UNEMI
UNIVERSIDAD ESTATAL DE MILAGRO

REPÚBLICA DEL ECUADOR

UNIVERSIDAD ESTATAL DE MILAGRO

VICERRECTORADO DE INVESTIGACIÓN Y POSGRADO

FACULTAD DE POSGRADOS

ARTÍCULOS PROFESIONALES DE ALTO NIVEL
PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

MAGÍSTER EN INTELIGENCIA ARTIFICIAL PARA LA EDUCACIÓN

TEMA:

FEEDBACK AUTOMATIZADO CON INTELIGENCIA ARTIFICIAL Y AUTORREGULACIÓN DEL APRENDIZAJE MATEMÁTICO EN EDUCACIÓN SECUNDARIA: UNA REVISIÓN SISTEMÁTICA DE LITERATURA 2020 A MARZO 2026.

Autores:

LCDA. NIDIA MARCELA ALARCÓN MICOLTA
LCDO. EDWIN MATÍAS VIDAL LANDIN
CPA. JOHANA PAMELA GALARZA BERMEO

Tutor:

MAE. DANIEL ANDRÉS ÁNGULO VELEZ

Milagro, 2026

Feedback automatizado con inteligencia artificial y autorregulación del aprendizaje matemático en educación secundaria: una revisión sistemática de literatura 2020 a marzo 2026

Automated feedback with artificial intelligence and self-regulation of mathematical learning in secondary education: a systematic literature review 2020 to March 2026

Lcda. Nidia Marcela Alarcón Micolta nnombrealarconmicoltanidiamarcelam@unemi.ed.ec https://orcid.org/0009-0003-4998-6712 Universidad Estatal de Milagro -Ecuador	Lcdo. Edwin Matías Vidal Landin evidall@unemi.edu.ec https://orcid.org/0009-0009-6626-7318 Universidad Estatal de Milagro -Ecuador	CPA. Johana Pamela Galarza Bermeo jgalarzab7@unemi.edu.ec https://orcid.org/0009-0007-3928-5049 Universidad Estatal de Milagro -Ecuador

RESUMEN

Esta revisión sistemática analiza el papel del feedback automatizado con inteligencia artificial (IA) en el desarrollo de la autorregulación del aprendizaje (ARA) matemático en educación secundaria. El estudio siguió el protocolo PRISMA 2020 y consultó siete bases de datos (Scopus, Web of Science, ERIC, SciELO, Redalyc, MDPI y Frontiers in Education) durante el período comprendido entre enero de 2020 y marzo de 2026. De 347 registros iniciales se llegó a 20 estudios para el análisis final. Los sistemas inteligentes de tutoría con modelado del conocimiento del estudiante presentan la evidencia más consistente de efectividad, seguidos por las plataformas adaptativas y los asistentes conversacionales basados en modelos de lenguaje. El monitoreo del error es la dimensión de la ARA más favorecida por el feedback automatizado; la planificación metacognitiva, en cambio, muestra resultados contradictorios que dependen del grado de autonomía que el sistema transfiere al estudiante. Las limitaciones más frecuentes son la retroalimentación superficial, la brecha digital y la tendencia a generar dependencia tecnológica cuando el diseño instruccional no contempla la progresiva transferencia de control. La revisión concluye que el feedback automatizado produce efectos sostenidos sobre la ARA únicamente cuando el docente interpreta activamente los datos del sistema y los articula con la práctica de aula.

Palabras clave: inteligencia artificial; feedback automatizado; autorregulación del aprendizaje; aprendizaje matemático; educación secundaria; revisión sistemática.

ABSTRACT

This systematic review examines the role of AI-based automated feedback in supporting self-regulated learning (SRL) in secondary school mathematics. The study followed the PRISMA 2020 protocol and searched seven databases (Scopus, Web of Science, ERIC, SciELO, Redalyc, MDPI, and Frontiers in Education) for the period January 2020 to March 2026. From 347 initial records, 20studies met the inclusion criteria for final analysis. Intelligent tutoring systems with student knowledge modeling show the most consistent evidence of effectiveness, followed by adaptive platforms and large language model-based conversational assistants. Error monitoring is the SRL dimension most consistently supported by automated feedback; metacognitive planning, by contrast, yields contradictory results depending on how much decision-making control the system transfers to the learner. The most recurrent limitations are superficial feedback, the digital divide, and the risk of technological dependency when instructional design does not include progressive transfer of control. The review concludes that automated feedback produces sustained effects on SRLonly when teachers actively interpret system data and integrate it into classroom practice.

KEYWORDS: artificial intelligence; automated feedback; self-regulated learning; mathematics learning; secondary education; systematic review.

1. INTRODUCCIÓN (OBJETIVO DEL ARTÍCULO)

Los resultados de PISA 2022 confirmaron lo que muchos docentes de secundaria ya percibían en el aula: las matemáticas siguen siendo el principal cuello de botella de la educación media. En América Latina, dos de cada tres estudiantes concluyen la secundaria sin alcanzar el nivel mínimo de competencia matemática (OCDE, 2023). Las explicaciones habituales, como la pobreza, las debilidades en la formación docente o la sobrecarga curricular, son válidas pero incompletas. La investigación sobre el aprendizaje apunta a un mecanismo más específico: el error matemático rara vez se corrige en el momento en que ocurre. En un aula de 35 estudiantes y 80 minutos de clase, el docente no puede revisar cada procedimiento ni explicar a cada alumno dónde se equivocó, de modo que la retroalimentación, cuando llega, lo hace tarde y suele reducirse a una calificación (Sánchez-Rivas et al., 2022).

En este escenario, la inteligencia artificial educativa ofrece una alternativa. Sistemas como MATHia, de Carnegie Learning, o plataformas como Khan Academy procesan miles de

interacciones por estudiante y devuelven retroalimentación inmediata ajustada a su patrón de error particular (Mousavinasab et al., 2021). Desde 2023 se suman los asistentes basados en modelos de lenguaje de gran escala, capaces de explicar procedimientos, responder preguntas de seguimiento y graduar la complejidad de sus explicaciones según el nivel del estudiante (Kasneci et al., 2023). La promesa consiste en hacer en segundos lo que el docente, por limitaciones de tiempo, no alcanza a hacer. Sin embargo, la cuestión de fondo no es si la tecnología funciona, sino qué tipo de aprendizaje produce.

Esa cuestión remite a la autorregulación del aprendizaje, entendida como el conjunto de procesos mediante los cuales el estudiante establece metas, planifica su trabajo, monitorea su avance y ajusta sus estrategias a partir de lo que descubre sobre su propio desempeño (Zimmerman, 2000; Panadero, 2017). En matemáticas este componente resulta decisivo, porque la disciplina es acumulativa: quien no domina las ecuaciones lineales difícilmente podrá abordar las cuadráticas, y quien no comprende el concepto de función tendrá dificultades con las derivadas. Identificar a tiempo la propia laguna conceptual marca, para el estudiante de secundaria, la diferencia entre avanzar o rezagarse durante todo un período académico.

La relación entre feedback automatizado y autorregulación es clara en la teoría, pero compleja en la evidencia empírica. En el modelo cíclico de Zimmerman, la fase de autorreflexión necesita información sobre el desempeño propio para activarse, y el feedback automático la proporciona de forma inmediata y sistemática. No obstante, no todo feedback produce el mismo efecto: un sistema que entrega la respuesta correcta sin más puede generar dependencia en lugar de autonomía (Shute, 2008), y una explicación excesivamente densa puede saturar al estudiante. Su eficacia depende del tipo de retroalimentación, del momento en que se ofrece, del nivel previo del aprendiz y, sobre todo, de si el sistema le deja margen para pensar o lo sustituye en su tarea cognitiva (Nicol & Macfarlane-Dick, 2006).

A pesar de su relevancia, la evidencia sobre cómo opera esta relación en la educación secundaria es dispersa. Buena parte de los estudios más citados se realizaron con población universitaria, con muestras reducidas o con herramientas que hoy ya no están en uso. Siguen abiertas preguntas como qué habilidades de autorregulación responden mejor al feedback generado por inteligencia artificial, qué tipo de sistema resulta más efectivo según el perfil del estudiante y bajo qué condiciones pedagógicas la herramienta produce un beneficio real.

En respuesta a este vacío, el objetivo de la presente investigación es analizar la contribución del feedback automatizado con inteligencia artificial al fortalecimiento de la autorregulación del aprendizaje matemático en estudiantes de educación secundaria, mediante una revisión sistemática de la literatura científica publicada entre 2020 y 2026. Para ello, el estudio se propone identificar las principales herramientas de inteligencia artificial empleadas para generar retroalimentación automatizada en matemáticas, describir sus aportes al desarrollo de las habilidades de autorregulación, analizar los beneficios, limitaciones y desafíos pedagógicos asociados a su uso, y sistematizar los hallazgos más relevantes del campo. De este modo, la revisión busca responder a la pregunta: ¿cómo contribuye el feedback automatizado con inteligencia artificial al fortalecimiento de la autorregulación del aprendizaje matemático en estudiantes de educación secundaria, según la literatura científica reciente?

2. MARCO TEÓRICO

La autorregulación del aprendizaje: fundamentos y modelos

La autorregulación del aprendizaje se define como el proceso activo y constructivo mediante el cual el estudiante establece metas para su aprendizaje y luego monitorea, regula y controla su cognición, su motivación y su conducta en función de esas metas y de las características del contexto (Pintrich, 2000). Esta noción desplaza la atención desde lo que el docente enseña hacia lo que el estudiante hace para aprender, y supone que el aprendiz no es un receptor pasivo de información, sino un agente capaz de dirigir sus propios procesos.

El modelo más difundido es el propuesto por Zimmerman (2000), que organiza la autorregulación en un ciclo de tres fases. En la fase de previsión, el estudiante analiza la tarea, fija metas, selecciona estrategias y activa creencias motivacionales como la autoeficacia y el interés. En la fase de ejecución, pone en marcha las estrategias elegidas mientras aplica procesos de autocontrol y autoobservación. En la fase de autorreflexión, evalúa su desempeño, atribuye causas a los resultados obtenidos y genera reacciones que orientan el siguiente ciclo. El carácter cíclico del modelo es su rasgo central: lo que el estudiante concluye al reflexionar alimenta la planificación de la tarea siguiente.

Winne y Hadwin (1998) ofrecen una perspectiva complementaria, de corte más cognitivo, en la que el aprendizaje autorregulado transcurre en cuatro etapas: definición de la tarea, establecimiento de metas y planes, puesta en práctica de tácticas y adaptación

metacognitiva. Su aporte más relevante para este estudio es la importancia que conceden al monitoreo, entendido como la comparación permanente entre el producto del trabajo y los estándares que el estudiante se ha propuesto. En este modelo, la retroalimentación externa es una fuente de información que el estudiante incorpora a su monitoreo interno, lo que la convierte en un insumo directo del proceso autorregulatorio.

Panadero (2017), tras comparar seis modelos teóricos, concluye que todos comparten la estructura básica de preparación, ejecución y evaluación, aunque difieren en el peso que otorgan a los componentes cognitivos, motivacionales y emocionales. Su revisión subraya además que la autorregulación puede enseñarse y que su desarrollo depende en gran medida de las oportunidades que el entorno ofrece para practicarla. Esta idea es clave: si la autorregulación es una competencia susceptible de ser promovida, cabe preguntarse qué tipo de entornos, incluidos los tecnológicos, la favorecen o la inhiben.

Metacognición y autorregulación en el aprendizaje de las matemáticas

La metacognición, entendida como el conocimiento que el sujeto tiene sobre sus propios procesos cognitivos y la regulación que ejerce sobre ellos (Flavell, 1979), constituye el núcleo de la autorregulación en tareas de resolución de problemas. En matemáticas, su papel ha sido estudiado con particular atención. Schoenfeld (1992) mostró que muchos estudiantes fracasan en la resolución de problemas no por falta de conocimientos, sino por deficiencias en el control del proceso: eligen un camino de solución y lo siguen sin verificar si los conduce a algún lugar, sin detenerse a evaluar alternativas ni a revisar sus supuestos.

Esta observación tiene consecuencias directas para la enseñanza. El aprendizaje matemático exige que el estudiante sepa cuándo aplicar un procedimiento, que detecte a tiempo que se ha equivocado y que sea capaz de reorganizar su estrategia. Como la disciplina se construye de manera acumulativa, los errores no identificados se arrastran de un contenido al siguiente. Por ello, el monitoreo del error, la autoevaluación del propio procedimiento y el ajuste de estrategias son componentes de la autorregulación especialmente decisivos en esta área.

En la educación secundaria, estas habilidades se encuentran todavía en desarrollo. Los adolescentes poseen una capacidad metacognitiva mayor que la de los niños, pero aún dependen en buena medida de señales externas para calibrar su desempeño. Esta condición intermedia hace que la retroalimentación desempeñe un papel especialmente importante: puede

actuar como andamiaje que favorece la autonomía o, si está mal diseñada, como muleta que la reemplaza.

La retroalimentación formativa

La retroalimentación es una de las influencias más poderosas sobre el aprendizaje, aunque su efecto puede ser tanto positivo como negativo. Hattie y Timperley (2007) la definen como la información proporcionada por un agente acerca de aspectos del desempeño o de la comprensión del estudiante, cuyo propósito es reducir la distancia entre la situación actual y la meta de aprendizaje. Para lograrlo, la retroalimentación debe responder a tres preguntas: hacia dónde voy, cómo lo estoy haciendo y qué debo hacer a continuación.

Estos autores distinguen cuatro niveles en los que puede operar la retroalimentación: la tarea, que informa si la respuesta es correcta; el proceso, que se refiere a las estrategias empleadas para resolverla; la autorregulación, que apunta a la capacidad del estudiante para monitorear y dirigir su propio trabajo; y el yo, que consiste en valoraciones personales como los elogios. Los niveles de proceso y autorregulación son los que se asocian con aprendizajes más profundos y transferibles, mientras que la retroalimentación centrada en la persona tiene escaso efecto sobre el aprendizaje. Esta clasificación resulta útil para analizar los sistemas de inteligencia artificial, ya que muchos de ellos operan principalmente en el nivel de la tarea.

El metaanálisis de Kluger y DeNisi (1996) había advertido que en torno a un tercio de las intervenciones de retroalimentación analizadas empeoraban el desempeño, lo que demuestra que no basta con proporcionar información para mejorar el aprendizaje. Shute (2008), en su revisión sobre retroalimentación formativa, precisó las condiciones que favorecen su eficacia: debe ser específica, centrarse en la tarea y no en la persona, ofrecerse en un formato manejable y, preferentemente, explicar por qué una respuesta es incorrecta en lugar de limitarse a señalarla. La autora distingue entre retroalimentación verificativa, que solo indica si la respuesta es correcta, y retroalimentación elaborada, que proporciona explicaciones, pistas o ejemplos. Esta segunda modalidad es la que mayor potencial tiene para activar procesos de revisión.

Nicol y Macfarlane-Dick (2006) vincularon de manera explícita la retroalimentación con la autorregulación. Su planteamiento sostiene que la buena retroalimentación no consiste en transmitir información del docente al estudiante, sino en fortalecer la capacidad del propio estudiante para evaluar su trabajo. Entre sus principios destacan la clarificación de los criterios

de buen desempeño, el fomento de la autoevaluación y la reflexión, y la oferta de oportunidades para cerrar la brecha entre el desempeño actual y el deseado. Desde esta perspectiva, la retroalimentación se valora por su capacidad para hacerse progresivamente innecesaria.

Inteligencia artificial y feedback automatizado en educación

La inteligencia artificial aplicada a la educación abarca un conjunto amplio de sistemas que analizan datos del desempeño estudiantil para adaptar la enseñanza, evaluar respuestas o generar retroalimentación. La revisión de Zawacki-Richter et al. (2019) identificó cuatro grandes áreas de aplicación: la elaboración de perfiles y predicción, la evaluación, los sistemas adaptativos y personalizados, y los sistemas inteligentes de tutoría. Los autores señalaron, además, la escasa presencia de reflexión pedagógica en buena parte de esta investigación, predominantemente desarrollada desde las ciencias de la computación.

Los sistemas inteligentes de tutoría constituyen la tradición más consolidada. Su arquitectura clásica integra un modelo del dominio, que representa el conocimiento a enseñar; un modelo del estudiante, que estima lo que este sabe y no sabe; y un modelo pedagógico, que decide qué intervención ofrecer en cada momento (Mousavinasab et al., 2021). El modelado del estudiante suele basarse en técnicas como el rastreo del conocimiento, propuesto por Corbett y Anderson (1995), que estima la probabilidad de que el estudiante domine cada habilidad a partir de sus respuestas sucesivas. Esta capacidad diagnóstica permite que la retroalimentación se dirija al paso específico del procedimiento donde se produjo el error. VanLehn (2011) encontró que los tutores que ofrecen retroalimentación a nivel de paso alcanzan una efectividad cercana a la de la tutoría humana, y el metaanálisis de Kulik y Fletcher (2016) confirmó efectos positivos consistentes de estos sistemas sobre el aprendizaje.

Las plataformas adaptativas representan una modalidad más sencilla. Seleccionan el siguiente ejercicio en función del historial reciente del estudiante y ofrecen pistas, videos o prácticas adicionales, sin necesariamente modelar en detalle el conocimiento subyacente. Su amplia disponibilidad, a menudo gratuita, explica su difusión en contextos escolares con recursos limitados.

La irrupción de los modelos de lenguaje de gran escala ha abierto una tercera vía. Estos asistentes conversacionales generan explicaciones en lenguaje natural, responden preguntas de seguimiento y reformulan sus explicaciones cuando el estudiante no las comprende. Kasneci et al. (2023) destacan sus oportunidades para personalizar el aprendizaje, pero advierten sobre

riesgos como la producción de contenido incorrecto, la dependencia del estudiante y la necesidad de nuevas competencias docentes para integrarlos críticamente. En matemáticas, la fiabilidad de sus respuestas es una preocupación particular, dado que un error presentado con fluidez y seguridad puede ser aceptado sin cuestionamiento. A ello se suman consideraciones éticas relacionadas con la privacidad, la equidad y la transparencia de los sistemas, que Holmes et al. (2022) proponen abordar como una agenda compartida por toda la comunidad educativa.

La relación entre feedback automatizado y autorregulación

La conexión teórica entre feedback automatizado y autorregulación se apoya en el papel que el monitoreo desempeña en los modelos descritos. Si la autorreflexión, en términos de Zimmerman (2000), requiere información sobre el propio desempeño, y si el monitoreo, en términos de Winne y Hadwin (1998), consiste en comparar el producto con un estándar, entonces un sistema que devuelve información inmediata y precisa sobre el error puede fortalecer ambos procesos. La inmediatez tiene un valor añadido en matemáticas: impide que el estudiante consolide un procedimiento erróneo a lo largo de varios ejercicios.

Sin embargo, esta relación no es automática. Koedinger y Aleven (2007) formularon el denominado dilema de la asistencia: ofrecer demasiada ayuda facilita el desempeño inmediato, pero puede privar al estudiante del esfuerzo cognitivo necesario para aprender; ofrecer demasiado poca puede conducir a la frustración y al abandono. El equilibrio adecuado depende del conocimiento previo del estudiante y de la naturaleza de la tarea. La investigación sobre búsqueda de ayuda en tutores inteligentes mostró, además, que muchos estudiantes usan las pistas de forma estratégicamente inadecuada, solicitándolas de manera sucesiva hasta obtener la respuesta sin procesar las explicaciones intermedias (Aleven et al., 2016). Este comportamiento ilustra cómo un sistema diseñado para apoyar el aprendizaje puede ser utilizado para evitarlo.

La teoría de la carga cognitiva aporta otro criterio relevante. Dado que la memoria de trabajo tiene una capacidad limitada (Sweller, 1988), una retroalimentación excesivamente extensa o compleja puede saturar al estudiante en lugar de ayudarle, especialmente cuando su dominio del contenido todavía es frágil. La calidad del feedback, por tanto, no se mide por la cantidad de información que ofrece, sino por su ajuste a lo que el estudiante puede procesar en ese momento.

Molenaar (2022) propone comprender estas dinámicas mediante el concepto de regulación híbrida entre humanos e inteligencia artificial. Según este planteamiento, la regulación del aprendizaje puede distribuirse entre el estudiante, el docente y el sistema, y la cuestión decisiva es cómo se reparte el control. Cuando el sistema asume de forma permanente las decisiones de planificación y monitoreo, el estudiante no tiene ocasión de ejercerlas; cuando el control se transfiere gradualmente al estudiante, la tecnología funciona como un andamiaje que se retira a medida que la competencia se consolida. Esta perspectiva permite explicar por qué un mismo tipo de herramienta puede fortalecer ciertas dimensiones de la autorregulación, como el monitoreo del error, y debilitar otras, como la planificación.

Finalmente, el docente ocupa un lugar central en esta relación. Los sistemas de inteligencia artificial generan información detallada sobre el desempeño individual y grupal, pero esa información solo adquiere valor pedagógico cuando alguien la interpreta y la traduce en decisiones de enseñanza. La retroalimentación automatizada, desde este marco, no sustituye la mediación docente, sino que la reorganiza: libera al profesor de tareas de corrección rutinaria y le permite concentrarse en orientar la reflexión metacognitiva de sus estudiantes.

3. METODOLOGÍA

El estudio adoptó el diseño de revisión sistemática de literatura siguiendo las directrices del protocolo PRISMA 2020 (Page et al., 2021; Moher et al., 2020). Este protocolo establece las fases de identificación, cribado, elegibilidad e inclusión que permiten que el proceso sea replicable por otro investigador.

Bases de datos y período de búsqueda

Las búsquedas se ejecutaron en siete bases de datos: Scopus, Web of Science, Education Resources Information Center (ERIC), Scientific Electronic Library Online (SciELO), Redalyc, MDPI y Frontiers in Education. La elección de este conjunto buscó equilibrar las bases internacionales de mayor impacto (Scopus, Web of Science) con repositorios latinoamericanos (SciELO, Redalyc) que tienden a quedar invisibles en revisiones realizadas exclusivamente en bases anglosajonas. El período de búsqueda cubrió desde enero de 2020 hasta mayo de 2026, ventana que recoge tanto los primeros estudios sobre sistemas inteligentes de tutoría de tercera generación como los trabajos más recientes sobre asistentes basados en modelos de lenguaje.

Palabras clave y ecuaciones de búsqueda

Las búsquedas se realizaron en español e inglés, combinando términos mediante operadores booleanos. La ecuación en inglés se estructuró del siguiente modo:

("automated feedback" OR "AI feedback" OR "intelligent tutoring system") AND ("self-regulated learning" OR "metacognition") AND ("mathematics" OR "math") AND ("secondary education" OR "high school")

La ecuación equivalente en español fue la siguiente:

("retroalimentación automatizada" OR "feedback con inteligencia artificial" OR "tutor inteligente") AND ("autorregulación del aprendizaje" OR "metacognición") AND ("matemáticas") AND ("educación secundaria" OR "bachillerato")

Criterios de inclusión y exclusión

Dentro de la Tabla 1 se presentan los criterios de inclusión establecidos dentro del presente estudio.

Tabla 1

Criterios de inclusión y exclusión

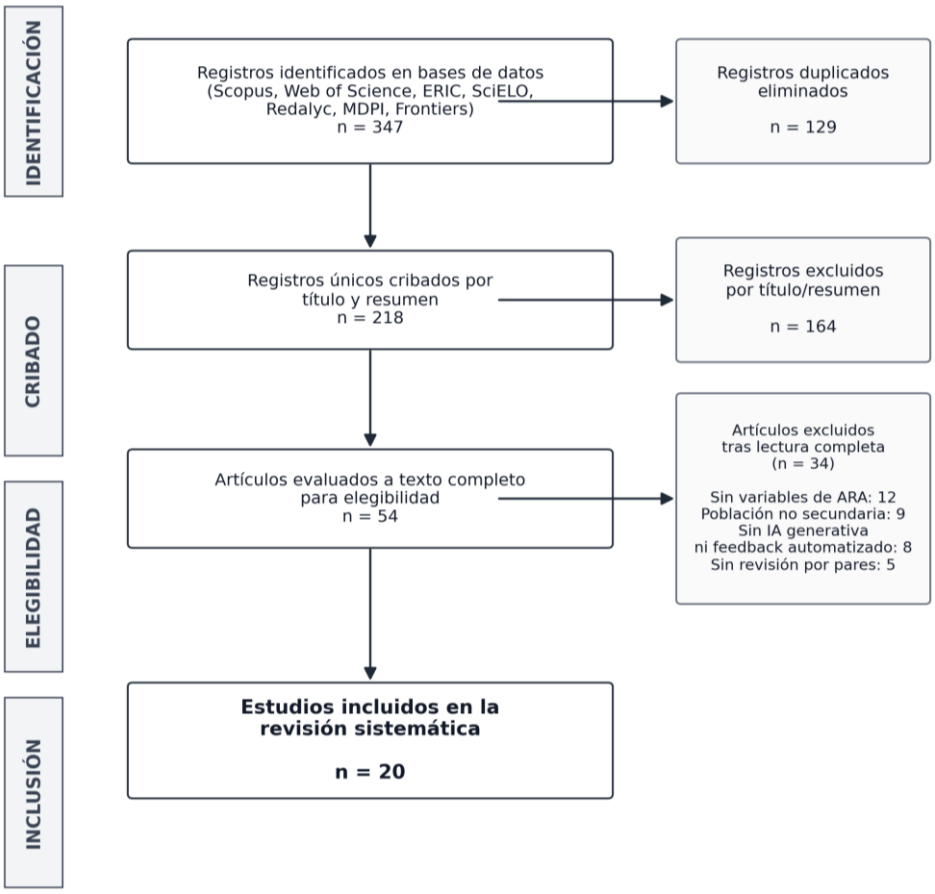
Criterios de inclusión	Criterios de exclusión
Artículos científicos con revisión de pares sobre inteligencia artificial en educación.	Documentos sin relación directa con educación o aprendizaje.
Estudios sobre feedback automatizado, sistemas inteligentes de tutoría o plataformas adaptativas.	Artículos centrados únicamente en aspectos técnicos de la inteligencia artificial sin aplicación educativa.
Investigaciones que aborden autorregulación, metacognición o aprendizaje autónomo.	Estudios que no incluyan variables relacionadas con autorregulación o metacognición.
Estudios aplicados al aprendizaje matemático o disciplinas cuantitativas afines.	Investigaciones sin relación con matemáticas ni con educación secundaria.
Publicaciones entre 2020 y 2026.	Documentos anteriores a 2020, duplicados o sin acceso al texto completo.
Artículos en español o inglés.	Opiniones, editoriales, blogs o documentos sin revisión científica.

Proceso de selección

La búsqueda inicial arrojó 347 registros distribuidos entre las bases consultadas. Tras eliminar 129 duplicados se obtuvieron 218 documentos únicos. La revisión de títulos y resúmenes excluyó 164 artículos cuyo contenido no se ajustaba al objeto de estudio, dejando 54 textos para evaluación a texto completo. La lectura íntegra de estos artículos, contrastando cada uno con los criterios de inclusión y exclusión, derivó en la exclusión adicional de 34 trabajos por las siguientes razones: ausencia de variables relacionadas con autorregulación (n = 12), población no perteneciente a educación secundaria (n = 9), ausencia de feedback automatizado o de inteligencia artificial generativa (n = 8) y publicación sin revisión por pares (n = 5). El corpus final quedó conformado por 20 estudios. El flujo completo del proceso se sintetiza en la Figura 1.

Figura 1

Diagrama de flujo para búsqueda y selección de investigaciones según PRISMA



Estrategia de análisis

El análisis se organizó en torno a una matriz de sistematización que registró, para cada uno de los 20 artículos seleccionados, las siguientes variables: año de publicación, país del estudio, diseño metodológico, tamaño y características de la muestra, herramienta de inteligencia artificial utilizada, dimensiones de autorregulación medidas, instrumento de medición empleado y principales resultados reportados. A partir de esta matriz se identificaron tres ejes temáticos que organizan los resultados presentados en la sección siguiente.

4. ANÁLISIS DE RESULTADOS

Herramientas de inteligencia artificial utilizadas para feedback automatizado

De los 20 estudios incluidos, 11 utilizaron sistemas inteligentes de tutoría como herramienta principal. MATHia y Carnegie Learning aparecen con mayor frecuencia, especialmente en investigaciones realizadas en Estados Unidos sobre álgebra de secundaria. La característica distintiva de estos sistemas es su capacidad de modelado del conocimiento del estudiante: no se limitan a registrar respuestas correctas e incorrectas, sino que infieren qué sub-habilidad domina y cuál no, y ajustan la secuencia de ejercicios en consecuencia (Pane et al., 2023). Esta capacidad diagnóstica permite que el feedback localice el error en el paso del procedimiento donde ocurrió, no en el resultado final.

Las plataformas adaptativas aparecen en ocho estudios. Khan Academy es la más utilizada, en buena medida porque su acceso gratuito facilita su uso en contextos con restricciones presupuestarias. A diferencia de los sistemas inteligentes de tutoría, las plataformas adaptativas operan con algoritmos de recomendación que seleccionan el siguiente ejercicio según el historial reciente, sin necesidad de modelar el conocimiento subyacente del estudiante. El feedback que producen es más sencillo: indica el error, ofrece una pista o un video de revisión, y propone una práctica adicional. Resulta efectivo para automatización procedimental, aunque más limitado para razonamiento conceptual.

Los asistentes conversacionales basados en modelos de lenguaje aparecen en seis estudios, todos publicados entre 2023 y 2026. El más mencionado es ChatGPT, seguido por integraciones de GPT-4 y Gemini en plataformas educativas específicas. Estos sistemas ofrecen algo que ningún sistema inteligente de tutoría clásico puede igualar: explicaciones en lenguaje natural, capacidad de responder preguntas de seguimiento y flexibilidad para

reformular una explicación si el estudiante no la entiende a la primera. La contrapartida es su fiabilidad matemática. Kasneci et al. (2023) reportan tasas de error conceptual entre 8% y 17% en problemas algebraicos de secundaria, lo que obliga a una supervisión docente activa.

Cinco estudios reportaron el uso de sistemas de evaluación automática de respuestas abiertas. Estos sistemas, basados en procesamiento de lenguaje natural, evalúan no solo si la respuesta numérica es correcta, sino la coherencia del razonamiento expresado. En problemas de justificación matemática, donde el estudiante debe explicar por qué su procedimiento es válido, este tipo de retroalimentación representa un avance respecto a los sistemas tradicionales de respuesta cerrada, aunque la calidad del análisis depende fuertemente del corpus de entrenamiento del sistema.

Aportes del feedback automatizado a la autorregulación del aprendizaje

Los efectos del feedback automatizado sobre las distintas dimensiones de la autorregulación no son uniformes. La revisión arroja un patrón claro: el monitoreo del error es la dimensión más favorecida, mientras que la planificación metacognitiva presenta resultados ambiguos.

Catorce de los 20 estudios analizados documentan mejoras consistentes en la capacidad del estudiante para identificar y categorizar sus propios errores. Esto coincide con la función primaria de cualquier sistema de feedback: hacer visible al estudiante la discrepancia entre su desempeño y el esperado. Lo interesante es que, cuando el feedback se acompaña de información acumulada sobre el patrón de errores (no solo el error puntual, sino la frecuencia con que se repite y los contextos en que aparece), los estudiantes desarrollan representaciones más sofisticadas de sus propias dificultades. Azevedo et al. (2022) muestran que esta capacidad se transfiere parcialmente a tareas sin retroalimentación: los estudiantes que han usado sistemas de inteligencia artificial durante varias semanas detectan sus errores con mayor precisión incluso cuando trabajan en lápiz y papel.

La autoevaluación (la capacidad de juzgar la calidad del propio trabajo antes de recibir retroalimentación externa) aparece fortalecida en diez estudios. El mecanismo es indirecto: cuando el estudiante observa, a lo largo de varias semanas, una correlación entre sus propios juicios y los resultados del sistema calibra mejor sus expectativas. Pero este efecto es desigual. Los estudiantes con niveles previos altos de autorregulación se benefician más; los estudiantes

con baja autonomía tienden a usar el sistema de manera reactiva, esperando la retroalimentación en lugar de anticiparla.

El ajuste de estrategias presenta resultados más heterogéneos. Aparece en ocho estudios, pero solo en aquellos donde el feedback no se limita a marcar el error, sino que ofrece una explicación del principio matemático violado y propone un enfoque alternativo. Cuando el sistema simplemente indica que la respuesta es incorrecta, los estudiantes tienden a reintentar con el mismo procedimiento o a abandonar la tarea, sin reorganización estratégica alguna.

La planificación es la dimensión más problemática. Solo cuatro estudios la abordan directamente, y los resultados se contradicen. Algunos reportan que el uso sostenido de plataformas adaptativas mejora la capacidad del estudiante para establecer metas; otros sugieren que cuando el sistema decide automáticamente qué contenido viene a continuación, el estudiante pierde la oportunidad de planificar y termina ejecutando una secuencia dictada por el algoritmo. Esta tensión, entre la eficiencia que ofrece la adaptación automática y la autonomía que requiere la metacognición, no está resuelta en la literatura.

Beneficios y limitaciones del uso de IA en el aprendizaje matemático

Entre los beneficios, la inmediatez aparece con regularidad en 16 de los 20 estudios. Recibir feedback en segundos, en lugar de días, evita que el estudiante consolide procedimientos erróneos y mantiene la continuidad de la tarea. Pane et al. (2023) reportaron, en un ensayo de gran escala con MATHia, tamaños de efecto sobre el rendimiento en álgebra que oscilan entre $d = 0{,}35$ y $d = 0{,}52$ después de ocho semanas de uso, magnitudes consideradas significativas en investigación educativa.

La personalización es el segundo beneficio más citado. Los sistemas inteligentes de tutoría y las plataformas adaptativas ajustan tanto el nivel de dificultad como el tipo de retroalimentación al historial individual, generando rutas de aprendizaje diferenciadas que serían inviables manualmente. En grupos numerosos, esta función resulta especialmente valiosa: lo que el docente no puede hacer (atender a 35 estudiantes en paralelo) lo hace el sistema.

Las limitaciones, sin embargo, son sustantivas. La retroalimentación superficial aparece como la más frecuente: cuando el feedback se limita a correcto/incorrecto o a la presentación de la solución, no produce efectos sobre la autorregulación, aunque sí los produzca sobre el

rendimiento inmediato (Shute, 2008). Este hallazgo es importante porque las plataformas más utilizadas en contextos escolares son precisamente las que ofrecen feedback más superficial.

La brecha digital es la segunda limitación más mencionada, y la más relevante para el contexto latinoamericano. Dos estudios realizados en México y Colombia documentaron que entre el 30% y el 35% de los estudiantes de secundaria en zonas rurales carece de conectividad suficiente para el uso continuo de plataformas educativas. Esta restricción no es solo técnica: rompe la lógica adaptativa del sistema, que requiere acumulación sostenida de datos para ajustar su funcionamiento.

El riesgo de dependencia tecnológica aparece en seis estudios. Cuando el estudiante recurre al sistema antes de intentar resolver por sí mismo, el feedback automatizado deja de cumplir su función de retroalimentación y pasa a sustituir el trabajo cognitivo. Este efecto, identificado como uno de los riesgos centrales del diseño instruccional pobre, depende menos de la tecnología en sí que del modo en que se incorpora a la práctica de aula.

Por último, cinco estudios señalan la ausencia de acompañamiento docente activo como factor moderador negativo. Los sistemas de inteligencia artificial generan reportes detallados sobre el desempeño individual y grupal, pero esos reportes son inútiles si el docente no los lee o no sabe interpretarlos. La tecnología, en estos casos, opera como un sustituto parcial del docente más que como un complemento de su práctica.

5. DISCUSIÓN

Los hallazgos de esta revisión permiten una respuesta que no es ni optimista ni pesimista sobre la tecnología: el feedback automatizado con IA contribuye a la ARA matemática en secundaria, pero esa contribución depende de variables que no están en el sistema, sino en cómo se diseña su uso pedagógico. El punto de convergencia más sólido entre los estudios es el efecto sobre el monitoreo del error. No sorprende: la función básica de cualquier sistema de feedback es hacer visible la discrepancia entre el desempeño actual y el esperado. Lo que añade la literatura reciente es que ese efecto se extiende a otras dimensiones de la ARA solo cuando el feedback comunica el principio matemático subyacente al error, no únicamente si la respuesta fue correcta. Mertens et al. (2022) sistematizan esta distinción en su meta análisis: el feedback con elaboración produce efectos sustancialmente mayores sobre

habilidades de orden superior. Esto tiene una implicación práctica directa: el tipo de feedback que más se usa en las plataformas escolares (el más barato de implementar, el que solo valida o rechaza respuestas) es también el menos efectivo para la ARA. La principal contradicción se localiza en la planificación metacognitiva. Algunos estudios reportan mejoras después del uso sostenido de plataformas adaptativas; otros no encuentran efecto o reportan deterioro. Mejeh et al. (2024) identifican la variable explicativa: cuando la plataforma controla la secuencia de contenidos, la planificación del estudiante se inhibe porque no hay nada que planificar. La implicación no es que los sistemas adaptativos sean malos, sino que su uso debe diseñarse para que el estudiante conserve agencia sobre al menos algunos aspectos de su trayectoria. Sobre cuál herramienta es más efectiva, la evidencia favorece a los ITS con modelado del conocimiento. Pero esa superioridad viene con un costo: licencias costosas, infraestructura tecnológica y la necesidad de un docente que interprete y use activamente los datos que el sistema produce (Bardach et al., 2026). Las plataformas adaptativas son más accesibles y más extendidas, pero sus efectos sobre la ARA son más débiles. Los asistentes basados en modelos de lenguaje son los más flexibles, pero su fiabilidad matemática es todavía insuficiente para uso autónomo sin verificación docente (Kasneci et al., 2023). Ninguna herramienta domina en todos los criterios relevantes. El rol del docente emerge como variable moderadora consistente. El título del artículo de Bardach et al. (2026) resume el hallazgo con precisión: Intelligent Tuto ring Systems Need Teachers. La IA no reemplaza la mediación pedagógica; la reorganiza. Cuando el docente interpreta los reportes del sistema y los conecta con decisiones de aula, los efectos sobre la ARA son significativamente mayores. Cuando el sistema opera como actividad paralela sin integración instruccional, los efectos se diluyen. Esta conclusión tiene implicaciones directas para la formación docente: el acceso a la tecnología es necesario, pero no suficiente; lo que hace la diferencia es la competencia del docente para darle un uso pedagógico (Holmes et al., 2022).

Una advertencia metodológica cierra la discusión. La mayor parte de los estudios incluidos provienen de contextos con conectividad estable e infraestructura tecnológica adecuada. Los efectos documentados probablemente sobreestiman lo que se observaría en implementaciones masivas en contextos latinoamericanos con brecha digital significativa. Extrapolar sin esa cautela sería metodológicamente irresponsable.

6. CONCLUSIÓN

La revisión derivó cinco conclusiones específicas, articuladas con los objetivos del estudio. Los sistemas inteligentes de tutoría con modelado del conocimiento del estudiante tienen la evidencia más consistente de efectividad sobre el rendimiento matemático y la ARA. Su capacidad diagnóstica localiza el error donde ocurre, condición necesaria para que el feedback active procesos de revisión metacognitiva. Las plataformas adaptativas son más accesibles, pero producen efectos más débiles sobre la ARA cuando su diseño no incluye mecanismos de transferencia progresiva de autonomía al estudiante. El monitoreo del error es el componente de la ARA más favorecido por el feedback automatizado. La planificación metacognitiva presenta resultados contradictorios que dependen de si el sistema controla o deja al estudiante la secuencia de aprendizaje. Esto sugiere que el diseño instruccional, no la sofisticación técnica del sistema determina qué dimensiones de la ARA se desarrollan y cuáles quedan inhibidas. El feedback automatizado produce efectos sostenidos sobre la ARA únicamente cuando el docente integra activamente los datos del sistema en su práctica de aula. La tecnología no sustituye la mediación pedagógica: la reorganiza. Las intervenciones con efectos más consistentes son aquellas donde el docente usa los reportes del sistema para ajustar su instrucción grupal y orientar la reflexión metacognitiva individual. La brecha digital no es una limitación periférica: afecta de forma estructural los beneficios que pueden obtenerse. Los estudiantes de zonas rurales o con conectividad irregular no acumulan el historial de interacciones que los sistemas adaptativos necesitan para funcionar. Cualquier política de adopción tecnológica que no contemple esta restricción amplía la desigualdad en lugar de reducirla.

El riesgo de dependencia tecnológica no es inherente al feedback automatizado; depende del diseño. Cuando el sistema entrega la ayuda antes de que el estudiante la solicite o decide todas las rutas de aprendizaje, la autonomía metacognitiva se deteriora. Cuando el feedback está disponible, pero requiere una acción del estudiante para activarse, la dependencia desaparece. Esta diferencia de diseño, aparentemente técnica, determina si la tecnología desarrolla la ARA o la reemplaza.

7. BIBLIOGRAFÍA

Azevedo, R., Taub, M., & Mudrick, N. V. (2022). Technologies supporting metacognition and self-regulated learning during hypermedia learning. En D. H. Schunk & J. A. Greene (Eds.), Handbook of self-regulation of learning and performance (2.ª ed., pp. 104–118). Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. https://doi.org/10.3102/003465430298487

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide agenda. Journal of Learning Analytics, 9(1), 1–23. https://doi.org/10.18608/jla.2022.7527

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2020). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLOS Medicine, 6(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Mousavinasab, E., Zarifsanaiey, N., Niakan Kalhori, S. R., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2021). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. Interactive Learning Environments, 29(1), 142–163. https://doi.org/10.1080/10494820.2018.1558257

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. Studies in Higher Education, 31(2), 199–218. https://doi.org/10.1080/03075070600572090

OCDE. (2023). PISA 2022 results (Volume I): The state of learning and equity in education. OECD Publishing. https://doi.org/10.1787/53f23881-en

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ, 372, n71. https://doi.org/10.1136/bmj.n71

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. Frontiers in Psychology, 8, 422. https://doi.org/10.3389/fpsyg.2017.00422

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2023). Effectiveness of cognitive tutor algebra I at scale. Educational Evaluation and Policy Analysis, 36(2), 127–144. https://doi.org/10.3102/01623737X14507447

Sánchez-Rivas, E., Ruiz-Palmero, J., & Sánchez-Rodríguez, J. (2022). Impacto del feedback formativo en el rendimiento matemático de estudiantes de educación secundaria. Revista de Educación, (395), 37–64. https://doi.org/10.4438/1988-592X-RE-2022-395-536

Shute, V. J. (2008). Focus on formative feedback. Review of Educational Research, 78(1), 153–189. https://doi.org/10.3102/0034654307313795

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? International Journal of Educational Technology in Higher Education, 16(1), 39. https://doi.org/10.1186/s41239-019-0171-0

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. En M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), Handbook of self-regulation (pp. 13–39). Academic Press.

UNEMI
UNIVERSIDAD ESTATAL DE MILAGRO

CARTA DE ACEPTACIÓN

Date: 2026-06-23

Autor(s): Nidia Marcela Alarcón Micolta[1], Edwin Matías Vidal Landín[2], Johana Pamela Galarza Bermeo[3].

Title: Feedback automatizado con inteligencia artificial y autorregulación del aprendizaje matemático en educación secundaria: una revisión sistemática de literatura 2020 a marzo 2026

Estimado autor(s), felicidades su artículo científico ha sido aceptado para publicación luego de una revisión por pares ciegos en **Annals Scientific Evolution**, página web, **https://magazineasce.com/** **ISSN: 3073-1178** en Publicación continua **Vol. 5. Núm. 3 (2026).** Las decisiones y acciones del Consejo Editorial se basan en principios éticos basados en la creencia de que las revistas científicas y los editores deben seguir los últimos avances de las revistas científicas.

Asce Magazine es una revista cientifica de acceso abierto con revisión doble ciego (pares ciegos) su principal propósito es diseminar los hallazgos de investigaciones en áreas multidisciplinarias. Dirigida a un público compuesto por docentes, investigadores, estudiantes y profesionales interesados en la investigación científica donde se evalúa la calidad y la relevancia de cada contribución.

Saludos.

Editor en Jefe
Ing. Mg.
Darwin Pico

UNEMI

UNIVERSIDAD ESTATAL DE MILAGRO

¡Evolución académica!

@UNEMIEcuador