

Uso de la inteligencia artificial (IA) en el proceso enseñanza de la lectoescritura

Use of artificial intelligence (AI) in the literacy teaching process

Uso dell'intelligenza artificiale (IA) nel processo di insegnamento della letto-scrittura

Sarina Magdalena Recalde Pincay¹
sarina.recaldepincay8272@upse.edu.ec
<https://orcid.org/0009-0002-4026-7781>

Luis Leonardo Zambrano-Vacacela²
lzambrano@upse.edu.ec
<https://orcid.org/0000-0002-2968-1351>

Correspondencia: sarina.recaldepincay8272@upse.edu.ec

Artículo de Investigación

*** Recibido:** 20 de junio de 2026 *** Aceptado:** 26 de julio de 2026 *** Publicado:** 28 de agosto de 2026

- I. Universidad Península de Santa Elena, Santa Elena – Ecuador.
- II. Universidad Península de Santa Elena, Santa Elena – Ecuador.

Resumen

La investigación midió los efectos del uso de herramientas de inteligencia artificial (IA) sobre el desempeño en lectoescritura de 50 estudiantes de Décimo Grado de Educación Básica de una institución educativa de la provincia de Pichincha, Ecuador, durante el periodo lectivo 2025–2026. Para lograr los resultados de investigación se optó por un enfoque cuantitativo de alcance explicativo, en la misma línea se desarrolló un diseño cuasiexperimental del pretest y posttest. El grupo experimental que formó parte de la investigación estuvo compuesto de 25 estudiantes a los que se les aplicó tres plataformas Duolingo ABC, Speechify y ChatGPT, dentro de este proceso de toma de información se tomó una prueba de lectoescritura a la que se le aplicó criterio de validez como V de Aiken de 0,87 y análisis de confiabilidad de alfa de Cronbach de 0.81, otro instrumento fue la rúbrica analítica de escritura CCI con 0,86 y además un cuestionario tipo Likert para analizar las percepciones estudiantiles. La investigación se desarrolló durante un periodo de doce semanas, en este proceso el grupo experimental alcanzó un resultado de posttest de 16,4 puntos sobre 20, frente al pretest de 13,1, además los resultados mostraron evidencia estadísticamente significativa con un $p < 0,001$ y una d de Cohen de 1,64 lo que evidenció resultados positivos. Entre los aspectos de mayor evidenciade mejora se encuentran la fluidez lectora con una d de Cohen de 2,12 y escritura argumentativa con d de Cohen 1,72 y comprensión literal con d de Cohen 1,09. Estos resultados ponen en evidencia que el uso de la IA y una planificación didáctica adecuada con la participación del docente mejora el rendimiento lector y escritor.

Palabras clave: inteligencia artificial; lectoescritura; fluidez lectora; comprensión lectora; escritura emergente.

Abstract

This research measured the effects of using artificial intelligence (AI) tools on the literacy performance of 50 tenth-grade students at an educational institution in the province of Pichincha, Ecuador, during the 2025–2026 academic year. To achieve the research results, a quantitative approach with an explanatory scope was chosen, employing a quasi-experimental pretest-posttest design. The experimental group consisted of 25 students who were administered three platforms: Duolingo ABC, Speechify, and ChatGPT. As part of the data collection process, a literacy test was administered, which was validated using Aiken's V (0.87) and Cronbach's alpha (0.81). Another instrument used was the CCI writing rubric (0.86), and a Likert-type questionnaire was also

administered to analyze student perceptions. The research was conducted over a twelve-week period. During this time, the experimental group achieved a post-test score of 16.4 out of 20, compared to a pre-test score of 13.1. Furthermore, the results showed statistically significant evidence ($p < 0.001$) and Cohen's d of 1.64, indicating positive outcomes. Among the areas showing the greatest improvement were reading fluency (Cohen's d of 2.12), argumentative writing (Cohen's d of 1.72), and literal comprehension (Cohen's d of 1.09). These results demonstrate that the use of AI and appropriate lesson planning with teacher participation improves reading and writing performance.

Keywords: artificial intelligence; literacy; reading fluency; reading comprehension; emergent writing.

Riassunto

Questa ricerca ha misurato gli effetti dell'uso di strumenti di intelligenza artificiale (IA) sulle prestazioni di letto-scrittura di 50 studenti del decimo anno dell'istruzione di base di un'istituzione educativa della provincia di Pichincha, Ecuador, durante l'anno scolastico 2025–2026. Per ottenere i risultati della ricerca è stato adottato un approccio quantitativo di portata esplicativa, con un disegno quasi-sperimentale pretest-posttest. Il gruppo sperimentale era composto da 25 studenti ai quali sono state applicate tre piattaforme: Duolingo ABC, Speechify e ChatGPT. Nel processo di raccolta dei dati è stata somministrata una prova di letto-scrittura, validata mediante V di Aiken pari a 0,87 e alfa di Cronbach pari a 0,81; è stata inoltre utilizzata la rubrica analitica di scrittura CCI con 0,86 e un questionario Likert per analizzare le percezioni degli studenti. La ricerca si è svolta per dodici settimane. Il gruppo sperimentale ha raggiunto un punteggio post-test di 16,4 su 20, rispetto a un pretest di 13,1; i risultati hanno mostrato significatività statistica ($p < 0,001$) e d di Cohen pari a 1,64. I maggiori miglioramenti si sono osservati nella fluidità di lettura ($d = 2,12$), nella scrittura argomentativa ($d = 1,72$) e nella comprensione letterale ($d = 1,09$). I risultati evidenziano che l'uso dell'IA, accompagnato da una pianificazione didattica adeguata e dalla partecipazione del docente, migliora le prestazioni di lettura e scrittura.

Parole chiave: intelligenza artificiale; letto-scrittura; fluidità di lettura; comprensione della lettura; scrittura emergente.

Introducción

Los datos correspondientes a las competencias relacionadas a las habilidades lectoras y de escritura en America Latina demuestran resultados alarmantes, ya que como lo menciona el informe ERCE 2019 de la UNESCO (2021) identifico que el 51% de los estudiantes del sexto año de Educación General Básica no logran niveles aceptables de lectura. En esta línea se muestra Ecuador con resultados iguales o inclusive menores ya que el Instituto Nacional de Evaluación Educativa (INEVAL, 2022) determinó que apenas el 37% de los estudiantes de sexto año de EGB obtiene resultados satisfactorios de comprensión lectora. Estas estadísticas demuestran que los jóvenes llegan al colegio sin lograr el desarrollo de competencias adecuadas relacionadas a la codificación lectora que deberían desarrollar dentro de ese nivel educativo.

Esta problemática es difícil de entender y afrontar y esta no solo se explica desde un análisis diagnóstico, ya que como lo mencionó por Ferreiro y Teberosky (1979) quienes mencionaron en su estudio que aprender a leer y escribir es un proceso activo que requiere compromiso del estudiante y no es solo un proceso de transferencia mecánica de actividades y análisis de palabras como códigos y más, sino más bien que es un proceso complejo de alto nivel cognitivo. Scarborough (2001) o menciono mediante una analogía como una cuerda trenzada que para ser eficiente depende de todas sus hebras, en este sentido las hebras son conciencia fonológica, vocabulario, sintaxis, conocimiento de género discursivo deben entretorse para lograr esta finalidad que es la lectura fluida. Así como lo menciona Gough y Tunmer (1986) en donde establecen que la decodificación es un factor de vital importancia y sin importar si las otras competencias relacionadas a la lectura estan bien no se logra esta finalidad de forma fluida.

Lo que sí ha cambiado es la disponibilidad de herramientas. Los sistemas de tutoría inteligente entornos computacionales que ofrecen instrucción adaptativa en tiempo real, según la definición que VanLehn (2011) popularizó llevan décadas demostrando efectividad en matemáticas. En este sentido tanto el procesamiento del lenguaje natural cierta las condiciones para el desarrollo de la lengua escrita. En la actualidad se presenta ciertas condiciones que la tecnología proyecta como el uso de plataformas que tienen la capacidad de analizar y evaluar producciones textuales, señalar errores de cohesión y brindar retroalimentación automatizada a las necesidades de los estudiantes no de forma genérica sino más bien adaptadas a las necesidades de los estudiantes de forma puntual y contextualizada (Warschauer y Grimes, 2008; Stevenson y Phakiti, 2019).

Aunque estas investigaciones se han desarrollado en países desarrollados en su mayoría existe poca o escasa investigación en países hispano hablantes como lo menciona Gómez-Merino et al. (2023), en donde presenta una investigación de revisión de la literatura brindando como resultado que la mayoría de las investigaciones que analizan dichas temáticas se encuentran en países anglosajones o asiáticos, lo que pone en evidencia que estos resultados que no se pueden interpolar a países como el nuestro que carece de infraestructura tecnológica y con limitaciones marcadas de conectividad que se encrujecen dentro de zonas rurales donde las instituciones educativas presentan muchas limitaciones tecnologías y de formación del docente sobre el manejo de estas. El estudio se desarrolló en una institución educativa de un entorno semiurbano de la provincia de Pichincha. Los docentes de Lengua y Literatura habían reportado, en diagnósticos previos, que sus estudiantes de Décimo Grado mostraban dificultades particulares en escritura argumentativa y en comprensión de textos expositivos complejos, justo el tipo de tareas donde los sistemas de IA generativa pueden ofrecer andamiaje adaptativo (Hmelo-Silver et al., 2022). Ese diagnóstico interno coincide con lo que la literatura regional viene señalando hace tiempo.

El campo tampoco carece de voces críticas. Popenici y Kerr (2017) advirtieron que la adopción acrítica de IA en educación puede reproducir sesgos pedagógicos previos si no se acompaña de diseño instruccional deliberado. Holmes et al. (2022) precisaron que la efectividad de los sistemas de IA educativa depende, en buena medida, de la calidad con que se integran al currículo. Esta perspectiva determina la importancia de usar tecnología y especialmente IA con finalidad pedagógica y con objetivos didácticos claros y es ahí donde el docente juega un papel fundamental en vincular por un lado la pedagogía y por otro el uso de tecnología.

Por otro lado, en Ecuador se existe mucha resistencia por parte de la comunidad docente por que el pasar de una metodología tradicional a una caracterizada por uso de tecnología y otros aspectos didácticos obligan a que se tengan que actualizar y reinventar la forma en como interactúan con sus estudiantes, es así como Chai et al. (2021) documentaron la importancia de la actitud que el docente presenta hacia la IA y su implementación y como esta relación predice la calidad de los procesos educativos con una correlación de 0,58. Es decir por más que una tecnología está bien diseñada si va a ser usada por un docente que no confía en esta forma de trabajo se va a tener como resultado una subutilización de esta, este aspecto tiene implicaciones directas ya que se pone en

énfasis la importancia de la capacitación docente y el cambio de actitud hacia el uso de tecnología de forma eficiente.

En base a cada uno de los aspectos mencionados se presenta la siguiente pregunta de investigación: ¿en qué medida el uso sistemático de herramientas de IA, integrado a la práctica docente habitual, mejora el desempeño en lectoescritura de estudiantes de Décimo Grado?, por lo que para responder a dicha pregunta se plantea el siguiente objetivo de investigación determinar los efectos de un programa de enseñanza mediado por IA sobre la fluidez lectora, la comprensión literal e inferencial y la producción escrita argumentativa. Y finalmente para el cumplimiento de este objetivo general se plantea los siguientes objetivos específicos: (a) comparar el rendimiento pre y postest entre grupos, (b) identificar las dimensiones con mayor ganancia diferencial y (c) indagar las percepciones de los estudiantes sobre la viabilidad de la integración tecnológica en su institución.

Modelos cognitivos de lectoescritura

Las habilidades de lectura y escritura son competencias que se desarrollan de forma conjunta y a estas no se las puede abordar de forma independiente. Como lo menciona Gough y Tunmer (1986) analiza la relación existente entre decodificación y comprensión lingüística multiplicativa, no aditiva, estas decir estas dos habilidades van de la mano, por lo que si un estudiante que no desarrolle una decodificación no logrará comprender el contenido sin importar cuantas palabras conozca de forma individual. De la misma mano lo mencionado por Scarborough (2001) con la idea de la cuerda trenzada con múltiples hebras haciendo una analogía de las hebras a elementos como conciencia fonológica, vocabulario, estructuras sintácticas, conocimiento de género discursivo, cada una de estas habilidades se conjugan y trabajan de forma integrada para lograr una lectura fluida, este punto de vista y definición permite identificar como para la lectura se requiere un conjunto de elementos que la conforman.

Lo mencionado por Defior et al. (2022) mencionan aspectos fundamentales como las habilidades de la capacidad de identificar y manipular los fonemas que conforman las palabras segmentar sílabas, aislar fonemas iniciales, producir rimas se convierten en predictores de las habilidades lectoras, este es un aspecto fundamental y necesario para estudiantes de habla hispana que por las características del idioma cuando los textos aumentan en complejidad se requiere el desarrollo de estas habilidades de forma conjunta (Defior y Serrano, 2021).

Otro aspecto de vital importancia es la capacidad del estudiante de lograr la decodificación y el convertir grafemas en fonemas, como lo menciona Perfetti (2007) el hecho de lograr la comprensión lectora se encuentra relacionada con la calidad de las representaciones léxicas almacenadas en la memoria a largo plazo. De manera seguida se requiere reconocer la palabra y acceder a su significado con rapidez para liberar recursos cognitivos hacia el procesamiento del texto. Por tal finalidad se da la diferenciación de que dos lectores con la misma rapidez pueden diferir en su capacidad de comprensión.

De manera seguida al combinar con la escritura se añade mayor nivel de complejidad, como lo menciona Hayes y Flower (1980), revisado posteriormente por Hayes (1996), en su modelo brinda una comprensión del proceso escrito como un ciclo que tiene una secuencialidad relacionada a planificación, textualización y revisión mediado por la memoria de trabajo. Entre las fases de mayor importancia se encuentra la revisión: Alamargot y Chanquoy (2001), en donde los escritores novatos tienden a repetir la revisión de aspectos como ortografía y puntuación, y no se enfocan en la organización de texto en su estructura global. En esta dificultad presentada las plataformas tecnológicas se convierten en una alternativa de calidad para lograr retroalimentación en este aspecto no tomado en cuenta como lo es la estructura integral de los documentos en su estructura argumentativa.

Inteligencia artificial en educación básica

La aplicación de la Inteligencia artificial dentro de la educación y especial en contextos educativos de primaria se ha venido dando de forma gradual, aunque en los últimos años se ha incrementado de forma significativa. Baker (2014) realizaron una revisión de los últimos veinte años de sistemas de tutorías inteligentes se han venido analizando dentro de contextos educativos relacionados a asignaturas como Lengua y Literatura y en el área de lectoescritura se da un incremento en la calidad de lectura y escritura.

En los últimos años el desarrollo de tecnologías de Inteligencia artificial generativa con el uso de modelos de lenguaje natural permite que los sistemas puedan entender el contexto de lenguaje de una forma fluida, en este sentido modelos de lenguaje de gran escala pueden generar texto coherente, dar retroalimentación a los estuantes para la mejora de estas competencias relacionadas a las habilidades lectoras y de escritura y dar una retroalimentación de forma directa y contextualizada a las necesidades de los estudiantes (Bender et al., 2021). De la misma forma

Zawacki-Richter et al. (2019) mencionan que estos sistemas de IA en sus primeras versiones no eran infalibles, pero con el paso de los años han mejorado de forma significativa, aunque es necesario que el docente sea el auditor de la calidad de las IAs y con ello se logre una mejora calidad en las retroalimentaciones.

Es necesario que para su implementación de forma adecuada se requiere de un diseño instruccional adecuado, el cual debe tomar en cuenta aspectos como: adaptarse a las habilidades de los estudiantes y por ende las actividades deben estar diseñadas para responder a dichas necesidades, por otro lado, ofrecer retroalimentación inmediata y personalizada (Woolf, 2010). Luckin et al. (2016) menciona que se requiere también del compromiso del estudiante ya que el estudiante debe desarrollar habilidades metacognitivas de autorregulación y autopercepción de sus debilidades y a partir de ello mejorar sus conocimientos y competencias en cuanto a la lectura y escritura.

En Ecuador no se encuentra contemplado dentro de las leyes que regulan el sistema educativo como la Ley Orgánica de Educación Intercultural (LOEI) dentro de estas no se contempla el uso de la IA como herramienta pedagógica, esto provoca que cada institución educativa y docente la usen según sus criterios y perspectivas. Coll (2018) manifestó que pese a que no hay una normativa que regule su uso o adopción permite que el docente sea quien promueva la innovación de forma responsable, pero esta debe estar direccionada bajo la panorámica de una orientación pedagógica y que su uso sea bajo un objetivo didáctico claro y con una metodología que le de un soporte.

Evidencia empírica sobre IA y adquisición lectora

Al revisar la literatura respecto a la evidencia empírica sobre la IA y el desarrollo lector, se observa que existe mucha heterogeneidad en las investigaciones. Suspitsyna et al. (2023) revisaron las investigaciones existentes sobre dicha temática e identificaron 47 estudios sobre herramientas de escritura asistida por IA, dentro de las mismas identificaron en promedio un efecto de 0,72 sobre la calidad textual, estos identificaron un mayor nivel de efecto sobre la argumentación que en aspectos descriptivos. Aunque se evidencia que de toda la literatura solo el 4% pertenece a investigaciones latinoamericanas.

En el contexto del habla hispana como lo menciona Rodríguez-García et al. (2023) investigaron una plataforma de evaluación automática de textos de estudiantes en nivel de educación secundaria, dentro del mismo identificaron que se presento mejoras significativas en torno a la cohesión y coherencia textual. Por otro lado, Morales-López y Hernández (2022), identificaron

resultados con cierto nivel de similitud en México con muestras pequeñas, lo que limitó la calidad de sus resultados y su efecto estadístico.

En esta misma línea la fluidez lectora ha sido analizada por Rasinski (2022) dentro de su estudio analizó evidencia sobre lecturas asistidas por tecnologías, entre los que se destacaban sistemas de reconocimiento de voz para identificar errores de decodificación, identificando que este tipo de tecnologías presentan resultados interesantes, aunque si se usan sin una finalidad pedagógica no logran consolidar las habilidades de lectura. La interacción que se presenta entre el uso de tecnologías e instrucción directa se busca comprender y analizar como se ha desarrollado dentro del contexto ecuatoriano.

Bajo esta panorámica la comprensión lectora es la finalidad de mayor dificultad de lograr como lo menciona Guthrie et al. (2004) evidenciaron que la motivación se convierte en un aspecto vital para lograr la instrucción y la comprensión lectora, ya que los estudiantes que no logran niveles de motivación adecuados muy difícilmente van a aprovechar de buena manera la tecnología. Por tal motivo es importante establecer que los sistemas de IA que pueden incorporar la gamificación como una estrategia de motivación mediante el uso de retroalimentación visual, progresión de niveles, recompensas simbólicas parecen tener mejores resultados (Dichev y Dicheva, 2017).

Otro aspecto vital para lograr una comprensión lectora adecuada en los estudiantes está relacionado con el docente y su actitud Chai et al. (2021) y su predisposición al uso de IA dentro de estos procesos, estos aspectos cuando se ejecutan de buena manera predicen $r = 0,58$ en su integración de forma efectiva dentro del aula. En este sentido, es evidente que la formación del docente se conforma como elementos vitales para incorporar la IA en procesos relacionados a las mejoras de la lectura y escritura por parte de los estudiantes.

Metodología

Enfoque de la investigación

La investigación adoptó un enfoque cuantitativo, dado que su finalidad consistió en medir los efectos del uso de herramientas de inteligencia artificial sobre el desempeño lector y escritor mediante variables operacionalizadas y análisis estadísticos. Hernández-Sampieri y Mendoza (2018) caracterizan este enfoque por el uso secuencial de procedimientos de recolección y procesamiento numérico de datos para contrastar hipótesis y establecer regularidades. La

comparación entre puntuaciones pre y postest, así como la magnitud del efecto entre grupos, requería un tratamiento numérico que el enfoque cualitativo no podía proporcionar con la precisión necesaria.

Tipo y diseño de investigación

Tipo

El estudio es de tipo aplicado, ya que buscó resolver un problema concreto del aula el bajo rendimiento en lectoescritura de Décimo Grado mediante la incorporación sistemática de herramientas tecnológicas con sustento didáctico. El alcance es explicativo: se propuso determinar la relación causal entre el uso de IA (variable independiente) y el desempeño lector y escritor (variable dependiente). Hernández-Sampieri y Mendoza (2018) señalan que los estudios explicativos van más allá de la descripción de conceptos o fenómenos y procuran establecer las causas de los hechos investigados.

Diseño

El diseño fue cuasiexperimental con el uso y aplicación de una preprueba y posprueba, con dos grupos no equivalentes (experimental y control). Esta decisión obedeció a una restricción práctica: los estudiantes ya estaban distribuidos en secciones por criterios administrativos institucionales, lo que impidió la asignación aleatoria individual. Campbell y Stanley (1966) demostraron que este tipo de diseño preserva la validez interna mediante mediciones previas que permiten estimar la equivalencia inicial entre grupos, y Shadish et al. (2002) sistematizaron los procedimientos estadísticos para controlar las diferencias preexistentes, principalmente el análisis de covarianza (ANCOVA). El esquema del diseño se representa en la Tabla 1.

Tabla 1

Esquema del diseño cuasiexperimental

Grupo	Pretest	Tratamiento	Posttest
Experimental (n = 25)	O ₁	X	O ₂
Control (n = 25)	O ₁	—	O ₂

Nota. O₁ = preprueba; X = aplicación del programa de enseñanza mediado por IA; O₂ = posprueba; — = ausencia de tratamiento experimental.

Población y muestra

Población

La población estuvo conformada por los 112 estudiantes matriculados en Décimo Grado de Educación Básica de la institución participante durante el periodo lectivo 2025–2026, distribuidos en cuatro paralelos. Bernal (2016) define la población como el conjunto de todos los elementos que comparten una serie de características comunes y respecto de los cuales se procura inferir o describir un fenómeno educativo.

Muestra

La muestra estuvo integrada por 50 estudiantes de dos paralelos completos, seleccionados mediante muestreo no probabilístico por conveniencia, atendiendo a criterios de accesibilidad institucional y disponibilidad de equipamiento tecnológico. El grupo experimental comprendió 25 estudiantes (13 mujeres, 12 hombres; edad media = 14,6 años; DT = 0,51) y el grupo de control, otros 25 (14 mujeres, 11 hombres; edad media = 14,7 años; DT = 0,48). Para determinar que estudiantes formaban parte del estudio se tomó en cuenta cuando presentaban las siguientes características: Estar matriculados, contar con un dispositivo con acceso a internet sea dentro de la institución educativa o en el hogar, además de ello se excluyó a dos estudiantes que contaban con necesidades educativas especiales ya que necesitaban software especializado y en la investigación no se contaba con el mismo. Finalmente, los dos grupos parte de la investigación se los comparó en la asignatura de Lengua y Literatura según sus calificaciones obtenidas en el año anterior con una p valor de 0,41.

Variables

Variable independiente: uso de herramientas de inteligencia artificial

Se definió como la incorporación sistemática de tres plataformas basadas en IA al trabajo de aula de Lengua y Literatura. Sus dimensiones fueron: (a) frecuencia de uso (dos sesiones semanales de 45 minutos durante doce semanas), (b) tipo de herramienta (Duolingo ABC para conciencia fonológica, Speechify para fluidez lectora y ChatGPT con prompts pedagógicos para retroalimentación escrita), (c) interacción pedagógica (mediada por el docente investigador) y (d) retroalimentación automatizada (devoluciones específicas por dimensión textual). Russell y Norvig (2021) definen la inteligencia artificial como el conjunto de sistemas capaces de ejecutar

tareas que normalmente exigen inteligencia humana, incluido el procesamiento del lenguaje natural y el análisis adaptativo de respuestas.

Variable dependiente: desempeño lector y escritor

Se operacionalizó a través de cuatro dimensiones medidas con la prueba ad hoc: fluidez lectora oral, comprensión literal, comprensión inferencial y producción escrita argumentativa. Ferreiro y Teberosky (1979) menciona que la lectoescritura se caracteriza por ser un proceso que permite la construcción de significados que está compuesto de elementos como representación gráfica del lenguaje y comprensión textual, bajo esa caracterización se estableció la operacionalización propuesta. Bajo esta panorámica esta dimensión se desarrolló sobre 20 puntos permitiendo de esta forma calcular las puntuaciones específicas y generales.

Hipótesis

Para Hernández-Sampieri y Mendoza (2018), la hipótesis constituye el elemento determinante que guía el proceso investigativo y que requiere ser validada, a continuación, se presentan las hipótesis de investigación:

Hipótesis nula (H_0). No existen diferencias estadísticamente significativas entre las medias del grupo experimental y del grupo de control en el postest de lectoescritura tras la intervención mediada por herramientas de IA. Expresada formalmente: $H_0: \mu_{GE} = \mu_{GC}$.

Hipótesis alterna (H_1). Los estudiantes que reciben un programa de enseñanza de la lectoescritura mediado por herramientas de IA obtienen, en el postest, puntuaciones significativamente superiores a las de sus pares con instrucción convencional. Expresada formalmente: $H_1: \mu_{GE} > \mu_{GC}$.

El contraste se realizó con un nivel de significación $\alpha = 0,05$. La regla de decisión adoptada fue rechazar H_0 si el valor p obtenido en la prueba t para muestras independientes resultaba inferior a 0,05 y, complementariamente, si el ANCOVA con el pretest como covariable confirmaba un efecto significativo de la intervención sobre el postest.

Técnicas e instrumentos de recolección de datos

Técnicas

Se aplicaron un conjunto de técnicas adicionales como las pedagógicas que permitió medir la calidad lectora y de escritura en dos momentos. Bajo esta panorámica como lo menciona Ander-

Egg (2011), permite ser una posibilidad de obtener información sobre actitudes, percepciones y comportamientos con la finalidad de tener información sobre las percepciones de los estudiantes en referencia a los instrumentos aplicados. De la misma forma se aplicó observación estructurada para identificar la participación de los estudiantes y la interacción que se generó dentro de las actividades académicas.

Instrumentos

Como instrumentos de recolección de datos se desarrolló una prueba de lectoescritura compuesta por tres secciones de forma integral y secuencial, dentro de la primera sección del instrumento se construyó con un protocolo de lectura cronometrada de un texto de 150 palabras, dentro de este también se registró errores relacionados a la decodificación, la segunda sección permitió medir la comprensión lectora mediante un texto de 400 palabras y 12 preguntas en relación a la misma y la tercera sección direccionada a una tarea de producción escrita argumentativa, todos ellos con la finalidad de tener una comprensión de las habilidades de lectura y escritura en los estudiantes. De manera adicional se aplicó un cuestionario de percepciones sobre el uso de la IA bajo escala de Likert para tener opiniones y percepciones por parte de los estudiantes.

Validez y confiabilidad de los instrumentos

Dentro del proceso de validación de instrumentos de investigación se realizó mediante la actividad de validez de contenido con la participación de cinco expertos de investigación dentro de las áreas de lengua escrita y tecnología educativa, orientados a determinar criterios de pertinencia, claridad y suficiencia por cada pregunta desarrollada dentro de los instrumentos, para luego aplicarse el estadístico de V de Aiken (1985) en el que se obtuvo un valor de 0,87, lo que arroja un resultado positivo por encima de los datos requeridos. De la misma forma la confiabilidad mediante el coeficiente de alfa de Cronbach con valores de 0,78, para fluidez lectora con 0,83 y para comprensión 0,81 dando resultados favorables para la investigación considerado bueno según los criterios de Koo y Li (2016).

Herramientas de IA utilizadas

El grupo experimental trabajó secuencialmente con tres plataformas durante doce semanas. En las primeras cuatro, los estudiantes usaron Duolingo ABC para ejercicios de conciencia fonológica y decodificación. De la semana cinco a la ocho, emplearon Speechify, una herramienta de lectura en

voz alta asistida por IA que permite ajustar la velocidad y resaltar palabras en tiempo real, orientada a mejorar la fluidez lectora.

En lo correspondiente a las semanas de la nueve a la doce, se utilizaron IAs entre las que destaco es ChatGPT, a la cual se le aplicaron Prompts para a través de esta proporcionar retroalimentación sobre los borradores de escritura argumentativa. Estas actividades no se desarrollaron de forma abierta sino mas bien mediante formularios y criterios de evaluación a lo que el sistema daba una respuesta personalizada acorde a la información presentada., con la finalidad de brindar retroalimentación como lo menciona Chen y Cheng (2008).

Procedimiento

La investigación se ejecutó en seis etapas. Primera, revisión bibliográfica y diseño de los instrumentos (cuatro semanas). Segunda, validación por juicio de expertos y prueba piloto con diez estudiantes ajenos a la muestra (dos semanas). Tercera, aplicación de la preprueba a ambos grupos en sesión simultánea de 90 minutos. Cuarta, implementación de la intervención durante doce semanas, con dos sesiones semanales para el grupo experimental y la programación habitual de Lengua y Literatura para el grupo de control. Quinta, aplicación de la posprueba bajo las mismas condiciones de la preprueba. Sexta, procesamiento estadístico de los datos e interpretación de resultados.

Análisis estadístico

Los datos se procesaron con SPSS v.28 y se cruzaron con tablas dinámicas de Microsoft Excel para verificación. Por la finalidad de la investigación se aplicó estadísticos descriptivos como la media, desviación estándar, percentiles y rangos intercuartílicos aplicado a los diferentes grupos de estudio y los momentos en los que se aplicó la investigación. De la misma forma como se realizó una toma de datos antes y después se aplicó la técnica estadística de T de Student para muestras independientes, con su aplicación de manera inicial de pruebas de normalidad mediante Shapiro-Wilk y homogeneidad de varianzas (Levene). Adicionalmente se aplicó prueba estadística de ANCOVA, con la finalidad de medir el efecto de la d de Cohen.

Consideraciones éticas

El estudio se ajustó a los principios éticos de investigación educativa con menores establecidos por la UNESCO (2021) y al código ético de la American Educational Research Association

(AERA, 2011). Antes de iniciar el trabajo de campo, la propuesta fue presentada al rectorado de la institución participante, que emitió autorización por escrito para la realización del estudio.

Para contar con criterios éticos y la investigación direccionarse a menores de edad se procedió a la firma de un consentimiento informado firmado por los padres de familia con la finalidad de brindarles información de la investigación y su finalidad, así como mencionarles la voluntariedad de desarrollar estas actividades.

Se procedió a codificar la información de los estudiantes para evitar exponer datos que puedan causar vulnerabilidad de los estudiantes, adicionalmente los datos solo fueron utilizados con finalidades académicas y de la investigación de forma particular y no se compartió esta información con terceras personas, por otro lado, no se solicitó ninguna información de carácter sensible a los estudiantes.

Resultados

Comparabilidad inicial de los grupos

La Tabla 2 presenta los estadísticos descriptivos de la preprueba para ambos grupos. Las puntuaciones totales fueron muy similares: el grupo experimental obtuvo una media de 10,8 (DT = 2,31) y el grupo de control, 10,5 (DT = 2,18). La prueba t no detectó diferencias estadísticamente significativas antes de la intervención ($t(48) = 0,49$; $p = 0,63$), lo que confirma la equivalencia inicial entre grupos. La prueba de Shapiro-Wilk verificó distribución normal en ambas muestras ($p > 0,05$) y la de Levene confirmó homogeneidad de varianzas ($F(1, 48) = 0,31$; $p = 0,58$).

Tabla 2

Estadísticos descriptivos del pretest por grupo (n = 50)

Variable	GE M (DT)	GC M (DT)	t(48)	p
Fluidez lectora	10,2 (2,44)	9,9 (2,38)	0,44	0,66
Comprensión literal	11,4 (2,61)	11,1 (2,55)	0,41	0,68
Comprensión inferencial	10,1 (2,22)	9,8 (2,19)	0,47	0,64
Escritura argumentativa	10,6 (2,38)	10,3 (2,21)	0,46	0,65
Puntuación total	10,8 (2,31)	10,5 (2,18)	0,49	0,63

Nota. GE = grupo experimental; GC = grupo de control. Puntuaciones sobre 20.

Resultados del postest

Tras las doce semanas de intervención, el grupo experimental registró ganancias sustanciales en todas las dimensiones evaluadas. La Tabla 3 muestra la comparación postest entre grupos.

Tabla 3

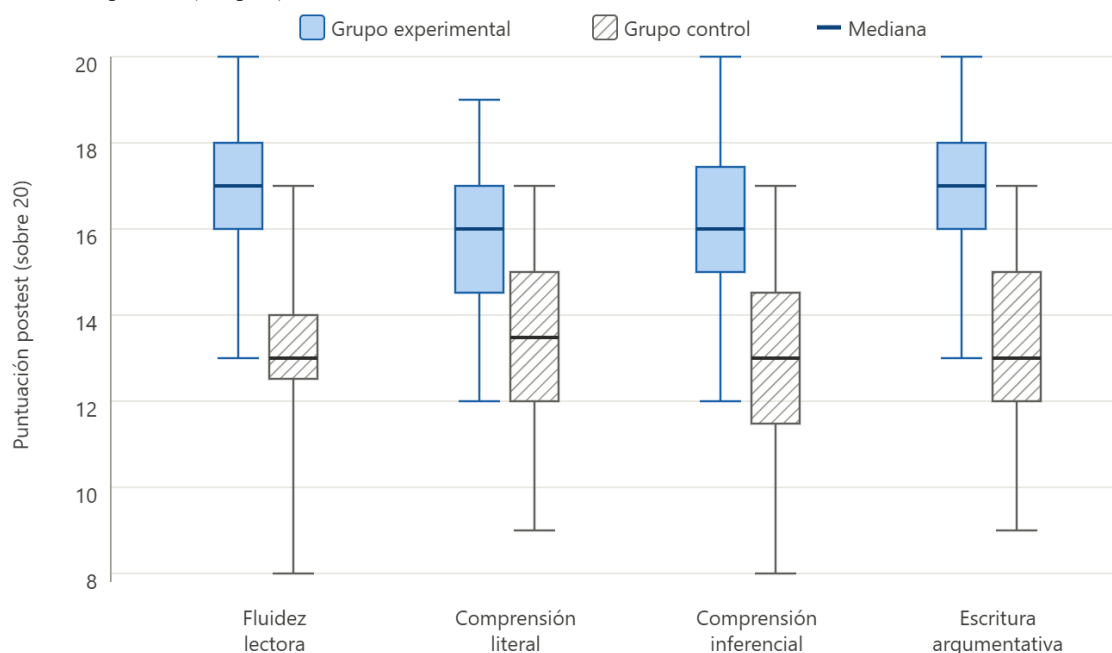
Comparación de medias en el postest entre grupo experimental y control (n = 50)

Variable	GE M (DT)	GC M (DT)	t(48)	p	d Cohen
Fluidez lectora	17,1 (1,88)	12,8 (2,22)	7,28	< 0,001	2,12
Comprensión literal	15,6 (1,95)	13,4 (2,11)	3,76	0,001	1,09
Comprensión inferencial	15,9 (1,91)	12,9 (2,08)	5,27	< 0,001	1,53
Escritura argumentativa	16,8 (1,94)	13,3 (2,17)	5,91	< 0,001	1,72
Puntuación total	16,4 (1,92)	13,1 (2,14)	6,21	< 0,001	1,64

Nota. GE = grupo experimental; GC = grupo de control. Puntuaciones sobre 20. El ANCOVA con el pretest como covariable confirmó el efecto significativo de la intervención sobre el postest total ($F(1, 47) = 38,62$; $p < 0,001$; η^2 parcial = 0,45).

Figura 1

Distribución postest (boxplot)



Nota. Los rectángulos indican el rango intercuartílico (P25–P75); la línea interior señala la mediana; los bigotes se extienden hasta los valores mínimo y máximo dentro de $1,5 \times \text{RIC}$. GE = grupo experimental ($n = 25$); GC = grupo control ($n = 25$). Puntuaciones sobre 20. Las cajas del GC se distinguen mediante patrón de rayas como clave secundaria al color.

Prueba de hipótesis

El contraste de hipótesis se realizó sobre la puntuación total del postest mediante prueba t para muestras independientes, con un nivel de significación $\alpha = 0,05$. Los resultados fueron evidentemente significativos con un p valor $< 0,001$ lo que determina significancia estadística. De manera seguida al análisis de covarianza para determinar la diferencia entre los tiempos de toma de información con un p valor $< 0,001$; η^2 parcial = 0,45 lo que indica que el 45 % de la varianza del postest, de forma adicional el tamaño del efecto con una d de Cohen de 1,64 tiene un rango de magnitud muy alto determinando el impacto de la investigación.

Con base en estos resultados se rechaza la hipótesis nula ($H_0: \mu_{GE} = \mu_{GC}$) y se acepta la hipótesis alterna ($H_1: \mu_{GE} > \mu_{GC}$). En términos sustantivos: los estudiantes que participaron del programa mediado por herramientas de IA obtuvieron puntuaciones de lectoescritura significativamente superiores a las de sus pares con instrucción convencional, y la magnitud de esa diferencia excede holgadamente el umbral de relevancia práctica.

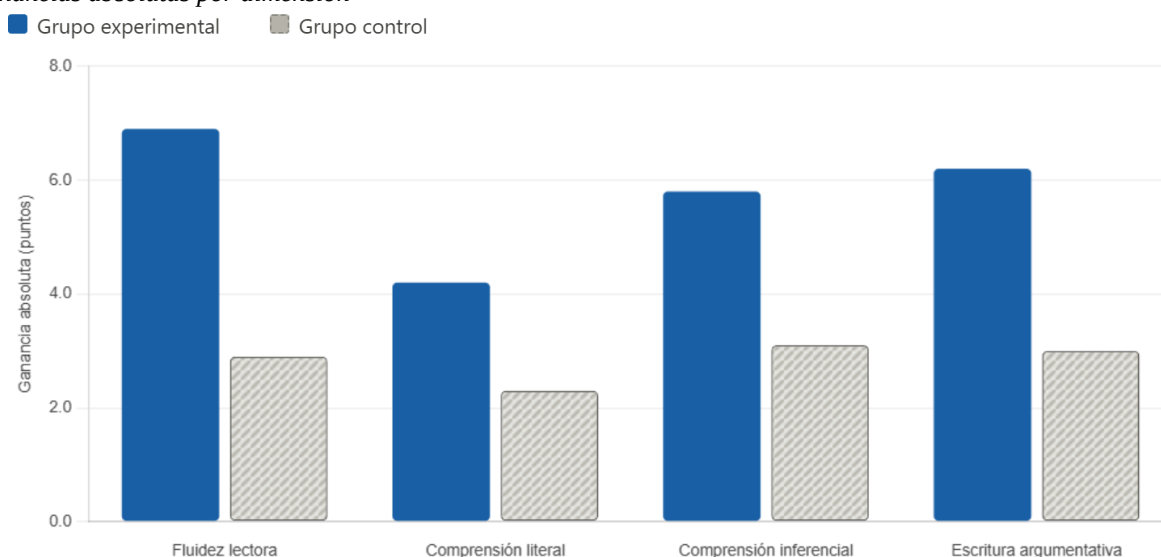
Patrones de ganancia diferencial

La ganancia absoluta del grupo experimental de 10,8 a 16,4 puntos; $\Delta = 5,6$ superó con amplitud la del grupo de control de 10,5 a 13,1; $\Delta = 2,6$. Se evidencia que en este mismo momento el grupo experimental presentó una ganancia significativa frente al grupo de control de casi el doble.

A estos datos mencionados se observa que la fluidez lectora mostrada en el estudio evidenció la mayor diferencia entre grupos con una d de Cohen de 2,12, de manera seguida en lo referente a la escritura argumentativa se obtuvo un valor de d de Cohen 1,72. Estos datos ponen en evidencia que los estudiantes tienen alta capacidad en lo referente a la fluidez lectora, sin embargo existen diferencias con los resultados obtenidos en comprensión literal con un valor de d de Cohen 1,09, aunque es un valor alto y marca con claridad el impacto que ha tenido la investigación en la mejora de la comprensión de lectura y escritura en los sujetos experimentales.

Figura 2

Ganancias absolutas por dimensión



Nota. Ganancia absoluta = Media posttest – Media pretest. GE = grupo experimental (n = 25); GC = grupo control (n = 25). Puntuaciones sobre 20. Las barras del GE utilizan relleno sólido; las del GC, relleno con patrón de rayas para facilitar la discriminación sin depender solo del color.

Dentro del grupo experimental, el análisis por percentiles reveló que los estudiantes ubicados en el tercil inferior del pretest obtuvieron ganancias proporcionalmente mayores (Δ promedio = 7,1 puntos) que los del tercil superior (Δ promedio = 3,8 puntos). Estos resultados ponen en evidencia que cuando los estudiantes empiezan con grandes deficiencias se orienta a que estos tengan mayor nivel de retroalimentación y por ende el efecto en la mejora se observa de forma notoria.

Por otro lado, el instrumento orientado a recoger las percepciones de ellos estudiantes identificó a un 84% de estudiantes que presentaron criterios de acuerdo con el uso de la IA y su aporte a la mejora de la escritura, frente al 60% de estudiantes que utilizaron metodologías tradicionales. En donde se identificó el mayor grado de diferencia se relaciona a la percepción de autonomía cuando se presentaban percepciones diferentes por parte de los estudiantes.

Discusión

Los resultados dialogan con la tendencia que Suspitsyna et al. (2023) documentaron en su metaanálisis, aunque el tamaño del efecto reportado aquí ($d = 1,64$ en puntuación total) excede el promedio del corpus reunido en aquel trabajo ($g = 0,72$). Se observa una ventaja de esta investigación frente a otras que han buscado abordar este tema de estudio ya que se desarrolló una intervención de doce semanas frente a investigaciones previas que en promedio tienen siete

semanas. Adicionalmente, en el estudio se utilizó tres instrumentos de recolección distintos dando un mayor entendimiento sobre el tema de estudio.

El mayor tamaño de efecto en fluidez lectora ($d = 2,12$) era esperable en parte. Rasinski (2022) documentó que los sistemas de lectura asistida por audio producen mejoras rápidas y medibles en velocidad y precisión lectora porque eliminan la barrera del desconocimiento decodificativo. Lo que no era evidente a priori era la magnitud del efecto. Speechify opera con reconocimiento de voz y ajuste de velocidad en tiempo real, lo que crea un andamiaje externo que se retira gradualmente conforme el estudiante gana autonomía, un mecanismo análogo en términos funcionales al concepto vygotskiano de zona de desarrollo próximo, aunque su implementación tecnológica difiere sustantivamente de lo que el autor concibió.

De manera seguida se observó que la escritura argumentativa presento un efecto muy alto de 1,72 y como lo menciona Cassany (2021), es uno de los aspectos de mayor dificultad en los sistemas educativos hispanohablantes, ya que esta es una de las competencias de mayor dificultad de desarrollo y por otro lado las características del sistema educativos de estos países se orientan de mayor forma a la copia y descripción mas que a un desarrolla analítico. Esto se debe a que la retroalimentación por parte de una IA como el caso de ChatGPT como un elemento potenciador de identificación y retroalimentación de ausencias de argumentos de conexión e inconsistencias dentro de los escritos. Rodríguez-García et al. (2023) identificaron estudios desarrollados en España con una buena calidad en las mejoras de escritura argumentativa.

Este estudio también presenta resultados muy interesantes debido a que se mejoró la riqueza léxica en los resultados posttest en el grupo experimental con una mejoría del 18% referente a los datos iniciales, lo que pone en evidencia que la retroalimentación a los estudiantes puede ser un potenciador en la calidad léxica de los estudiantes.

La investigación presenta datos de interés en referencia a la comprensión literal con valores de 1,09 y la comprensión inferencial con valores de 1,53, lo que evidencia que el uso de la IA puede ser un potenciador de la mejora de las habilidades relacionadas a el análisis de los datos y como esta puede lograr que se solventen falencias en la comprensión lectora y por ende ayuda a que el estudiante mejore estas competencias. Estos datos van de la mano con Holmes et al. (2022) sugieren sobre la complementariedad entre IA y enseñanza directa: es decir la IA y tecnología tiene la capacidad de cubrir estas falencias que con la educación tradicional no se pueden solventar.

Por otro lado, los estudiantes que mejores resultados y mayores niveles de ganancia lograron es el tercil inferior como lo menciona Woolf (2010) los sistemas de retroalimentación mejoran de forma más significativa a estudiantes que ingresan con habilidades muy bajas y que a través de estas estrategias mejoran de forma significativa. Morales-López y Hernández (2022) también identificaron estas ventajas en estos grupos que presentan dificultades de forma inicial.

Las percepciones que se tomo por parte de los estudiantes demostraron que cuando un estudiante comprende lo que hace un sistema y sabe su finalidad se logran mejores resultados (Luckin et al. 2016), por otro lado se evidencio diferencias en los estudiantes en cuanto al uso de la IA, por un lado los estudiantes que tenían mayor capacidad no requerían mayor retroalimentación por parte de la IA a diferencia de los que tenía falencias requerían que la IA los retroalimente de forma secuencial.

Conclusiones

Los estudiantes de Décimo Grado que trabajaron con herramientas de IA durante doce semanas alcanzaron, en todas las dimensiones evaluadas, puntuaciones significativamente superiores a las de sus pares con instrucción convencional. La magnitud del efecto resultó alta algo poco frecuente en investigación educativa, lo que respalda la pertinencia del programa de enseñanza mediado por IA en contextos similares al estudiado.

La comparación pre y posttest confirmó la equivalencia inicial de los grupos y la ventaja del experimental tras la intervención; la fluidez lectora y la escritura argumentativa fueron las dimensiones con mayor ganancia diferencial; y las percepciones estudiantiles, además de mostrar valoración mayoritariamente positiva, revelaron que la utilidad subjetiva del andamiaje varía con el nivel inicial, lo que tiene consecuencias para el diseño de futuras intervenciones.

La implicación pedagógica más directa es que las herramientas de IA no sustituyen al docente: amplifican su acción. El diseño instruccional que orientó el uso de cada plataforma lo elaboró el profesor, quien también medió la interacción de los estudiantes con los sistemas. Sin esa estructura, los resultados habrían sido probablemente menores.

Para las instituciones educativas ecuatorianas, el hallazgo tiene dos consecuencias prácticas. Se destaca la importancia de generar inversión en capacitación docente con especial énfasis en didáctica mediada por la IA, es decir uso de tecnología, pero con finalidades didácticas. Por otro

lado, construir normativas y protocolos que regulen el uso de la IA de forma pedagógica con la finalidad de lograr en estudiantes mejores resultados académicos y uso de la IA de forma ética.

Referencias

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Alamargot, D., & Chanquoy, L. (2001). *Through the models of writing*. Kluwer Academic Publishers.
- American Educational Research Association. (2011). *Code of ethics*. AERA.
- Ander-Egg, E. (2011). *Aprender a investigar: nociones básicas para la investigación social*. Editorial Brujas.
- Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. En J. A. Larusson y B. White (Eds.), *Learning analytics* (pp. 61–75). Springer. https://doi.org/10.1007/978-1-4614-3305-7_4
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. <https://doi.org/10.1145/3442188.3445922>
- Bernal, C. A. (2016). *Metodología de la investigación* (4.ª ed.). Pearson.
- Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research*. Rand McNally.
- Cassany, D. (2021). La escritura extensiva: propuesta para la didáctica de la composición. *Revista de Educación*, 393, 11–34. <https://doi.org/10.4438/1988-592X-RE-2021-393-486>
- Chai, C. S., Koh, J. H. L., & Tsai, C. C. (2021). A review of the quantitative measures of technological pedagogical content knowledge (TPACK). En T. Foulger y K. Graziano (Eds.), *Technological pedagogical content knowledge* (pp. 87–106). Springer.
- Chen, C. F. E., & Cheng, W. Y. E. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2.ª ed.). Lawrence Erlbaum Associates.
- Coll, C. (2018). La personalización del aprendizaje escolar: una exigencia de la nueva ecología del aprendizaje. *Revista Iberoamericana de Educación*, 76, 13–36. <https://doi.org/10.35362/rie760175>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. <https://doi.org/10.1007/BF02310555>
- Defior, S., & Serrano, F. (2021). Conciencia fonológica e intervención en dislexia: una revisión sistemática. *Revista de Logopedia, Foniatría y Audiología*, 41(3), 121–133. <https://doi.org/10.1016/j.rlfa.2021.01.002>
- Defior, S., Serrano, F., & Marín-Cano, M. J. (2022). Dificultades lectoras y escritoras: bases neuropsicológicas y programas de intervención. *Pirámide*.
- Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed and what remains uncertain. *International Journal of Educational Technology in Higher Education*, 14(1), 9. <https://doi.org/10.1186/s41239-017-0042-5>
- Ferreiro, E., & Teberosky, A. (1979). *Los sistemas de escritura en el desarrollo del niño*. Siglo XXI.

- Gómez-Merino, A. I., Ruiz-Palmero, J., & Sánchez-Rodríguez, J. (2023). Inteligencia artificial en la educación lingüística: revisión sistemática de la literatura. *Revista Complutense de Educación*, 34(2), 329–342. <https://doi.org/10.5209/rced.78645>
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading, and reading disability. *Remedial and Special Education*, 7(1), 6–10. <https://doi.org/10.1177/074193258600700104>
- Guthrie, J. T., Wigfield, A., Barbosa, P., Perencevich, K. C., Taboada, A., Davis, M. H., Scaffiddi, N. T., & Tonks, S. (2004). Increasing reading comprehension and engagement through concept-oriented reading instruction. *Journal of Educational Psychology*, 96(3), 403–423. <https://doi.org/10.1037/0022-0663.96.3.403>
- Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. En C. M. Levy y S. Ransdell (Eds.), *The science of writing* (pp. 1–27). Lawrence Erlbaum Associates.
- Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. En L. W. Gregg y E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). Lawrence Erlbaum Associates.
- Hernández-Sampieri, R., & Mendoza, C. P. (2018). *Metodología de la investigación: las rutas cuantitativa, cualitativa y mixta*. McGraw-Hill.
- Hmelo-Silver, C. E., Liu, L., & Jordan, R. (2022). Scaffolding the development of inquiry skills in game-based learning environments. *Journal of Science Education and Technology*, 30(2), 194–210. <https://doi.org/10.1007/s10956-020-09862-0>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. 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), 20–45. <https://doi.org/10.18608/jla.2022.7227>
- Instituto Nacional de Evaluación Educativa. (2022). *Resultados de las evaluaciones educativas nacionales 2021–2022*. Ministerio de Educación del Ecuador.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 22(140), 1–55.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson.
- Morales-López, H., & Hernández, C. A. (2022). Escritura digital y retroalimentación automatizada en educación secundaria: un estudio exploratorio. *Pixel-Bit: Revista de Medios y Educación*, 63, 131–152. <https://doi.org/10.12795/pixelbit.91305>
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Rasinski, T. V. (2022). Fluency and comprehension: An important connection. *Reading Today*, 40(1), 26–28.
- Rodríguez-García, A. M., Raso-Sánchez, F., & Ruiz-Palmero, J. (2023). Evaluación automatizada de textos académicos mediante inteligencia artificial: efectos sobre la coherencia y cohesión textual en Educación Secundaria. *Educación XX1*, 26(1), 167–192. <https://doi.org/10.5944/educxx1.32148>
- Russell, S., & Norvig, P. (2021). *Artificial intelligence: A modern approach* (4.^a ed.). Pearson.

- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. En S. B. Neuman y D. K. Dickinson (Eds.), *Handbook of early literacy research* (pp. 97–110). Guilford Press.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Stevenson, M., & Phakiti, A. (2019). The effects of computer-generated feedback on the quality of writing. *Assessing Writing*, 19, 51–65. <https://doi.org/10.1016/j.asw.2013.11.007>
- Suspitsyna, T., Rueda, M., & Dillard, C. (2023). Artificial intelligence and student writing: A meta-analysis of intervention studies. *Computers & Education*, 196, 104736. <https://doi.org/10.1016/j.compedu.2022.104736>
- UNESCO. (2021). *Los aprendizajes fundamentales en América Latina y el Caribe: Estudio Regional Comparativo y Explicativo (ERCE 2019)*. Oficina Regional de Educación para América Latina y el Caribe.
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist*, 46(4), 197–221. <https://doi.org/10.1080/00461520.2011.611369>
- Warschauer, M., & Grimes, D. (2008). Automated writing assessment in the classroom. *Pedagogies: An International Journal*, 3(1), 22–36. <https://doi.org/10.1080/15544800701771580>
- Woolf, B. P. (2010). *Building intelligent interactive tutors: Student-centered strategies for e-learning*. Morgan Kaufmann.
- 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>