

Transformación de la evaluación educativa mediante herramientas digitales e inteligencia artificial***Transforming educational assessment through digital tools and artificial intelligence***

Cedeño Murillo Bel Anahis, Quiñonez García Lilibeth Ximena, Ronquillo Valencia Leandro David, Núñez Michuy Carlos Manuel.

**CIENCIA, TECNOLOGÍA Y
SOCIEDAD**

Julio - diciembre, V°7-N°2; 2026

Recibido: 15-08-2026

Aceptado: 17-08-2026

Publicado: 31-12-2026

PAIS

- Quevedo - Ecuador
- Quevedo - Ecuador
- Quevedo - Ecuador
- Quevedo - Ecuador

INSTITUCION

- Universidad Técnica Estatal de Quevedo, UTEQ.
- Universidad Técnica Estatal de Quevedo, UTEQ.
- Universidad Técnica Estatal de Quevedo, UTEQ.
- Universidad Técnica Estatal de Quevedo, UTEQ.

CORREO:

- ✉ nonaq@uteq.edu.ec
- ✉ vegag@uteq.edu.ec
- ✉ itigselema@uteq.edu.ec
- ✉ cnunezm@uteq.edu.ec

ORCID:

- ✉ <https://orcid.org/0000-0002-7205-7562>
- ✉ <https://orcid.org/0000-0001-5098-4608>
- ✉ <https://orcid.org/0000-0002-3668-5684>
- ✉ <https://orcid.org/0000-0003-2298-7697>

FORMATO DE CITA APA

Oña, N. Vega, G. Tigselema, I. & Núñez, C. (2026). *Transformación de la evaluación educativa mediante herramientas digitales e inteligencia artificial*. Revista G-ner@ndo, V°7 (N°2.). P. 3442 – 3459.

Resumen

La incorporación de herramientas digitales e inteligencia artificial (IA) abre nuevas posibilidades para transformar la evaluación educativa, especialmente mediante una retroalimentación más oportuna y ajustada a las necesidades de los estudiantes. Este estudio tuvo como objetivo analizar el impacto de estas tecnologías en los procesos evaluativos y, a partir de los hallazgos, proponer un modelo orientado a la personalización del aprendizaje y la equidad evaluativa en el contexto ecuatoriano. Se desarrolló una investigación de enfoque mixto, con diseño secuencial exploratorio, no experimental y transeccional, de alcance descriptivo-correlacional y propositivo. Participaron 20 docentes y 30 estudiantes de cinco instituciones educativas. Se utilizaron revisión de literatura, encuesta, entrevistas semiestructuradas y análisis documental. Los resultados muestran una percepción favorable hacia el uso de estas tecnologías, principalmente por sus posibilidades para mejorar la retroalimentación y personalizar el aprendizaje, aunque persisten inquietudes relacionadas con el acceso, la privacidad, la transparencia y la intervención humana. En los estudiantes se encontró una asociación positiva entre retroalimentación y personalización mediada por tecnología y desempeño académico ($p = 0,591$; $p < 0,001$). A partir de estos hallazgos se propone un modelo que integra diagnóstico, seguimiento formativo, retroalimentación personalizada y gobernanza ética de los datos. Se concluye que la tecnología aporta valor a la evaluación cuando se integra con criterios pedagógicos, equidad de acceso y acompañamiento docente.

Palabras clave: evaluación educativa; inteligencia artificial; herramientas digitales; retroalimentación formativa; personalización; equidad.

Abstract

The integration of digital tools and artificial intelligence (AI) is creating new opportunities to transform educational assessment, particularly through more timely feedback tailored to students' needs. This study aimed to analyze the impact of these technologies on assessment practices and based on the findings, propose a model focused on personalized learning and assessment equity in the Ecuadorian context. A mixed-methods approach was adopted, using a sequential exploratory, non-experimental, cross-sectional design with a descriptive-correlational and propositional scope. The study involved 20 teachers and 30 students from five educational institutions. Data were collected through a literature review, surveys, semi-structured interviews, and document analysis. The findings showed an overall positive perception of these technologies, particularly regarding their potential to improve feedback and personalize learning. However, concerns remained about access, privacy, transparency, and the preservation of human involvement in assessment. Among students, a positive association was found between technology-mediated feedback and personalization and academic performance ($p = 0.591$; $p < 0.001$). Based on these findings, an assessment model was proposed that integrates diagnosis, formative monitoring, personalized feedback, and ethical data governance. The study concludes that technology adds meaningful value to assessment when combined with sound pedagogical criteria, equitable access, and teacher guidance.

Keywords: educational assessment; artificial intelligence; digital tools; formative feedback; personalization; equity.

Introducción

La evaluación educativa ha sido concebida históricamente como el mecanismo institucional por excelencia para medir y certificar el aprendizaje. Sin embargo, en las últimas dos décadas su función ha experimentado una reconfiguración epistemológica profunda: de un modelo predominantemente sumativo y clasificador se transita hacia paradigmas que privilegian el proceso, la retroalimentación y la construcción activa del conocimiento (Black y Wiliam, 2009). Este desplazamiento coincide con la irrupción de herramientas digitales e inteligencia artificial en los entornos educativos, generando lo que algunos autores denominan una «disrupción pedagógica evaluativa» (Holmes et al., 2019).

En el contexto ecuatoriano, esta transformación se desarrolla en un escenario de tensiones: infraestructuras tecnológicas heterogéneas, deudas históricas de equidad y una comunidad docente que demanda formación continua para incorporar con sentido crítico las nuevas tecnologías. Según el Informe de Competencias Digitales del Ministerio de Educación del Ecuador (2022), más del 60 % de los docentes de secundaria no ha recibido formación formal en plataformas digitales evaluativas, lo que limita el impacto pedagógico de las herramientas disponibles.

La IA ha comenzado a redefinir las posibilidades de la evaluación formativa mediante sistemas capaces de analizar patrones de aprendizaje y generar retroalimentación personalizada en tiempo real (Zawacki et al., 2019), aunque su implementación conlleva riesgos éticos que la comunidad educativa apenas comienza a debatir con rigor: sesgo algorítmico, erosión de la agencia docente y amenazas a la integridad académica (Luckin et al., 2016; Pinto et al., 2022). En esa línea, Fandiño y Muñoz (2026) confirman que la IA ha modificado sustancialmente las prácticas evaluativas en la educación superior, señalando además la vulneración de la privacidad de los datos y la deshumanización de la evaluación como riesgos adicionales, frente a oportunidades como la personalización, la retroalimentación automatizada y la mayor accesibilidad lo que exige un marco de integración ética y estratégica antes que una adopción acrítica. El análisis bibliométrico de Galán et al. (2025) respalda esta lectura al evidenciar un crecimiento sostenido de la producción científica sobre IA y evaluación en educación superior entre 2014 y 2024, con una aceleración marcada desde 2022, lo que confirma que el fenómeno estudiado se encuentra aún en plena consolidación teórica.

Ante este panorama, el presente estudio se plantea tres preguntas centrales: ¿Qué herramientas digitales e IA están siendo efectivamente utilizadas en los procesos evaluativos de las instituciones educativas ecuatorianas? ¿Cómo perciben docentes y estudiantes los beneficios, limitaciones y desafíos de dicha integración? ¿Qué características debe tener un

modelo evaluativo mediado por tecnología que responda a los principios de equidad, personalización y retroalimentación formativa, adaptado al contexto ecuatoriano? Responder estas interrogantes tiene relevancia científica y una urgencia práctica: las decisiones que hoy adopten las instituciones educativas en materia de evaluación digital determinarán, en gran medida, las oportunidades de aprendizaje de las próximas generaciones.

Para dar respuestas a las preguntas de investigación planteadas, se propuso como objetivo general de esta investigación; analizar el impacto de la incorporación de herramientas digitales e inteligencia artificial en la transformación de los procesos de evaluación educativa, con el fin de proponer un modelo pedagógico innovador que potencie la retroalimentación oportuna, la personalización del aprendizaje y la equidad evaluativa en contextos de educación formal del Ecuador. Para alcanzarlo, se plantearon tres objetivos específicos: identificar las principales herramientas digitales e inteligencias artificiales aplicadas en procesos de evaluación educativa a nivel nacional e internacional, valorando su pertinencia pedagógica y accesibilidad en el contexto latinoamericano; determinar la percepción de docentes y estudiantes respecto a los beneficios, limitaciones y desafíos que representa la integración de la tecnología digital en las prácticas evaluativas actuales en instituciones de educación secundaria y superior; y diseñar una propuesta de modelo evaluativo mediado por herramientas digitales e inteligencia artificial que responda a los principios de equidad, personalización y retroalimentación formativa continua, adaptada a las realidades educativas del Ecuador.

La evaluación formativa y sumativa como procesos complementarios

La evaluación formativa se concibe como un proceso deliberado mediante el cual docentes y estudiantes recaban evidencia sobre el aprendizaje para tomar decisiones pedagógicas que lo mejoren mientras aún está en curso, situando la retroalimentación y la autorregulación del estudiante en el centro del acto evaluativo. Black y Wiliam (2009) sostienen que una evaluación solo es formativa cuando la evidencia recogida se utiliza efectivamente para modificar la enseñanza, y no como un mero registro de desempeño; esta perspectiva es clave para comprender el papel de las herramientas digitales e IA, pues amplían la capacidad de generar evidencia continua sobre el aprendizaje, aunque Valdez et al. (2023) enfatizan que la calidad de la retroalimentación, claridad, oportunidad y orientación hacia la mejora, determina el impacto formativo más que la sofisticación técnica del instrumento. La evaluación sumativa, en cambio, cumple una función complementaria: certificar mediante un juicio final el nivel de logro alcanzado al término de un período o programa. Dautović (2025) advierte que esta modalidad corre el riesgo de reducir la evaluación a un ejercicio de clasificación cuando se aplica de forma aislada, y que los sistemas digitales de calificación automatizada, si bien aportan consistencia y

rapidez, pueden introducir rigidez cuando privilegian respuestas estandarizadas sobre formas legítimas de demostrar comprensión; esta tensión es especialmente relevante para las instituciones ecuatorianas, que deben equilibrar la eficiencia de las plataformas digitales con la validez pedagógica y la justicia evaluativa hacia la diversidad de trayectorias de aprendizaje.

Inteligencia artificial aplicada a la educación (IAEd)

La inteligencia artificial aplicada a la educación integra aprendizaje automático, procesamiento de lenguaje natural y analítica de datos para apoyar la enseñanza, el aprendizaje y la evaluación. Zawacki-Richter et al. (2019) concluyen, tras una revisión sistemática, que sus aplicaciones se concentran en cuatro áreas, perfilado y predicción del estudiante, evaluación y retroalimentación automatizada, sistemas adaptativos y agentes conversacionales de apoyo docente y advierten que persiste un vacío en la investigación sobre las implicaciones pedagógicas y éticas de estas herramientas. Los mismos autores señalan que gran parte de esta literatura proviene de contextos anglosajones y asiáticos con alta disponibilidad tecnológica, mientras que las realidades de países en desarrollo, incluidos los latinoamericanos, están notablemente subrepresentadas, lo que refuerza la pertinencia de investigaciones situadas como la presente.

La retroalimentación como mecanismo de mejora del aprendizaje

La retroalimentación es uno de los factores con mayor capacidad de influir en el rendimiento académico cuando se implementa adecuadamente. Hattie y Timperley (2007) proponen que debe responder a tres preguntas ¿hacia dónde voy?, ¿cómo voy progresando? y ¿cuáles son los siguientes pasos? y distinguen cuatro niveles en los que puede operar: la tarea, el proceso, la autorregulación y el yo personal, siendo este último, como los elogios genéricos, el menos eficaz para el aprendizaje. Esta clasificación resulta útil para valorar la calidad de las retroalimentaciones generadas por sistemas de IA, al permitir distinguir entre señales genéricas de desempeño y orientaciones sobre el proceso cognitivo y la autorregulación; los autores enfatizan que la retroalimentación de mayor impacto es la que reduce la brecha entre el desempeño actual y la meta de aprendizaje, criterio empleado en este estudio para valorar la pertinencia pedagógica de las plataformas digitales analizadas.

Personalización del aprendizaje y brecha digital

La personalización del aprendizaje refiere a la capacidad de adaptar contenidos, ritmo y estrategias didácticas a las características individuales de cada estudiante, superando el modelo homogéneo de enseñanza. Luckin et al. (2016) sostienen que los sistemas de aprendizaje adaptativo pueden modelar el estado de conocimiento del estudiante en tiempo real y ajustar la dificultad en consecuencia, aunque advierten que este potencial no debe idealizarse, pues su

eficacia depende de la calidad de los datos de entrenamiento, la mediación docente y el acceso real a dispositivos y conectividad estable. Frente a este optimismo moderado, Selwyn (2016) introduce una lectura más crítica al cuestionar si dichas condiciones de acceso se verifican efectivamente en contextos de alta desigualdad estructural como el ecuatoriano: buena parte del discurso sobre tecnología educativa asume condiciones de acceso universal que rara vez se cumplen en la práctica, lo que puede profundizar las brechas de aprendizaje existentes en lugar de reducirlas. Esta perspectiva es central para Ecuador, donde coexisten instituciones con infraestructura tecnológica consolidada junto a otras con recursos limitados, particularmente en zonas rurales; la equidad educativa mediada por tecnología no se garantiza con la sola disponibilidad de herramientas, sino que requiere políticas institucionales deliberadas de formación docente y provisión de acceso, sin las cuales la innovación corre el riesgo de beneficiar preferentemente a quienes ya cuentan con mayores ventajas socioeconómicas.

Ética algorítmica e integridad académica en la evaluación digital

La incorporación de sistemas de inteligencia artificial en la evaluación plantea dilemas éticos que trascienden la eficiencia técnica de las herramientas. Holmes et al. (2019) señalan que los algoritmos de evaluación automatizada pueden reproducir y amplificar sesgos presentes en los datos de entrenamiento, afectando de manera desproporcionada a estudiantes de determinados grupos socioculturales o lingüísticos, y advierten sobre la falta de transparencia de muchos sistemas comerciales, cuyos criterios de calificación no siempre resultan auditables por docentes e instituciones. Identifican además la erosión de la agencia docente cuando las decisiones evaluativas se delegan de manera acrítica en sistemas automatizados, así como riesgos para la integridad académica asociados al uso de inteligencia artificial generativa por parte de los propios estudiantes, y concluyen que la implementación responsable de estas tecnologías exige marcos normativos claros, capacitación docente en alfabetización algorítmica y mecanismos institucionales de supervisión humana, lineamientos directamente pertinentes para el modelo evaluativo propuesto en esta investigación.

Síntesis crítica del marco teórico

Las categorías teóricas revisadas, evaluación formativa y sumativa, inteligencia artificial aplicada a la educación, retroalimentación, personalización, brecha digital y ética algorítmica no operan de manera aislada, sino que se articulan en tensiones productivas para el objeto de estudio. Mientras Black y Wiliam (2009) y Hattie y Timperley (2007) aportan los criterios pedagógicos para juzgar la calidad de una retroalimentación mediada por tecnología, Luckin et al. (2016) y Zawacki-Richter et al. (2019) documentan el potencial técnico de la IA para generarla a escala; sin embargo, Selwyn (2016) y Holmes et al. (2019) advierten que dicho potencial no es

neutro ni automáticamente equitativo, sino que depende de condiciones de acceso, formación docente y supervisión ética que distan de estar garantizadas en el contexto ecuatoriano. Esta triangulación teórica, pedagógica, técnica y crítica, constituye el lente analítico que orienta el diseño metodológico y la interpretación de los resultados de la presente investigación.

Materiales y Métodos

Este estudio usó un enfoque mixto, es decir, se combinó información cuantitativa con información cualitativa. Esto permitió no solo medir las percepciones de los participantes, sino también entender mejor las condiciones pedagógicas, tecnológicas y éticas que rodean el uso de estas herramientas (Hernández-Sampieri & Mendoza, 2018).

La investigación se basó en el paradigma sociocrítico con orientación interpretativa (Miranda & Ortiz, 2020). Desde esta mirada, la evaluación mediada por tecnología no se analizó solo como una innovación, sino también tomando en cuenta las desigualdades de acceso, los problemas éticos y las condiciones institucionales que pueden facilitar o dificultar su uso.

Se usó un diseño mixto secuencial exploratorio, no experimental, transeccional y de campo (Hernández et al., 2014; Arias, 2020). El estudio se organizó en dos fases. La primera fue una revisión documental y bibliográfica, que sirvió para definir las categorías de análisis y construir los instrumentos. La segunda fue el trabajo de campo, con un alcance descriptivo-correlacional y propositivo, para conocer las percepciones de los participantes, explorar relaciones entre variables y aportar evidencia para el modelo evaluativo.

En la primera fase se hizo una revisión sistemática de literatura siguiendo las directrices PRISMA (Torres et al., 2024). Se consultaron Scopus, Web of Science, ERIC, Dialnet, Redalyc y Latindex, usando descriptores en español e inglés como evaluación educativa, herramientas digitales, inteligencia artificial, retroalimentación formativa, assessment y AI in education. Se priorizaron investigaciones publicadas entre 2018 y 2026.

En la segunda fase se analizaron las experiencias y percepciones de docentes y estudiantes sobre la utilidad pedagógica de estas herramientas, su aporte a la retroalimentación y personalización del aprendizaje, las condiciones de accesibilidad y las preocupaciones éticas. Esta fase también incluyó una parte propositiva: los hallazgos empíricos y los referentes teóricos se juntaron para construir un modelo de evaluación mediado por tecnología, pensado para el contexto educativo ecuatoriano (Lima & Soto, 2020).

La población fueron docentes y estudiantes de instituciones de educación secundaria y superior del Ecuador. Se usó un muestreo no probabilístico intencional por criterios (Reyes & Boente, 2019). Para los docentes, se pidió que estuvieran en ejercicio activo y tuvieran al menos dos años de experiencia. Para los estudiantes, que pertenecieran a bachillerato o educación

superior y usaran plataformas digitales con regularidad. Y para las instituciones, que ya tuvieran alguna herramienta tecnológica incorporada en sus procesos evaluativos.

Participaron 50 personas en total: 20 docentes y 30 estudiantes, de cinco instituciones educativas de contextos urbanos y rurales. Para la parte cualitativa se eligió a los participantes de forma intencional hasta llegar a un punto de saturación de la información, un criterio común en estudios cualitativos para saber cuándo dejar de recolectar datos (Hernández et al., 2014).

La muestra respondió al carácter exploratorio y contextual del estudio. Por eso, los resultados se interpretaron dentro de las características propias de las instituciones participantes, sin pretender generalizarlos a todo el sistema educativo ecuatoriano.

Se usaron cuatro técnicas: revisión sistemática de literatura, encuesta, entrevista semiestructurada y análisis documental. Esta combinación permitió comparar la evidencia científica con las experiencias de docentes y estudiantes, y con las condiciones institucionales en las que ocurren las prácticas evaluativas.

La encuesta fue un cuestionario de 12 ítems, dividido en cuatro dimensiones: utilidad pedagógica; retroalimentación y personalización; accesibilidad tecnológica y brecha digital; y ética, transparencia y confianza. Se respondió en escala Likert de cinco puntos, de 1 (totalmente en desacuerdo) a 5 (totalmente de acuerdo). Se aplicó de forma virtual, a través de Google Forms.

La validez de contenido se revisó con el juicio de cinco especialistas en tecnología educativa y evaluación. Después se calculó la consistencia interna con el coeficiente Alfa de Cronbach, tomando como adecuado un valor igual o mayor a 0,70.

Las entrevistas semiestructuradas incluyeron preguntas abiertas sobre cuatro categorías: utilidad pedagógica de las herramientas digitales e IA; barreras de acceso y brecha digital; aspectos éticos e integridad académica; y propuestas para avanzar hacia una evaluación más equitativa. Durante el análisis se dejó abierta la posibilidad de que surgieran nuevas subcategorías a partir de lo que los participantes fueran contando.

El análisis documental complementó esta información con la revisión de currículos, planes institucionales y normativas educativas relacionadas con evaluación y uso de tecnologías.

En el caso de los estudiantes, también se registró su promedio académico del último período, en una escala de 0 a 10. Este dato se usó para explorar si tenía alguna relación con la dimensión de retroalimentación y personalización mediada por herramientas digitales e IA.

Los datos cuantitativos se procesaron con SPSS. Se calcularon frecuencias, porcentajes, medias y desviaciones estándar para describir las respuestas. La consistencia interna del cuestionario se evaluó con el Alfa de Cronbach.

Para ver si existía relación entre la retroalimentación y personalización mediada por herramientas digitales e IA y el desempeño académico de los estudiantes, se usó el coeficiente Rho de Spearman, con un nivel de significación de $p < 0,05$. Como el diseño fue no experimental y transeccional, esta relación se interpretó en términos de asociación y no de causalidad.

La información de las entrevistas se analizó mediante codificación abierta y axial, organizando los contenidos según las categorías definidas de antemano y las que fueron surgiendo durante el análisis. Después, estos hallazgos cualitativos se compararon con los resultados de la encuesta y con la evidencia documental.

Por último, ambos componentes se integraron en una matriz de convergencia, que permitió identificar coincidencias, complementos y diferencias entre las distintas fuentes de información. Esta triangulación fue el puente entre los hallazgos empíricos y la construcción del modelo evaluativo propuesto.

La participación fue voluntaria, y antes de empezar se informó a cada participante sobre los objetivos y características del estudio. Se garantizó la confidencialidad de los datos, el anonimato de los participantes y la posibilidad de retirarse de la investigación en cualquier momento sin ninguna consecuencia. También se contó con la autorización de las instituciones participantes y se respetaron los principios éticos que aplican a este tipo de investigaciones con seres humanos.

Análisis de Resultados

Siguiendo el orden de los objetivos específicos, en este apartado se articulan los hallazgos documentales con la información cuantitativa y cualitativa recogida en el trabajo de campo. La fase empírica reunió las respuestas de 20 docentes y 30 estudiantes de cinco instituciones educativas, a quienes se aplicó un cuestionario de 12 ítems en escala Likert de cinco puntos. Estos ítems se organizaron en cuatro dimensiones: utilidad pedagógica, retroalimentación y personalización, accesibilidad tecnológica, y condiciones éticas asociadas al uso de herramientas digitales e inteligencia artificial.

El instrumento mostró una consistencia interna sólida, con un Alfa de Cronbach de 0,905, un valor que confirma la homogeneidad entre los ítems y respalda el análisis conjunto de las cuatro dimensiones.

Herramientas digitales e inteligencia artificial utilizadas en la evaluación

En respuesta al primer objetivo específico, la revisión documental permitió reconocer que la incorporación tecnológica en la evaluación educativa se concentra principalmente en plataformas de gestión del aprendizaje, aplicaciones para cuestionarios y actividades

interactivas, sistemas de retroalimentación automatizada y herramientas de inteligencia artificial generativa.

Más que una sustitución de los procedimientos tradicionales, los hallazgos muestran una convivencia entre recursos convencionales y digitales. Las plataformas institucionales continúan siendo utilizadas para organizar actividades, registrar calificaciones y aplicar pruebas, mientras que las herramientas basadas en inteligencia artificial empiezan a incorporarse con mayor frecuencia en tareas vinculadas con la generación de preguntas, revisión de respuestas, elaboración de rúbricas y retroalimentación inmediata.

Esto sugiere que la transformación tecnológica de la evaluación todavía atraviesa una etapa de transición. El uso de inteligencia artificial no aparece como una práctica consolidada, sino como un recurso que se va integrando de forma selectiva, sobre todo en aquellas actividades donde docentes y estudiantes perciben una ventaja concreta en rapidez y acompañamiento.

Percepción de docentes y estudiantes

La Tabla 1 resume las puntuaciones alcanzadas en las cuatro dimensiones evaluadas.

Tabla 1.

Percepción de docentes y estudiantes sobre el uso de herramientas digitales e IA en la evaluación

Dimensión	Docentes		Estudiantes	
	Media	DE	Media	DE2
Utilidad pedagógica	3,85	0,64	3,72	0,59
Retroalimentación y personalización	3,63	0,51	3,74	0,54
Accesibilidad tecnológica	3,6	0,47	3,5	0,63
Ética, transparencia y confianza	3,42	0,63	3,42	0,65
Valoración global	3,63	0,45	3,6	0,52

Nota. Escala de 1 = totalmente en desacuerdo a 5 = totalmente de acuerdo. Fuente: Elaborado por los autores.

La valoración global fue favorable en ambos grupos, aunque no al punto de hablar de una aceptación irrestricta: los docentes obtuvieron una media general de 3,63, y los estudiantes, 3,60. Esta cercanía entre ambas puntuaciones deja ver que los dos grupos reconocen posibilidades educativas en estas tecnologías, pero también conservan reservas frente a ciertas condiciones de implementación.

Entre los docentes, la dimensión mejor valorada fue la utilidad pedagógica ($M = 3,85$; $DE = 0,64$), lo que sugiere que las herramientas digitales se perciben sobre todo como un apoyo para diversificar los procedimientos evaluativos, reducir tareas repetitivas y contar con información más rápida. En las entrevistas, esta valoración se reflejó en comentarios sobre el ahorro de tiempo y la posibilidad de revisar de forma más sistemática el avance de los estudiantes.

En el caso de los estudiantes, la puntuación más alta correspondió a retroalimentación y personalización ($M = 3,74$; $DE = 0,54$). Valoraron especialmente poder conocer sus errores de inmediato y contar con nuevas oportunidades para corregirlos — un hallazgo que resulta revelador, porque muestra que, desde su experiencia, la tecnología cobra sentido cuando ayuda a comprender el propio desempeño, y no simplemente cuando reemplaza una prueba escrita por una actividad digital.

Las diferencias entre ambos grupos fueron mínimas: los docentes valoraron ligeramente más la utilidad pedagógica, y los estudiantes, la retroalimentación y personalización. Tiene sentido si se piensa en el rol que cada grupo ocupa dentro del proceso evaluativo: para el docente, la tecnología es también una herramienta de organización y seguimiento; para el estudiante, su beneficio se percibe con más claridad cuando se traduce en una orientación concreta para mejorar.

Accesibilidad y brecha digital

La accesibilidad obtuvo medias de 3,60 en docentes y 3,50 en estudiantes, por debajo de lo observado en utilidad pedagógica y retroalimentación. Esto no refleja una percepción negativa generalizada, pero sí deja ver que las condiciones materiales siguen siendo una limitación importante.

Las entrevistas ayudaron a matizar esta situación: el acceso tecnológico no se describió como un problema uniforme. En algunas instituciones, la conectividad y la disponibilidad de dispositivos permiten incorporar plataformas evaluativas de manera relativamente estable; en otras, sobre todo cuando los estudiantes dependen de conexiones domésticas o de teléfonos móviles compartidos, las mismas actividades generan dificultades de participación.

Este contraste explica por qué contar con una plataforma no garantiza, por sí solo, una evaluación equitativa. Los participantes señalaron que los problemas de conectividad pueden retrasar la entrega de actividades, dificultar la participación en evaluaciones sincrónicas y dejar a ciertos estudiantes en desventaja, sin que eso tenga necesariamente relación con lo que saben.

Todo esto sugiere que la brecha digital no es solo un problema tecnológico, se visualiza como una condición que puede afectar la validez misma de la evaluación. Cuando el desempeño

observado depende en parte de la calidad del acceso a internet o del dispositivo disponible, se vuelve difícil atribuir el resultado únicamente al aprendizaje alcanzado. **Ética, transparencia y confianza**

La dimensión con menor puntuación en ambos grupos fue la relacionada con ética, transparencia y confianza, con una media de 3,42 tanto para docentes como para estudiantes.

Este resultado merece atención porque contrasta con la valoración positiva de la utilidad pedagógica. Los participantes reconocen las ventajas de las herramientas, pero no manifiestan el mismo grado de seguridad respecto de la forma en que los sistemas de inteligencia artificial procesan información, generan recomendaciones o utilizan los datos personales.

Las entrevistas permitieron identificar tres preocupaciones recurrentes. La primera se relacionó con la privacidad de la información; la segunda, con la dificultad para comprender cómo determinadas plataformas producen sus respuestas o recomendaciones; y la tercera, con el riesgo de que el docente delegue excesivamente decisiones evaluativas en sistemas automatizados.

Estas inquietudes no se tradujeron en un rechazo generalizado de la inteligencia artificial. Más bien reflejan una posición prudente: los participantes aceptan su incorporación siempre que exista supervisión humana y que la institución defina reglas claras sobre su utilización. En varias respuestas cualitativas apareció la idea de que la tecnología puede ayudar a evaluar, pero no debería constituirse en la única instancia que determine el desempeño de un estudiante.

Uso formativo de inteligencia artificial y desempeño académico

Para profundizar en el posible vínculo entre el uso formativo de la inteligencia artificial y el aprendizaje, se analizó la relación entre la puntuación de los 30 estudiantes en la dimensión de retroalimentación y personalización y su promedio académico del último período.

Dado el tamaño de la muestra y la naturaleza ordinal de parte de la información, se optó por el coeficiente de correlación de Spearman. El análisis arrojó una asociación positiva moderada y estadísticamente significativa entre ambas variables ($p = 0,591$; $p < 0,001$).

Tabla 2.*Relación entre uso formativo de IA y desempeño académico*

Variables	p de		
	Spearman	p	n
Retroalimentación/personalización mediante IA – promedio académico	0,591	< 0,001	30

Fuente: Elaborado por los autores.

El promedio académico del grupo estudiantil fue de 7,88 puntos (DE = 0,89). La dirección positiva de la correlación indica que los estudiantes que valoraron más el uso de herramientas digitales e IA para recibir retroalimentación y ajustar su proceso de aprendizaje también tendieron a obtener mejores resultados académicos.

Con todo, este hallazgo debe leerse con cautela. La asociación encontrada no demuestra que la inteligencia artificial produzca directamente una mejora en el rendimiento: el diseño no experimental y transeccional del estudio no permite establecer causalidad. Es posible, por ejemplo, que los estudiantes con mayor autonomía, mejores condiciones de acceso o mayor dominio digital sean también quienes sacan más provecho de estas herramientas. El hallazgo, entonces, debe entenderse como una relación relevante pero que requiere investigaciones posteriores con diseños longitudinales o experimentales para confirmarse.

Integración de los resultados cuantitativos y cualitativos

La triangulación dejó ver una convergencia clara entre las distintas fuentes de información. Las puntuaciones cuantitativas muestran una valoración favorable de la utilidad pedagógica y de la retroalimentación, y las entrevistas ayudan a entender por qué estas dimensiones fueron especialmente apreciadas: para docentes y estudiantes, el valor de la tecnología está ligado sobre todo a la rapidez para disponer de información sobre el aprendizaje y a la posibilidad de intervenir antes de que las dificultades se consoliden.

Algo similar ocurre con las puntuaciones más moderadas en accesibilidad y ética, que encuentran respaldo en los testimonios cualitativos. Las dificultades de conectividad, la disponibilidad desigual de dispositivos, las dudas sobre privacidad y la necesidad de mantener supervisión humana aparecieron una y otra vez en el discurso de los participantes.

La triangulación revela, así, una postura menos polarizada de lo que podría suponerse: la comunidad educativa no se sitúa entre aceptar o rechazar la inteligencia artificial, sino que muestra una disposición favorable a utilizarla cuando existe una finalidad pedagógica clara, siempre que venga acompañada de límites bien establecidos.

Modelo evaluativo derivado de los hallazgos

Los resultados obtenidos sirvieron de fundamento para la construcción de un modelo evaluativo híbrido articulado en cuatro componentes: diagnóstico, seguimiento formativo, retroalimentación personalizada y gobernanza ética de los datos.

El diagnóstico considera tanto los conocimientos previos como las condiciones reales de acceso tecnológico del estudiante. El seguimiento formativo incorpora evidencia continua sobre el progreso y permite identificar dificultades antes de la evaluación final. La retroalimentación personalizada utiliza recursos digitales e inteligencia artificial como apoyo para proporcionar orientaciones oportunas, pero mantiene al docente como responsable último de las decisiones evaluativas.

Finalmente, la gobernanza ética establece la necesidad de definir criterios institucionales sobre privacidad, transparencia, uso responsable de datos y supervisión de los sistemas automatizados. Este último componente responde directamente a las preocupaciones expresadas por los participantes y evita entender la innovación únicamente como incorporación de tecnología.

En conjunto, los resultados muestran que las herramientas digitales y la inteligencia artificial poseen potencial para transformar la evaluación educativa, especialmente cuando fortalecen la retroalimentación y permiten responder de manera más diferenciada a las necesidades de aprendizaje. Sin embargo, el efecto positivo de estas tecnologías no depende únicamente de sus capacidades técnicas. El acceso equitativo, la formación de docentes y estudiantes, la claridad de los criterios éticos y la permanencia del juicio pedagógico humano constituyen condiciones necesarias para que dicha transformación tenga un sentido verdaderamente formativo.

Discusión

La discusión se organiza retomando los hallazgos expuestos en los resultados y contrastándolos con múltiples autores, con el fin de interpretar su alcance, sus coincidencias y las tensiones que plantean frente a la literatura especializada.

Ante esta perspectiva, los resultados muestran que las herramientas digitales y la inteligencia artificial están cambiando la forma en que docentes y estudiantes entienden la evaluación. Pero ese cambio no depende solo de tener nuevas tecnologías disponibles, sino del sentido pedagógico con el que se usan. La valoración favorable que se encontró en ambos grupos, sobre todo en utilidad pedagógica y en las posibilidades de retroalimentación y personalización, sugiere que la tecnología importa más cuando ayuda a entender el progreso del estudiante y permite actuar a tiempo sobre sus dificultades.

Esto se relaciona con lo que plantean Black y Wiliam (2009), quienes ven la evaluación formativa como parte del propio proceso de aprendizaje, y con Hattie y Timperley (2007), para quienes la retroalimentación funciona mejor cuando muestra la distancia entre lo que el estudiante hace ahora y la meta que debe alcanzar. Desde esta idea, la rapidez de las herramientas digitales no es una mejora por sí sola: lo que realmente aporta es que la información generada le diga al estudiante qué debe corregir y cómo seguir avanzando.

La asociación positiva entre retroalimentación y personalización mediante IA y el desempeño académico ($p = 0,591$; $p < 0,001$) suma otro elemento a esta lectura. El diseño del estudio no permite decir que una cosa cause la otra, pero sí sugiere que los estudiantes que aprovechan mejor estas herramientas tienden también a tener mejores resultados. Esto coincide con lo que dicen Luckin et al. (2016) sobre el potencial de los sistemas adaptativos para responder a las diferencias entre estudiantes, aunque ellos mismos advierten que todo depende de la calidad de los datos, del acompañamiento docente y de las condiciones en que se usan estas herramientas.

Sin embargo, aceptar estas tecnologías no significó confiar en ellas del todo. La dimensión de ética, transparencia y confianza fue la que obtuvo las puntuaciones más bajas, tanto en docentes como en estudiantes. Las dudas sobre privacidad, sobre cómo funcionan realmente los sistemas automatizados y sobre un posible desplazamiento del criterio docente coinciden con lo que advierten Zawacki-Richter et al. (2019) y Holmes et al. (2019): que usar IA en educación plantea preguntas que van más allá de si funciona bien técnicamente. Aquí aparece algo interesante entre los participantes: hay disposición a usar la IA, pero no a dejar que ella tome decisiones que requieren interpretación pedagógica.

La accesibilidad es otra condición que matiza estos beneficios. Las diferencias en conectividad y en disponibilidad de dispositivos muestran que una misma estrategia tecnológica puede vivirse de forma muy distinta según el contexto. Esto respalda lo que plantea Selwyn (2016), quien cuestiona la idea de que todos tienen más o menos el mismo acceso a la tecnología. En Ecuador, la brecha digital puede trasladarse directamente al proceso evaluativo y hacer que un problema de conexión se vea, sin serlo, como una diferencia de desempeño académico.

Al juntar todos estos hallazgos, se puede ir más allá de pensar solo en ventajas o riesgos de la inteligencia artificial. La pregunta no es si la tecnología debe estar en la evaluación educativa, sino en qué condiciones puede estarlo sin debilitar la equidad ni el lado humano del proceso. Las percepciones de los participantes, las entrevistas y la evidencia documental

apuntan en la misma dirección: hacia una integración acompañada de criterios pedagógicos, accesibilidad y supervisión docente.

En este punto cobra sentido el modelo evaluativo propuesto, que gira en torno al diagnóstico, el seguimiento formativo, la retroalimentación personalizada y la gobernanza ética de los datos. Su aporte principal no es meter más tecnología en la evaluación, sino definir condiciones para usarla de forma pedagógicamente adecuada. La IA puede ampliar el seguimiento y la personalización, pero interpretar el aprendizaje, entender el contexto y tomar decisiones que afectan al estudiante sigue siendo trabajo del docente.

Los hallazgos sugieren que la transformación digital de la evaluación educativa debe verse como un proceso donde la tecnología y el criterio pedagógico se complementan. No se trata de reemplazar al docente por sistemas automatizados, sino de aprovechar lo que la tecnología puede hacer con rapidez y a gran escala, dejando la responsabilidad humana en todo lo que necesita comprensión, juicio y sensibilidad frente a las diferencias de cada estudiante y cada contexto.

Conclusiones

La incorporación de herramientas digitales e inteligencia artificial amplía las posibilidades de la evaluación educativa, particularmente en el seguimiento del aprendizaje, la retroalimentación oportuna y la atención diferenciada a las necesidades de los estudiantes. Sin embargo, su pertinencia pedagógica no depende únicamente de las capacidades de las herramientas disponibles, sino de las condiciones en que estas son incorporadas a la práctica educativa. En el contexto analizado, las diferencias de acceso, conectividad y disponibilidad de dispositivos continúan siendo factores que pueden limitar una participación equitativa.

Docentes y estudiantes mostraron una percepción favorable hacia la integración de estas tecnologías, especialmente por sus posibilidades para agilizar la retroalimentación y favorecer experiencias de aprendizaje más personalizadas. Esta aceptación coexistió con preocupaciones relacionadas con la privacidad, la transparencia de los sistemas automatizados y el riesgo de reducir la intervención humana en la evaluación. Los hallazgos evidencian, por tanto, una disposición hacia el uso de la IA acompañada de una demanda clara de supervisión docente y criterios institucionales para orientar su aplicación.

En los estudiantes se encontró una asociación positiva entre la retroalimentación y personalización mediada por herramientas digitales e IA y el desempeño académico ($p = 0,591$; $p < 0,001$). Este resultado sugiere que un aprovechamiento formativo de estas herramientas puede relacionarse con mejores resultados académicos; no obstante, debido al carácter no experimental y transeccional del estudio, esta asociación no permite establecer una relación

causal y deberá profundizarse mediante investigaciones posteriores con muestras más amplias y diseños longitudinales.

A partir de la integración de la evidencia cuantitativa, cualitativa y documental, se plantea un modelo evaluativo articulado en cuatro componentes: diagnóstico, seguimiento formativo, retroalimentación personalizada y gobernanza ética de los datos. Su aplicación requiere considerar las condiciones tecnológicas de cada institución y mantener al docente como responsable de interpretar la información y tomar las decisiones evaluativas que afectan al estudiante.

Los resultados permiten concluir que la transformación digital de la evaluación no consiste en trasladar las prácticas tradicionales a una plataforma ni en delegar la evaluación a sistemas automatizados. Su verdadero potencial se encuentra en articular las capacidades de la tecnología con el criterio pedagógico, de manera que la inteligencia artificial contribuya a una evaluación más oportuna y personalizada sin comprometer la equidad, la transparencia ni la dimensión humana del proceso educativo.

Referencias bibliográfica

- Arias, F. G. (2020). El proyecto de investigación: Introducción a la metodología científica (7.^a ed.). Editorial Episteme.
[https://books.google.com/books?hl=es&lr=&id=W5n0BgAAQBAJ&oi=fnd&pg=PA9&dq=Arias,+F.+G.+\(2020\).+El+proyecto+de+investigaci%C3%B3n:+Introducci%C3%B3n+a+l+a+metodolog%C3%ADa+cient%C3%ADfica+\(7.%C2%AA+ed.\).+Editorial+Episteme.&ots=kZqOesunqb&sig=t-W2ha2HL_FcRDeRAam5qrLpE88](https://books.google.com/books?hl=es&lr=&id=W5n0BgAAQBAJ&oi=fnd&pg=PA9&dq=Arias,+F.+G.+(2020).+El+proyecto+de+investigaci%C3%B3n:+Introducci%C3%B3n+a+l+a+metodolog%C3%ADa+cient%C3%ADfica+(7.%C2%AA+ed.).+Editorial+Episteme.&ots=kZqOesunqb&sig=t-W2ha2HL_FcRDeRAam5qrLpE88)
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Dautović, A. M. (2025). *Evaluación formativa en la enseñanza de la lengua española en la escuela primaria-actitudes de los alumnos y padres* (Master's thesis, Sveučilište u Zagrebu, Sveučilište u Zagrebu, Filozofski fakultet, Odsjek za romanistiku). <https://repositorij.unizg.hr/object/ffzg:13802>
- Fandiño, Y. J., & Muñoz Barriga, A. (2026). Retos y oportunidades de la inteligencia artificial para la evaluación académica en la educación superior. *Revista de la Universidad de La Salle*, 95, 1–22. <https://doi.org/10.19052/ruls.vol1.iss95.9>
- Galán-Íñigo, A., Ruiz-Lázaro, J., & Jiménez-García, E. (2025). La inteligencia artificial en los procesos de evaluación en educación superior: un análisis bibliométrico (2014-2024). *Bordón, Revista de Pedagogía*, 77(3), 131–154. <https://doi.org/10.13042/Bordon.2025.107797>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hernández-Sampieri, R., & Mendoza, C. (2018). Metodología de la investigación: Las rutas cuantitativa, cualitativa y mixta. <https://www.academia.edu/download/64591365/Metolog%C3%ADvestigaci%C3%B3n.%20Rutas%20cuantitativa,%25cualitativa%20y%20mixta.pdf>
- Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
- Lima, J., & Soto, D. (2020). Paulo Freire y la pedagogía crítica: Su legado para una nueva pedagogía desde el Sur. *Revista iberoamericana de Estudos em Educação*, 15(3), 1072–1093. <http://redalyc.org/journal/6198/619865690003/html/>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson. <https://discovery.ucl.ac.uk/id/eprint/1475756/>
-

- Ministerio de Educación del Ecuador. (2022). Informe de competencias digitales docentes 2022. MinEduc.
- Miranda, S., & Ortiz, J. (2020). Los paradigmas de la investigación: Un acercamiento teórico para reflexionar desde el campo de la investigación educativa. *Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 11(21), 1–18. <https://www.scielo.org.mx/pdf/ride/v11n21/2007-7467-ride-11-21-e064.pdf>
- Pinto-Llorente, A. M., Sánchez-Gómez, M. C., García-Peñalvo, F. J., & Casillas-Martín, S. (2022). Students' perceptions and attitudes towards asynchronous technological tools in blended-learning training to improve grammatical competence in English as a second language. *Computers in Human Behavior*, 117, Artículo 106628. <https://doi.org/10.1016/j.chb.2020.106628>
- Selwyn, N. (2016). *Education and technology: Key issues and debates* (2.^a ed.). Bloomsbury Academic. [https://books.google.com/books?hl=es&lr=&id=reKPdzC6toC&oi=fnd&pg=PR5&dq=Selwyn,+N.+\(2016\).+Education+and+technology:+Key+issues+and+debates+\(2.%C2%AA+ed.\).+Bloomsbury+Academic.&ots=QhQiWhGZvz&sig=aek_SqcsDzNLzK70e5Hz1u5eB_w](https://books.google.com/books?hl=es&lr=&id=reKPdzC6toC&oi=fnd&pg=PR5&dq=Selwyn,+N.+(2016).+Education+and+technology:+Key+issues+and+debates+(2.%C2%AA+ed.).+Bloomsbury+Academic.&ots=QhQiWhGZvz&sig=aek_SqcsDzNLzK70e5Hz1u5eB_w)
- Torres, R., Mejía, N., & Huayta, Y. (2024). Problems and challenges of educational public policies in Latin America: A systematic review. *Comuni@cción: Revista de Investigación en Comunicación y Desarrollo*, 15(2), 167–180. <https://doi.org/10.33595/2226-1478.15.2.1052>
- Valdez Valdez, L. S., Sánchez Uscamayta, J. O., & Lescano López, G. S. (2023). Evaluación formativa: retroalimentación, estrategias e instrumentos. *Revista Educación*, 47(2), 794–812. https://www.scielo.sa.cr/scielo.php?pid=S2215-26442023000200794&script=sci_arttext
- 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), Artículo 39. <https://doi.org/10.1186/s41239-019-0171-0>
-