato el promedio de Matemática fue 696 puntos y el 51,8 % alcanzó el umbral mínimo de competencia (≥ 700); mientras que en básica elemental los promedios nacionales se situaron en torno a 678 puntos ese mismo periodo. Respecto a dicha evaluación, la UNESCO (2025) especificó que la cobertura muestral en el país alcanzó a 47.339 estudiantes de 1.187 instituciones.

En este orden de ideas, la evidencia reciente asocia al desempeño del estudiante con experiencias de cierre escolar y al aprendizaje virtual. También lo relaciona con la forma en que se usa la tecnología en clase, donde un uso moderado de recursos



digitales se asocia con mejores resultados, mientras que las distracciones tecnológicas se relacionan con menores logros. Aún existen brechas de acceso, de capacidades docentes y de gestión de recursos; por esto, para lograr utilizar estas herramientas dentro de las aulas de clases, estas condiciones recurren a la implementación de plataformas de inteligencia artificial en el aprendizaje al área de matemáticas.

Con base a lo expuesto, se formula la siguiente pregunta de investigación: ¿Qué evidencia existe sobre la eficacia de herramientas de inteligencia artificial en el desarrollo de habilidades matemáticas en estudiantes de básica media de instituciones públicas durante el periodo del 2020 a 2025?

El presente tema se eligió por su pertinencia académica y educativa ante la expansión de soluciones de IA, entre ellas sistemas generativos y tutores adaptativos, en escenarios escolares reales. Miao y Holmes (2023) en documentos de orientación internacional describen el alcance, los usos y precauciones para entornos de aprendizaje y evaluación. En la literatura se registran avances y vacíos, por un lado, se reportaron impactos positivos en el desempeño matemático con IA, aunque con heterogeneidad de diseños, variables y contextos (Martin et al., 2024; Son, 2024).

Para esto, se plantea el objetivo general: explorar la evidencia disponible sobre la eficacia de herramientas de inteligencia artificial en el desarrollo de habilidades matemáticas en estudiantes de básica media de instituciones públicas durante el periodo de 2020 a 2025.

MÉTODOS

Para el desarrollo de la revisión sistemática, se formuló la siguiente pregunta PICO: ¿En estudiantes de básica media de instituciones públicas, la aplicación de herramientas de IA orientadas al aprendizaje matemático, en comparación con métodos tradicionales de enseñanza, demuestra eficacia en el desarrollo de habilidades matemáticas durante el periodo 2020 a 2025?

Se definieron los criterios de inclusión: se consideraron estudiantes de grados equivalentes a 5º - 7º año de educación general básica subnivel medio, contextos escolares públicos, intervenciones con tutores, IA, chatbots y analítica; también



resultados vinculados a logro o habilidades matemáticas, diseños empíricos, idioma español, inglés o portugués; además de disponibilidad de texto completo con periodo de publicación entre 2020 a 2025.

Se omitieron textos repetidos, incompletos o no legibles, donde no se apreció la presencia de uso de IA; también se excluyeron estudios con pequeños grupos; igual, se descartaron investigaciones de institutos privados o de nivel educativo diferente al subnivel medio de educación básica; por último, se dejaron de lado documentos no pertinentes con matemáticas, resúmenes o textos sin número y que no eran de las fechas proporcionadas.

Para identificar la evidencia, se buscó en Web of Science Core Collection, *Scopus/Science Direct* (Elsevier), *Education Resources Information Center* (ERIC), *Scientific Electronic Library Online* (SciELO) y Redalyc; se buscó literatura gris en sitios de tesis y reportes de instituciones. La primera búsqueda se realizó el 25 de agosto de 2025 y la última revisión actualización el 5 de septiembre de 2025.

Luego, los registros se descargaron con RIS y se organizaron en Rayyan, donde se les hizo filtrado, deduplicación y etiquetado con diseño, idioma, año de publicación y contexto escolar. Las palabras “*Artificial intelligence*”, “*Mathematics education*”, “*Secondary education*”, y “*Public schools*”, y sus variantes en español y portugués fueron seleccionadas según el tesauro de la UNESCO. La búsqueda se realizó mediante la combinación de “AND”, “OR” y “NOT”, el uso de comillas, en base a los criterios de elegibilidad. También se unificaron y pilotaron cadenas por base de datos. Ejemplo de una cadena de búsqueda implementada:

ERIC: ((DE “*Artificial Intelligence*” OR “*intelligent tutoring systems*” OR *chatbot** OR “*adaptive learning*” OR “*learning analytics*”) AND (DE “*Mathematics Achievement*” OR *Mathematics* OR *Algebra*) AND ((DE “*Middle Schools*” OR “*Junior High Schools*” OR “*Lower Secondary Education*” OR “*grade 5*” OR “*grade 6*” OR “*grade 7*”) AND (DE “*Public Schools*”))) *Peer reviewed*; *Publication Date*: 2020-2021; *Language*: *English* OR *Spanish* OR *Portuguese*; NOT “*Higher Education*”;

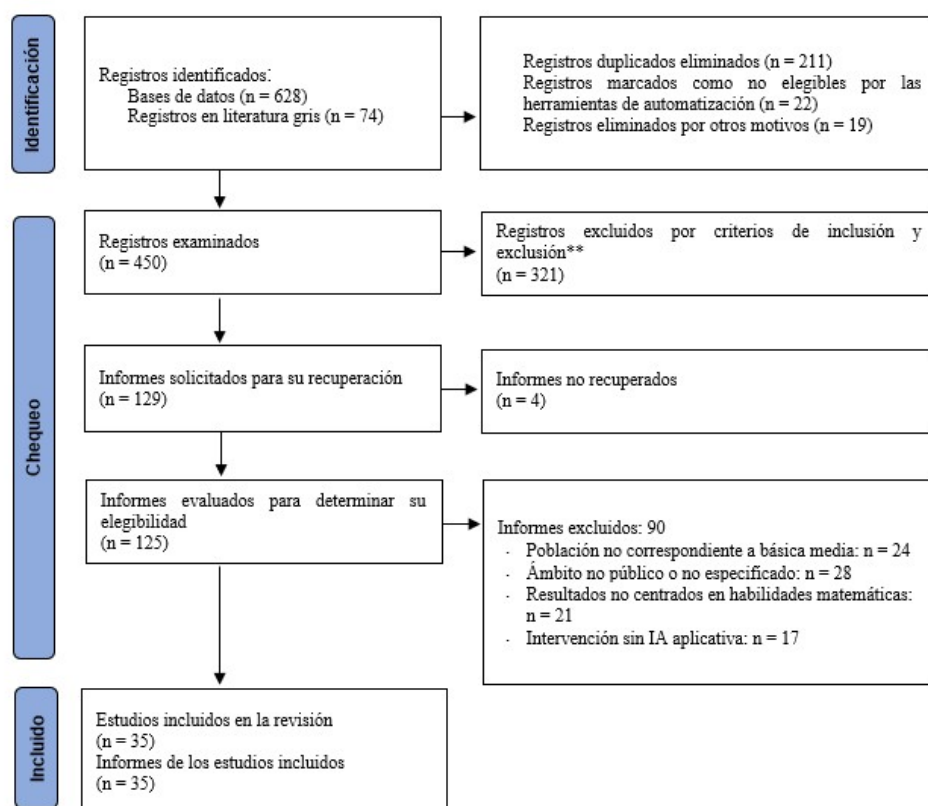
En cuanto al proceso de selección se desarrolló en dos fases, el cribado de títulos y resúmenes, seguido de lectura a texto completo. El investigador y el tutor del programa de maestría evaluaron de manera independiente los estudios; las discrepancias se



Marlon Andrés Suscal Orellana; Jorge Luis Armijos Carrión

resolvieron por consenso entre ambos. El reporte adoptó el método PRISMA 2020 (Figura 1), dejando en la síntesis final 35 artículos que respondieron a la pregunta PICO y cumplieron los filtros definidos, de acuerdo con el flujo cuantificado.

Figura 1. Diagrama de flujo PRISMA.



Por otra parte, para la extracción de los datos se construyó una matriz con los campos como: N°, autor, año y país, título, tipo de estudio/metodología, población y muestra, principales resultados, herramienta de IA, habilidad matemática evaluada, indicadores de eficacia, entre otros. Se realizó una matriz adicional, que fue resultado de una consolidación de los artículos analizados con los criterios de la base de datos, la cadena de búsqueda y los filtros aplicados, fue por el registro de los artículos recuperados y evaluados, que se excluyeron y se incluyeron.

A partir del registro de los estudios recuperados, excluidos e incluidos, se organizó la síntesis de acuerdo con el diseño, herramienta, la habilidad que se trabajó y el nivel educativo del país, con el fin de señalar lo investigado y que procedimientos se utilizaron. Se realizó un análisis bibliométrico que calculó frecuencias anuales, exploró

la concurrencia de los indicadores y la distribución por revistas y regiones. En este sentido, se elaboró una síntesis narrativa que contuvo comparabilidad entre los indicadores de eficiencia, se desarrolló una discusión temática y se organizaron los resultados en parámetros.

Para la estimación de consistencia interna, calidad de evidencia y los sesgos se aplicó el programa de habilidades de evaluación crítica (CASP). Se registraron los ítems en formato sí, no, no aplica y las reglas de decisión que se establecieron por ítems que se calcularon se cumplieron y se emitió un juicio global. A la par, las revisiones añadidas se juzgaron con *A Measurement Tool to Assess Systematic Reviews 2* (AMSTAR 2). A su vez, la gradación de la evidencia (GRADE) se aplicó con niveles alta, moderada, baja o muy baja.

RESULTADOS

Análisis bibliométrico

Los resultados se mostraron por medio de tablas de resumen y planos de prueba. Después, la primera tabla juntó rasgos de cada estudio y nivel de la evidencia; además, otra guía anotó, por base de datos, los criterios para la elección de la evidencia y los números resultantes según la aplicación de cada filtro. Adicional, se realizó una descripción de la evidencia según la zona o región de donde proviene la evidencia.

Dado el carácter de las fuentes vistas, la revisión usó solo fuentes de acceso abierto; por eso, no hizo falta permiso de un comité de ética. Igualmente, se sostuvieron normas de claridad, rigor técnico y mención total. Por tanto, cada elección analítica quedó registrada. Además, se atendieron permisos de entrada libre y derechos de autor; sumado a esto, se avisaron posibles conflictos de interés y se guardaron los recursos de búsqueda junto con diagramas para futuras comprobaciones.

En términos de distribución geográfica, el corpus de estudios se concentró en Norteamérica (13 estudios) y Europa (12), seguidos por Asia (7) y América Latina (3). En Norteamérica, Europa y Asia, la calidad se valoró como “buena” por el peso de estudios observacionales y predictivos; en cambio, en América Latina fue “alta” por la presencia de diseños experimentales y revisiones con criterios explícitos. Por su parte,



la producción norteamericana se apoyó en series a gran escala de plataformas con tutoría inteligente y analítica, entre ellas, *ASSISTments* y *ALEKS* (Spitzer et al., 2024; Drijvers et al., 2021; Matayoshi et al., 2025).

En términos temporales, el mayor volumen de publicaciones recientes se concentra entre 2023 y 2025. En el plano editorial, destacan *Journal of Educational Data Mining*, con artículos que utilizaron pruebas de predicción y pruebas en plataformas; de igual manera, *Computers & Education Open*, con evaluaciones longitudinales y *Journal of Computer Assisted Learning*, con trabajos metodológicos y de minería de aprendizaje (Vanacore et al., 2024; Spitzer et al., 2024; Létourneau et al., 2025; Hermann Spitzer y Musslick, 2021; Hermann Spitzer y Moeller, 2024).

Respecto a los esquemas de investigación, el análisis, junta cinco pruebas o test con reparto al azar (Pardos y Bhandari, 2024; de-Morais y Jaques, 2022; Uriarte Portillo et al., 2023; Lim et al., 2025; Matayoshi et al., 2025), cerca de ocho análisis casi experimentales o espontáneos (Shih et al., 2023; Mavrikis et al., 2022; Tomasik et al., 2020), cinco resúmenes con metaanálisis, repasos sistemáticos y de cobertura; además de un grupo mayor de análisis contemplativos y adivinatorios basados en puentes de uso.

Síntesis narrativa

Diversidad de herramientas

La evidencia acumulada converge en que la tutoría inteligente y la práctica adaptativa mantienen o incrementan el desempeño en matemáticas en escenarios de aula y de enseñanza remota, con resultados consistentes en registros longitudinales de *bettermarks* y comparaciones antes-durante cierres escolares (Spitzer et al., 2024; Hermann Spitzer y Moeller, 2023). Mientras que metaanálisis y revisiones recientes confirman efectos positivos para plataformas adaptativas como *ALEKS* y para diversos tutores virtuales en K-12, si bien con heterogeneidad por diseño y duración (Sun et al., 2021; Létourneau et al., 2025).

Impacto en habilidades

En cuanto a las intervenciones controladas, surgen señales diferenciadas, la ayuda generada por un modelo de lenguaje en un tutor abierto (*OATutor*) produjo ganancias significativas frente al control y equivalentes a la ayuda redactada por tutores humanos



Marlon Andrés Suscal Orellana; Jorge Luis Armijos Carrión

(Pardos y Bhandari, 2024). En experimentos utilizados una plataforma, el umbral de dominio modula la retención a corto plazo con una posterior atenuación y el uso de informes sobre tareas por parte del profesor, estimado utilizando un diseño de alimentos aleatorios, incrementan la finalización de la siguiente tarea. En la educación secundaria, la realidad aumentada con un tutor utilizando lógica difusa mostró mayores ganancias que la RA sin tutoría (Uriarte Portillo et al., 2023).

Factores contextuales

Dada la configuración de pistas y retroalimentación que afecta a estudiantes con tendencia a *gaming the system* y la demora de ayudas puede mejorar resultados en ese subgrupo (Vanacore et al., 2024). La modalidad de entrada influye en la fluidez y la cantidad de ejercicios resueltos; la escritura manuscrita dentro del tutor aceleró la resolución, aunque sin diferencias en ganancia inmediata (de-Morais y Jaques, 2022). En Europa, las investigaciones sobre enseñanza remota detallan condiciones de implementación y barreras de acceso que condicionan la lectura de resultados en sistemas adaptativos (Drijvers et al., 2021). En consecuencia, la tabla 1 facilita lectura rápida:

Tabla 1. Síntesis por ejes analíticos.

Eje	Número de artículos	Herramientas/Enfoques dominantes	Autores citados
Diversidad de herramientas	12	Tutoría por pasos, chat generativo integrado, RA con lógica difusa, reconocimiento manuscrito, rastreo de conocimientos	Levonian et al. (2025); Uriarte-Portillo et al. (2023); de Moraes y Jaques (2022); Lu et al. (2024)
Impacto en habilidades	14	STI/adaptativos para fracciones/álgebra; experimentos y comparaciones longitudinales	Pardos et al. (2024); Spitzer et al. (2023, 2024); Sun et al. (2021)
Factores contextuales	9	Configuración de pistas/feedback, uso docente de reportes, umbrales de maestría, condiciones de acceso	Vanacore et al. (2024); Lim et al. (2025); Matayoshi et al. (2025); Drijvers et al. (2021)

En la tabla 2, los Sistemas Tutores Inteligentes (STI) concentran la mayor evidencia, con mejoras en precisión, completitud y retención de corto plazo y calificaciones altas o buenas de la evidencia. Asimismo, la Realidad Aumentada con reglas (RA) muestra ventajas en geometría y las plataformas adaptativas presentan progresos cuando la



Marlon Andrés Suscal Orellana; Jorge Luis Armijos Carrión

práctica se prolonga, mientras el chat con modelo de lenguaje (LLM) ofrece apoyos comparables a tutoría humana.

Tabla 2. Síntesis de herramienta de IA y la habilidad de matemática trabajada.

Herramienta de IA	Habilidad matemática trabajada	Principales resultados y nivel de evidencia
Sistema Tutor Inteligente (STI) (n=18)	Fracciones, álgebra, aritmética y geometría; resolución paso a paso	Mejora de desempeño y precisión; aumentos de completitud; efectos en retención de corto plazo; sensibilidad a la duración. GRADE: alta 3; buena 15.
Realidad Aumentada (RA) con tutor inteligente (n=7)	Geometría y álgebra con soporte visual y reglas	Ventaja frente a versiones sin tutor; incrementos de precisión y logro; efectos consistentes en tareas espaciales. GRADE: alta 3; buena 4.
Plataforma adaptativa (n=4)	Cálculo básico y contenidos curriculares de primaria/secundaria	Tendencias de mejora en rendimiento y seguimiento; ausencia de brechas en periodos largos; efectos ligados a práctica sostenida. GRADE: alta 2; buena 2.
Sin herramienta de IA declarada (n=2)	Funciones, representación y visualización; aritmética simbólica	Mejoras vinculadas a decisiones de diseño instruccional en línea y RA educativa; sensibilidad a la duración. GRADE: buena 2.
Analítica de autorregulación del aprendizaje (SRL) con aprendizaje automático (ML) (n=1)	Resolución de problemas y estrategias	Variaciones por subgrupos; efectos relacionados con tiempo de exposición. GRADE: buena 1.
Chat conversacional con modelo de lenguaje (LLM) (n=1)	Procedimientos de álgebra y elementos de estadística	Soporte comparable a asistencia humana en tareas procedimentales; efectos dependientes del tiempo de uso. GRADE: buena 1.
Sistema adaptativo con retroalimentación formativa (n=1)	Competencias curriculares en plataforma	Incrementos en crecimiento de rendimiento y precisión dentro del sistema. GRADE: buena 1.
Otro (n=1)	Resolución de ejercicios en enseñanza remota	Incrementos de desempeño con seguimiento analítico; variaciones por grupo. GRADE: buena 1.

Nota. n= cantidad de artículos que integra cada temática



DISCUSIÓN

El análisis revela una gama extensa de ideas con IA en el mundo de las matemáticas escolares: plataformas de enseñanza con tutorización inteligente o STI, interacciones conversacionales con modelos de lenguaje, realidad aumentada (RA) con normas un poco confusas, lectura de escritura a mano e ideas conectadas con métodos de atención. Los apoyos creados por modelos de lenguaje y la diálogo unificado enseñaron que son útiles y confiables con sistemas de protección y búsqueda de datos; por eso se siguió usándolos sin parar en salones de clases de verdad (Pardos y Bhandari, 2024; Levonian et al., 2025).

Por su parte, los Sistemas Tutoriales Inteligentes (STI) y las plataformas adaptativas reunieron la mayor cantidad de estudios; además, se utilizaron *bettermarks*, *ASSISTments*, *ALEKS* y *HINTS*, junto con combinaciones de Entornos de Aprendizaje Exploratorio (ELE) que integraron práctica guiada con actividades de indagación. Respectos a los temas, sobresalieron fracciones, álgebra y porcentajes, tocando un poco los aspectos de representación y estructuración (Hermann Spitzer y Moeller, 2022; Mavrikis et al., 2022; del-Olmo Muñoz et al., 2023). Asimismo, la RA con reglas difusas reveló avances versus modos sin guía (Uriarte Portillo et al., 2023).

Además, los estudios largos usando datos como *bettermarks* dejaron ver más aciertos y menos fallos durante la educación a distancia manteniéndose en evidencia la dificultad de la profesión y las responsabilidades que integra. Asimismo, los resultados se replicaron en varios países europeos y latinoamericanos, con patrones de mejora en estudiantes con rendimiento previo más bajo y convergencia en segmentos intermedios (Hermann Spitzer y Moeller, 2023; Hermann Spitzer y Moeller, 2024).

A continuación, los ensayos controlados muestran efectos medibles en resultados académicos, la ayuda generada por modelos de lenguaje en *OATutor* produjo ganancias en pruebas frente al control y equivalencia con ayuda humana, con reducción de errores mediante verificación automática. De forma complementaria, la RA con un tutor de reglas difusas incrementó la ganancia en geometría y la motivación, mientras un STI mejoró el desempeño en fracciones (Pardos y Bhandari, 2024; Uriarte Portillo et al., 2023; Shih et al., 2023).



Adicional, en los resultados diversos elementos del contexto de uso condicionaron la eficacia de las herramientas. A la vez, el ritmo y el momento de las ayudas alteraron el rendimiento de quienes tienden a jugar con el sistema, mientras el diferimiento de pistas mostró ventajas frente a la entrega inmediata y ciertas variantes de gamificación resultaron neutras o desfavorables según la configuración (Vanacore et al., 2024). Por último, el incentivo aleatorizado para que los profesores consulten reportes incrementó la probabilidad de completar la siguiente tarea a estudiantes y los umbrales de maestría afectaron la retención temprana con atenuación posterior (Lim et al., 2025; Matayoshi et al., 2025; Quintos Almeda y Shaun Baker, 2020).

También se observaron efectos vinculados con la modalidad de interacción y las condiciones del entorno. De igual modo, la escritura manuscrita dentro del tutor aceleró la resolución y aumentó el número de ejercicios sin cambios claros en la ganancia inmediata, mientras la enseñanza remota expuso barreras de acceso y variaciones en apoyo institucional que condicionaron el uso de plataformas (de-Morais y Jaques, 2022; Drijvers et al., 2021).

En cuanto a las diferencias halladas, se deben a cómo se actúa, al modo de entrar datos y al uso de informes por maestros; además, escribir a mano hizo más veloz sin ventajas claras, los informes hicieron más extenso y el nivel de saber incluyó en guardar memoria, con patrones estables por tipo de ítem u orden de presentación (de-Morais y Jaques, 2022; Lim et al., 2025; Matayoshi et al., 2025; Rohani et al., 2024; Lu et al., 2024). De igual manera, respecto a los parecidos, los observacionales usaron métricas internas y los experimentos pre-post externos; mayor duración rindió mejores progresos, con impacto docente diferido y condicionantes de conectividad en Europa (Sun et al., 2021; Hermann Spitzer y Moeller, 2023; Drijvers et al., 2021).

En las implicaciones prácticas, la variedad de soluciones guía decisiones de diseño, como el diálogo con modelos de lenguaje en tareas procedimentales, tutoría por pasos para descomponer operaciones y realidad aumentada con reglas para visualización. Cabe estudiar la integración de IA en contenidos complejos como aprendizaje de fracciones y álgebra.

En el ámbito del currículo, las indicaciones del tutor indican progresiones en cadena y efecto sobre la retención a corto plazo, con consecuencias en la secuenciación



curricular y en las tareas a corto plazo. En lo que respecta a las repercusiones teóricas y políticas, es necesario contar con criterios analíticos, documentación de flujos y filtros que aseguren la trazabilidad, incluyendo escritura a mano cuando sea pertinente.

CONCLUSIONES

La evidencia respecto a las habilidades matemáticas en estudiantes de básica media converge en un conjunto de factores que acompañan la eficacia de las herramientas, como el diseño y temporización de ayudas, personalización adaptativa, secuencias de contenido coherentes con dependencias entre habilidades, monitoreo docente mediante reportes, dificultad graduada de los ítems y duración suficiente de la práctica. Asimismo, la interacción con el sistema influye a través de la modalidad de entrada, la forma de retroalimentación y la disponibilidad de recursos, circunstancias que condicionan la participación, fluidez de resolución y continuidad del trabajo, además de la calidad técnica del despliegue y de las políticas de acceso escolar.

Las investigaciones articularon trayectorias metodológicas complementarias que incluyeron ensayos y cuasiexperimentos con pruebas antes y después; además, llevaron a cabo seguimientos a largo plazo en escenarios reales y análisis observacionales a gran escala sobre registros de plataforma; igualmente, integraron diseños naturalistas y esquemas de asignación por aliento para estimar el impacto del uso del docente de informes de forma simultánea, se desarrollaron modelos para trazar el conocimiento y clasificar eventos con validación de igual manera se aplicaron mapas de relación entre subtemas mediante el análisis de redes; finalmente, las revisiones sistemáticas llevaron a cabo búsquedas claras con cribado independientemente; con esto, quedaron delineadas tendencias y condiciones para la implementación en la educación matemática.

En conjunto, las rutas de intervención se organizan en tres líneas complementarias: tutoría gradual con práctica adaptativa y umbrales de dominio; además, chat conversacional con verificación de respuesta. También, aprendizaje exploratorio articulado con ejercicios estructurados, de modo que combinan explicación, práctica escalonada e indagación guiada por recursos digitales. En esa línea se incorporan



variantes que refinan la interacción y el seguimiento: realidad aumentada (RA) con reglas difusas para la visualización geométrica; igual que el reconocimiento de escritura para procedimientos algebraicos; por su lado, paneles docentes con indicadores de control del tiempo de las pistas y ajustes de gamificaciones destinados a reducir usos improductivos durante la resolución.

Para concluir, existen vacíos de datos en África y Oceanía, además de una escasa presencia de sistemas públicos en América Latina; al mismo tiempo, se pueden notar descripciones incompletas de características institucionales y de procesos de formación docente relacionados con el uso de esta herramienta; a su vez algunos estudios se basan en métricas internas con intervenciones breves y con escasa conexión en comparación con pruebas externas. En consecuencia, la evidencia existente aún no cuenta con investigaciones multicéntricas que utilicen medidas comparables y horizontales temporales amplios; se necesitan análisis sobre la equidad y la seguridad de los chats en el aula, así como evaluaciones de costos y capacidades; de manera similar, siguen surgiendo preguntas sobre la continuidad, la transferencia y la priorización de contenidos en el subnivel medio.

REFERENCIAS

- Bolaño García, M., & Duarte Acosta, N. (2024). Una revisión sistemática del uso de la inteligencia artificial en la educación. *Revista Colombiana de Cirugía*, 39(1), 51-63. <https://doi.org/10.30944/20117582.2365>
- Corica, A., Sureda, P., Parra, V., Schiaffino, S., & Godoy, D. (2025). Educación e inteligencia artificial: desempeño de chatbots y profesores de matemática en la resolución de problemas geométricos. *Areté, Revista Digital del Doctorado en Educación*, 10(1). <https://doi.org/10.55560/arete.2024.ee.10.9>
- del-Olmo Muñoz, J., González Calero, J., Diago, P., Arnau, D., & Arevalillo Herráez, M. (2023). Intelligent tutoring systems for word problem solving in COVID-19 days: could they have been (part of) the solution? *ZDM – Mathematics Education*, 55(1), 35-48. <https://doi.org/10.1007/s11858-022-01396-w>



- de-Morais, F., & Jaques, P. (2022). Does handwriting impact learning on math tutoring systems? *Informatics in Education*, 21(1), 55-90. <https://doi.org/10.15388/infedu.2022.03>
- Drijvers, P., Thurm, D., Vandervieren, E., Klinger, M., Moons, F., van-der-Ree, H., Doorman, M. (2021). Distance mathematics teaching in Flanders, Germany, and the Netherlands during COVID-19 lockdown. *Educational Studies in Mathematics*, 108(1), 35-64. <https://doi.org/10.1007/s10649-021-10094-5>
- Gbolade Opesemowo, O., & Omoponle Adewuyi, H. (2024). A systematic review of artificial intelligence in mathematics education: The emergence of 4IR. *EURASIA Journal of Mathematics, Science and Technology Education*, 20(7), 1-11. <https://doi.org/10.29333/ejmste/14762>
- Hermann Spitzer, M., & Moeller, K. (2022). Predicting fraction and algebra achievements online: A large-scale longitudinal study using data from an online learning environment. *Journal of Computer Assisted Learning*, 38(6), 1797-1806. <https://doi.org/10.1111/jcal.12721>
- Hermann Spitzer, M., & Moeller, K. (2023). Performance increases in mathematics during COVID-19 pandemic distance learning in Austria: Evidence from an intelligent tutoring system for mathematics. *Trends Neurosci Educ*, 31(100203). <https://doi.org/10.1016/j.tine.2023.100203>
- Hermann Spitzer, M., & Moeller, K. (2024). Performance increases in mathematics within an intelligent tutoring system during COVID-19 related school closures: a large-scale longitudinal evaluation. *Computers and Education Open*, 6(1). <https://doi.org/10.1016/j.caeo.2024.100162>
- Hermann Spitzer, M., & Musslick, S. (2021). Academic performance of K-12 students in an online-learning environment for mathematics increased during the shutdown of schools in wake of the COVID-19 pandemic. *PLOS One*, 16(8). <https://doi.org/10.1371/journal.pone.0255629>
- IEA. (2023a). *Trends in International Mathematics and Science Study*. Asociación Internacional para la Evaluación del Logro Educativo: <https://n9.cl/gldhu>
- IEA. (2023b). *South Africa: Highlights Report*. Department of Basic Education. <https://n9.cl/dwcyi>



- IMCO. (2023). *PISA 2022: Dos de cada tres estudiantes en México no alcanzan el nivel básico de aprendizajes en Matemáticas*. Instituto Mexicano para la Competitividad A.C. <https://n9.cl/baw3s>
- INEVAL. (2023). *Informe Nacional de Resultados. Ser Estudiante. Subnivel Básica Media*. Ministerio de Educación del Ecuador. Instituto Nacional de Evaluación Educativa. <https://n9.cl/n3v8hk>
- Létourneau, A., Deslandes Martineau, M., Charland, P., Alexander Karran, J., Boasen, J., & Majorique Léger, P. (2025). A systematic review of AI-driven intelligent tutoring systems (ITS) in K-12 education. *npj Science of Learning*, 10(29). <https://doi.org/10.1038/s41539-025-00320-7>
- Levonian, Z., Henkel, O., Li, C., & Postle, M. (2025). Designing Safe and Relevant Generative Chats for Math Learning in Intelligent Tutoring Systems. *Journal of Educational Data Mining*, 17(1), 66-97. <https://doi.org/10.5281/zenodo.14751365>
- Lim, W., Heffernan, N., & Sales, A. (2025). Evaluating the Effects of Assignment Report Usage on Student Outcomes in an Intelligent Tutoring System: A Randomized-Encouragement Design. *Journal of Educational Data Mining*, 17(1), 248-275. <https://doi.org/10.5281/zenodo.15366697>
- Lu, Y., Tong, L., & Cheng, Y. (2024). Advanced Knowledge Tracing: Incorporating Process Data and Curricula Information via an Attention-Based Framework for Accuracy and Interpretability. *Journal of Educational Data Mining*, 16(2), 58-84. <https://doi.org/10.5281/zenodo.13712553>
- Martin, F., Zhuang, M., & Schaefer, D. (2024). Systematic review of research on artificial intelligence in K-12 education (2017–2022). *Computers and Education: Artificial Intelligence*, 6(1). <https://doi.org/10.1016/j.caeai.2023.100195>
- Matayoshi, J., Cosyn, E., Uzun, H., & Kurd Misto, E. (2025). Using a Randomized Experiment to Compare Mastery Learning Thresholds. *Journal of Educational Data Mining*, 17(1), 308-336. <https://doi.org/10.5281/zenodo.15698758>
- Mavrikis, M., Rummel, N., Wiedmann, M., Loibl, K., & Holmes, W. (2022). Combining exploratory learning with structured practice educational technologies to foster both conceptual and procedural fractions knowledge. *Educational technology*

- research and development*, 70(1), 691-412. <https://doi.org/10.1007/s11423-022-10104-0>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO (United Nations Educational, Scientific and Cultural Organization). <https://doi.org/10.54675/EWZM9535>
- OECD. (2022). *Description of the PISA mathematics literacy proficiency levels: 2022*. Organization for Economic Cooperation and Development. <https://n9.cl/v7qx6>
- OECD. (2022). *Singapore Student performance (PISA)*. Organización para la Cooperación y el Desarrollo Económicos. <https://gpseducation.oecd.org/CountryProfile?primaryCountry=SGP&topic=PI>
- OECD. (2023). *PISA 2022 Results The State of Learning and Equity in Education*. Organización para la Cooperación y el Desarrollo Económicos. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Pardos, Z., & Bhandari, S. (2024). ChatGPT-generated help produces learning gains equivalent to human tutor-authored help on mathematics skills. *PLOS One*, 19(5): e0304013. <https://doi.org/10.1371/journal.pone.0304013>
- Ponce Vértiz, M., Altabás Kajatt, F., Cuadros Espinoza, M., Ñañez Aldaz, D., & Vera Tudela, D. (2024). *El Perú en PISA. Informe nacional de resultados*. Ministerio de Educación del Perú. <https://n9.cl/r5p5j1>
- Quintos Almeda, M., & Shaun Baker, R. (2020). Predicting Student Participation in STEM Careers: The Role of Affect and Engagement during Middle School. *Journal of Educational Data Mining*, 12(2). <https://doi.org/10.5281/zenodo.4008053>
- Rohani, N., Rohani, B., & Manataki, A. (2024). ClickTree: A Tree-based Method for Predicting Math Students' Performance Based on Clickstream Data. *Journal of Educational Data Mining*, 16(2), 32-57. <https://doi.org/10.5281/zenodo.13627655>
- Sanabria Medina, G., & Regil Vargas, L. (2024). Inteligencia artificial para la recomendación de recursos en educación en línea. *Apertura (Guadalajara, Jal.)*, 16(2). <https://doi.org/10.32870/ap.v16n2.2542>

- Shih, S., Chang, C., Kuo, B., & Huang, Y. (2023). Mathematics intelligent tutoring system for learning multiplication and division of fractions based on diagnostic teaching. *Education and Information Technologies*, 28(1), 9189-9210. <https://doi.org/10.1007/s10639-022-11553-z>
- Son, T. (2024). Intelligent Tutoring Systems in Mathematics Education: A Systematic Literature Review Using the Substitution, Augmentation, Modification, Redefinition Model. *Computers*, 13(10), 270. <https://doi.org/10.3390/computers13100270>
- Spitzer, M., Bardach, L., Strittmatter, Y., Meyer, J., & Moeller, K. (2024). Evaluating the content structure of intelligent tutor systems—A psychological network analysis. *Computers and Education Open*, 7(1). <https://doi.org/10.1016/j.caeo.2024.100198>
- Sun, S., Else Quest, N., Hodges, L., French, A., & Dowling, R. (2021). The Effects of ALEKS on Mathematics Learning in K-12 and Higher Education: A Meta-Analysis. *Investigations in Mathematics Learning*, 13(3), 182-191. <https://doi.org/10.1080/19477503.2021.1926194>
- Tomasik, T., Helbling, L., & Moser, U. (2020). Educational gains of in-person vs. distance learning in primary and secondary schools: A natural experiment during the COVID-19 pandemic school closures in Switzerland. *International Journal of Psychology*, 56(4), 566-576. <https://doi.org/10.1002/ijop.12728>
- UNESCO. (2022). *El estudio ERCE 2019 y los niveles de aprendizaje en matemáticas: ¿Qué nos dicen y cómo usarlos para mejorar los aprendizajes de los estudiantes?* UNESCO Office Santiago and Regional Bureau for Education in Latin America and the Caribbean. <https://n9.cl/q819>
- UNESCO. (2023). *Tecnología en la educación: ¿Una herramienta en los términos de quien?* Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. <https://doi.org/10.54676/NEDS2300>
- UNESCO. (2025, Abril 25). “Ser Estudiante” 2023-2024. *La evaluación nacional permite fortalecer la calidad educativa en Ecuador*. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura: <https://n9.cl/ngf67b>

Marlon Andrés Suscal Orellana; Jorge Luis Armijos Carrión

- Uriarte Portillo, A., Zatarain Cabada, R., Barrón Estrada, M., Ibáñez, M., & González Barrón, L. (2023). Intelligent Augmented Reality for Learning Geometry. *Information*, 14(4), 245. <https://doi.org/10.3390/info14040245>
- Vanacore, K., Gurung, A., Sales, A., & Heffernan, N. (2024). Effect of Gamification on Gamers: Evaluating Interventions for Students Who Game the System. *Journal of Educational Data Mining*, 16(1). <https://doi.org/10.5281/zenodo.11549799>

