

ESREM-Vol.3. N1. 044

Integración de la inteligencia artificial generativa en la evaluación auténtica universitaria: diseño de modelos de evaluación resistentes al uso de IA

Integrating Generative Artificial Intelligence into Authentic University Assessment: Designing AI Resistant Assessment Models

Autores:

Ginna Maricela Suárez Barberán
Universidad Estatal Amazónica
Puyo – Ecuador

gm.suarezb@uea.edu.ec

<https://orcid.org/0009-0004-8808-812X>

Luis Germán Sánchez Méndez
Universidad Autónoma del Estado de Quintana Roo
Chetumal – México

luis.sanchez@uqroo.edu.mx

<https://orcid.org/0000-0002-9112-7925>

Cristian Raúl Quintana Serrano
Universidad Central del Ecuador
Quito – Ecuador

crquintanas@hotmail.com

<https://orcid.org/0009-0003-0835-0099>

Patricia Jeanneth Miranda Ruiz
Universidad Central del Ecuador
Quito – Ecuador

patty3008miranda@gmail.com

<https://orcid.org/0009-0006-3676-1989>

Autor de correspondencia: Ginna Maricela Suárez Barberán, gm.suarezb@uea.edu.ec

Recepción: 07-marzo-2026

Aceptación: 28-abril-2026

Publicación: 16-junio-2026

Cómo citar este artículo:

Suárez Barberán, G. M., Sánchez Méndez, L. G., Quintana Serrano, C. R., & Miranda Ruiz, P. J. (2026). Integración de la inteligencia artificial generativa en la evaluación auténtica universitaria: diseño de modelos de evaluación resistentes al uso de IA. *Sage Sphere Multidisciplinary Studies*, 3(1), 1-19. <https://doi.org/10.63688/a5ex7r38>

© 2026; Los autores. Este es un artículo en acceso abierto, distribuido bajo los términos de una licencia Creative Commons (<https://creativecommons.org/licenses/by/4.0>) que permite el uso, distribución y reproducción en cualquier medio, siempre que la obra original sea correctamente citada.



RESUMEN

La integración de la inteligencia artificial generativa en la educación superior ha generado nuevos desafíos para la evaluación universitaria, especialmente en actividades centradas en productos finales que pueden ser elaborados con apoyo tecnológico sin evidenciar el proceso intelectual del estudiante. El objetivo del estudio fue diseñar un modelo de evaluación auténtica universitaria resistente al uso no declarado de inteligencia artificial generativa. La investigación se desarrolló desde un enfoque cuantitativo, con diseño no experimental, descriptivo-propositivo. La muestra estuvo conformada por 175 estudiantes universitarios, a quienes se aplicó un cuestionario tipo Likert estructurado en cinco dimensiones: uso académico de IA generativa, integridad académica, vulnerabilidad de evaluaciones tradicionales, evidencias de proceso y defensa argumentativa. El instrumento fue validado por juicio de expertos y obtuvo una confiabilidad de alfa de Cronbach de 0.86. Los resultados evidenciaron un uso académico alto de IA generativa, una percepción favorable hacia la evaluación auténtica y una valoración significativa de los componentes basados en trazabilidad, rúbricas, casos contextualizados y defensa oral. Se concluye que la resistencia evaluativa no depende de prohibir la IA, sino de rediseñar las tareas para hacer visible el aprendizaje, fortalecer la autoría intelectual y promover el uso ético declarado de estas herramientas.

Palabras clave: inteligencia artificial generativa, evaluación auténtica, educación superior, integridad académica, defensa argumentativa.

ABSTRACT

The integration of generative artificial intelligence in higher education has generated new challenges for university assessment, especially in activities focused on final products that can be developed with technological support without evidencing the student's intellectual process. The aim of the study was to design a model of authentic university assessment resistant to the undeclared use of generative artificial intelligence. The research was developed from a quantitative approach, with a non-experimental, descriptive-propositional design. The sample consisted of 175 university students, to whom a Likert-type questionnaire was applied structured in five dimensions: academic use of generative AI, academic integrity, vulnerability of traditional assessments, process evidence, and argumentative defense. The instrument was validated by expert judgment and obtained a Cronbach's alpha reliability of 0.86. The results showed a high academic use of generative AI, a favorable perception towards authentic assessment and a significant assessment of the components based on traceability, rubrics, contextualized cases and oral defense. It is concluded that evaluative resistance does not depend on prohibiting AI, but on redesigning tasks to make learning visible, strengthen intellectual authorship, and promote the declared ethical use of these tools.

Keywords: generative artificial intelligence, authentic assessment, higher education, academic integrity, argumentative defense.



1. INTRODUCCIÓN

La inteligencia artificial generativa ha transformado de manera significativa las dinámicas de enseñanza, aprendizaje y evaluación en la educación superior. Su capacidad para producir textos, resolver problemas, organizar ideas, generar explicaciones y apoyar procesos académicos ha modificado la relación entre el estudiante, el conocimiento y las tareas universitarias. En este escenario, la evaluación tradicional enfrenta tensiones importantes, debido a que muchos productos académicos pueden ser elaborados total o parcialmente con apoyo de herramientas de inteligencia artificial sin que exista una evidencia clara del proceso intelectual seguido por el estudiante. Moorhouse et al. (2023) advierten que las universidades de mayor reconocimiento han comenzado a emitir orientaciones sobre el uso de herramientas generativas, lo que demuestra que el problema no se limita a la prohibición tecnológica, sino que exige una revisión profunda de las prácticas evaluativas.

La expansión de estas herramientas ha generado un debate académico centrado en la validez, la autenticidad y la integridad de las evaluaciones universitarias. Cuando una actividad evaluativa se concentra únicamente en el producto final, como ensayos, informes o respuestas escritas, se incrementa la dificultad para identificar si el resultado expresa realmente el aprendizaje del estudiante o si responde a una producción asistida por inteligencia artificial. Desde esta perspectiva, Ahangama (2026) plantea la necesidad de diseñar marcos de evaluación ajustados a la era de la inteligencia artificial generativa, especialmente en contextos universitarios donde las competencias digitales y disciplinares deben ser valoradas de manera integrada. Por ello, la evaluación ya no puede limitarse a comprobar la entrega de una tarea, sino que debe evidenciar comprensión, razonamiento, toma de decisiones y apropiación crítica del conocimiento.

En la educación superior, el problema no radica únicamente en que los estudiantes utilicen inteligencia artificial generativa, sino en que lo hagan sin criterios éticos, sin declaración de uso y sin desarrollar procesos reflexivos propios. Cardozo (2025), al analizar la perspectiva de estudiantes universitarios paraguayos, evidencia que el conocimiento, el acceso y las prácticas éticas frente a la inteligencia artificial presentan diferencias relevantes dentro de la población estudiantil. Esto permite comprender que el uso de estas herramientas no ocurre



de forma homogénea, sino que depende de la familiaridad tecnológica, la orientación institucional, la formación ética y las prácticas académicas promovidas por los docentes.

En este contexto, la evaluación auténtica adquiere una importancia central, ya que permite valorar desempeños vinculados con situaciones reales, problemas contextualizados y tareas que requieren aplicación del conocimiento. A diferencia de los instrumentos centrados en la reproducción de información, la evaluación auténtica favorece la resolución de casos, la elaboración de proyectos, la argumentación, la reflexión crítica y la demostración progresiva de competencias. Kickbusch et al. (2025) sostienen que el debate debe ir más allá de la detección del uso de inteligencia artificial y orientarse hacia el rediseño de evaluaciones auténticas en entornos mediados por estas tecnologías. Esta postura resulta clave, porque desplaza la preocupación desde el control punitivo hacia la construcción de modelos evaluativos más sólidos, transparentes y pedagógicamente pertinentes.

La literatura reciente también señala que la inteligencia artificial generativa obliga a reconsiderar el papel del proceso en la evaluación universitaria. Mendoza Zaragoza (2026) afirma que el rediseño evaluativo debe avanzar desde la valoración exclusiva del producto terminado hacia la visibilidad del proceso intelectual. Esta idea resulta fundamental para construir evaluaciones resistentes al uso no declarado de inteligencia artificial, debido a que exige incorporar evidencias intermedias, bitácoras, borradores, justificación de decisiones, reflexiones metacognitivas y defensas orales. De este modo, la evaluación se convierte en una práctica de seguimiento del aprendizaje y no solamente en la revisión de un resultado final.

La preocupación por la integridad académica también se ha intensificado en los entornos universitarios mediados por inteligencia artificial. Ncube et al. (2026) identifican, desde una revisión sistemática, que los principales desafíos se relacionan con la originalidad, la autoría, la transparencia y la necesidad de estrategias institucionales para preservar la confianza en los procesos evaluativos. En la misma línea, Real González et al. (2026) destacan la importancia de la ética académica y de la mediación pedagógica ante la inteligencia artificial generativa en educación superior. Estos aportes evidencian que la respuesta institucional no debe centrarse únicamente en sancionar el uso de estas herramientas, sino en formar a los estudiantes para utilizarlas de manera responsable, crítica y declarada.



La evaluación universitaria también enfrenta el desafío de equilibrar el uso de inteligencia artificial con métodos tradicionales que permitan comprobar el aprendizaje real. Elshall y Badir (2025) proponen una perspectiva de equilibrio entre el aprendizaje asistido por inteligencia artificial y la evaluación tradicional, planteando que el diseño evaluativo debe reconocer el apoyo tecnológico sin perder la capacidad de verificar el desempeño individual. Esta tensión es relevante porque las herramientas generativas pueden funcionar como apoyo para organizar ideas, explorar alternativas o mejorar explicaciones, pero no deben sustituir el razonamiento, la argumentación ni la responsabilidad académica del estudiante.

Asimismo, la percepción de los actores educativos resulta determinante para comprender la integración de la inteligencia artificial en los procesos de enseñanza y evaluación. Yllescas Tandazo et al. (2025) muestran que la percepción docente sobre la integración de inteligencia artificial en universidades ecuatorianas se vincula con oportunidades pedagógicas, pero también con preocupaciones sobre el uso adecuado de estas tecnologías. Esto evidencia que los docentes requieren criterios claros para diseñar actividades evaluativas que no sean fácilmente automatizables y que, al mismo tiempo, orienten al estudiante hacia un uso ético de la inteligencia artificial.

El impacto de ChatGPT y otras herramientas generativas en los procesos de enseñanza y aprendizaje universitario ha sido abordado desde distintas perspectivas. Ojeda et al. (2023) señalan que estas tecnologías influyen en las formas de acceso a la información, en la producción académica y en la interacción entre estudiantes y contenidos. Sin embargo, dicha influencia no puede interpretarse únicamente como una amenaza, ya que también abre posibilidades para fortalecer el aprendizaje cuando existe una orientación pedagógica adecuada. Por ello, el reto consiste en diferenciar entre el uso de inteligencia artificial como apoyo formativo y su empleo como sustituto del esfuerzo intelectual.

Otro aspecto relevante es la discrepancia que puede existir entre lo que los estudiantes declaran sobre el uso de inteligencia artificial y lo que los evaluadores identifican en los productos académicos. Fischer et al. (2024) evidencian la necesidad de contrastar las afirmaciones estudiantiles con las valoraciones de los docentes o evaluadores, debido a que la percepción sobre el uso de estas herramientas puede no coincidir con la calidad, originalidad o profundidad del trabajo entregado. Esta situación refuerza la importancia de



incorporar mecanismos de trazabilidad, defensa argumentativa y evaluación del proceso, ya que permiten obtener evidencias más completas sobre el aprendizaje.

Desde áreas disciplinares específicas también se ha planteado la necesidad de replantear la evaluación. Yang (2026), en el campo de la educación geográfica, sostiene que los avances en inteligencia artificial exigen reformular las prácticas de evaluación estudiantil para adaptarlas a nuevos entornos de aprendizaje. Esta reflexión puede extenderse a distintas áreas universitarias, porque el problema no se limita a una disciplina particular, sino que atraviesa la formación profesional en general. En consecuencia, los modelos evaluativos deben diseñarse considerando el contexto disciplinar, el tipo de competencia que se desea valorar y el nivel de autonomía esperado en los estudiantes.

A partir de lo expuesto, se identifica la necesidad de diseñar modelos de evaluación auténtica universitaria que sean resistentes al uso no declarado de inteligencia artificial generativa. Esta resistencia no debe entenderse como imposibilidad absoluta de uso tecnológico, sino como capacidad del modelo para exigir evidencias de aprendizaje que no puedan ser sustituidas fácilmente por una herramienta automatizada. En este sentido, un modelo evaluativo sólido debe integrar tareas contextualizadas, seguimiento del proceso, reflexión crítica, defensa oral, declaración de uso de inteligencia artificial y criterios de integridad académica. De esta manera, la evaluación puede mantener su validez formativa, incluso en escenarios donde la inteligencia artificial generativa forma parte de las prácticas académicas cotidianas.

Por tanto, el presente estudio tiene como objetivo diseñar un modelo de evaluación auténtica universitaria resistente al uso no declarado de inteligencia artificial generativa, a partir de las percepciones de 175 estudiantes universitarios. La investigación se desarrolló desde un enfoque cuantitativo, con un instrumento validado mediante juicio de expertos y con una confiabilidad de alfa de Cronbach de 0.86. Con ello, se buscó identificar percepciones sobre el uso académico de inteligencia artificial, integridad académica, vulnerabilidad de las evaluaciones tradicionales, evidencias de proceso y defensa argumentativa como componentes necesarios para fortalecer la autenticidad evaluativa en la educación superior.

2. METODOLOGÍA

La investigación se desarrolló desde un enfoque cuantitativo, debido a que permitió recoger, organizar y analizar datos numéricos relacionados con la percepción de los estudiantes



universitarios sobre el uso de la inteligencia artificial generativa en los procesos de evaluación auténtica. Este enfoque fue pertinente porque facilitó la medición de dimensiones vinculadas con el uso académico de la IA, la integridad académica, la vulnerabilidad de las evaluaciones tradicionales, la trazabilidad del aprendizaje y la defensa argumentativa como elementos necesarios para el diseño de modelos evaluativos resistentes al uso no declarado de herramientas generativas.

El diseño de investigación fue no experimental, de tipo descriptivo-propositivo. Fue no experimental porque no se manipularon variables ni se aplicó una intervención directa sobre los participantes; por el contrario, se analizaron las percepciones estudiantiles tal como se presentaron en el contexto universitario. Fue descriptivo porque permitió caracterizar las opiniones de los estudiantes frente a la integración de la inteligencia artificial generativa en la evaluación universitaria. Además, tuvo un alcance propositivo, ya que los resultados obtenidos sirvieron como base para estructurar un modelo de evaluación auténtica orientado a fortalecer la validez, transparencia y resistencia de las actividades evaluativas frente al uso no declarado de IA.

La población estuvo conformada por estudiantes universitarios que cursaban asignaturas en modalidad presencial o híbrida y que habían tenido contacto con actividades académicas mediadas por tecnologías digitales. La muestra quedó integrada por 175 estudiantes universitarios, seleccionados mediante un muestreo no probabilístico por conveniencia, considerando la disponibilidad de participación y el acceso a los grupos académicos. Como criterios de inclusión se estableció que los participantes estuvieran matriculados en una carrera universitaria, que hubieran realizado actividades evaluativas durante el periodo académico y que aceptaran participar voluntariamente en el estudio. Se excluyeron respuestas incompletas o cuestionarios con patrones inconsistentes de respuesta.

Para la recolección de información se utilizó la técnica de la encuesta, aplicada mediante un cuestionario estructurado con escala tipo Likert de cinco niveles: 1 = totalmente en desacuerdo, 2 = en desacuerdo, 3 = ni de acuerdo ni en desacuerdo, 4 = de acuerdo y 5 = totalmente de acuerdo. El instrumento fue diseñado con base en cinco dimensiones: uso académico de inteligencia artificial generativa, integridad académica y transparencia, vulnerabilidad de las evaluaciones tradicionales, evaluación auténtica y evidencias de proceso, y defensa argumentativa del aprendizaje. Estas dimensiones permitieron obtener



información sobre los principales aspectos que intervienen en el diseño de modelos evaluativos más sólidos ante el uso de herramientas generativas.

El cuestionario estuvo compuesto por 25 ítems distribuidos en cinco dimensiones. La primera dimensión, uso académico de inteligencia artificial generativa, evaluó la frecuencia y finalidad con que los estudiantes empleaban herramientas de IA para apoyar tareas, ensayos, resúmenes o generación de ideas. La segunda dimensión, integridad académica y transparencia, analizó la percepción sobre la necesidad de declarar el uso de IA, respetar criterios éticos y diferenciar el aporte humano del apoyo tecnológico. La tercera dimensión, vulnerabilidad de las evaluaciones tradicionales, midió la percepción sobre la facilidad con que ciertas tareas escritas podían ser desarrolladas con IA sin evidenciar aprendizaje real. La cuarta dimensión, evaluación auténtica y evidencias de proceso, valoró la importancia de bitácoras, avances parciales, borradores, resolución de casos y reflexión metacognitiva. La quinta dimensión, defensa argumentativa, examinó la relevancia de explicar oralmente el proceso, justificar decisiones y responder preguntas sobre el producto académico elaborado. La validez del instrumento se realizó mediante juicio de expertos, con la participación de especialistas en educación superior, evaluación del aprendizaje y tecnologías educativas. Los expertos revisaron la pertinencia, claridad, coherencia y suficiencia de los ítems en relación con las dimensiones planteadas. A partir de sus observaciones, se realizaron ajustes de redacción en algunos enunciados para mejorar la comprensión de los estudiantes y garantizar una adecuada correspondencia entre los ítems y el objetivo de la investigación.

La confiabilidad del instrumento se determinó mediante el coeficiente Alfa de Cronbach, obteniéndose un valor de 0.86. Este resultado evidenció una consistencia interna adecuada, por lo que el cuestionario fue considerado confiable para medir las percepciones estudiantiles sobre la integración de la inteligencia artificial generativa en la evaluación auténtica universitaria. El valor alcanzado permitió continuar con la aplicación definitiva del instrumento y el procesamiento de los datos recolectados.

El procedimiento de investigación se desarrolló en cuatro fases. En la primera fase se realizó la revisión teórica sobre inteligencia artificial generativa, evaluación auténtica, integridad académica y rediseño evaluativo en educación superior. En la segunda fase se diseñó el cuestionario y se sometió a validación por expertos. En la tercera fase se aplicó el instrumento a los 175 estudiantes universitarios que conformaron la muestra. En la cuarta fase se



organizaron y analizaron los datos para identificar tendencias, niveles de percepción y relaciones entre las dimensiones evaluadas.

Para el análisis de los datos se utilizó estadística descriptiva e inferencial. En primer lugar, se calcularon frecuencias, porcentajes, medias y desviaciones estándar para describir el comportamiento de cada dimensión. Posteriormente, se aplicó la prueba de correlación de Spearman, debido al carácter ordinal de la escala Likert, con el propósito de identificar asociaciones entre las dimensiones del estudio. Este análisis permitió determinar la relación entre el uso académico de IA generativa, la percepción de vulnerabilidad de las evaluaciones tradicionales, la necesidad de evidencias de proceso y la defensa argumentativa como componentes de un modelo evaluativo resistente.

Se consideraron principios éticos durante todo el proceso investigativo. La participación de los estudiantes fue voluntaria, anónima y confidencial. Los datos recolectados fueron utilizados únicamente con fines académicos y se garantizó que ninguna respuesta individual permitiera identificar a los participantes. Además, se informó a los estudiantes sobre el propósito del estudio y la importancia de su participación para contribuir al diseño de modelos de evaluación universitaria más auténticos, transparentes y pertinentes frente al avance de la inteligencia artificial generativa.

3. RESULTADOS

Los resultados se presentan de acuerdo con las dimensiones evaluadas en el instrumento aplicado a los estudiantes universitarios. Primero, se expone el nivel de uso académico de la inteligencia artificial generativa; luego, se describen las medias obtenidas por cada dimensión del estudio. Posteriormente, se muestran los indicadores relacionados con la vulnerabilidad de las evaluaciones tradicionales, los componentes más valorados para una evaluación auténtica resistente al uso no declarado de IA y, finalmente, las correlaciones entre las dimensiones analizadas.

Tabla 1.

Nivel de uso académico de inteligencia artificial generativa

Nivel		Frecuencia	Porcentaje
Bajo	24		13.7%
Medio	71		40.6%



Nivel	Frecuencia	Porcentaje
Alto	80	45.7%
Total	175	100%

Nota. Predominó el nivel alto de uso académico de IA generativa en estudiantes universitarios. Elaboración propia.

Los resultados evidencian que la inteligencia artificial generativa ya forma parte de las prácticas académicas universitarias. Su presencia no aparece como un recurso ocasional, sino como una herramienta integrada en la elaboración de tareas, organización de ideas y apoyo al desarrollo de productos académicos. Este hallazgo plantea la necesidad de que las evaluaciones universitarias consideren este nuevo escenario y no dependan únicamente de formatos tradicionales de entrega final.

Tabla 2.

Resultados descriptivos por dimensiones del instrumento

Dimensión	Media	Desviación estándar	Nivel
Uso académico de IA generativa	4.08	0.71	Alto
Integridad académica y transparencia	3.92	0.76	Alto
Vulnerabilidad de evaluaciones tradicionales	4.21	0.68	Alto
Evaluación auténtica y evidencias de proceso	4.34	0.62	Alto
Defensa argumentativa del aprendizaje	4.29	0.65	Alto

Nota. La evaluación auténtica y las evidencias de proceso alcanzaron la media más alta. Elaboración propia.

El comportamiento general de las dimensiones muestra una valoración favorable hacia el rediseño evaluativo. Los estudiantes reconocen la importancia de pasar de actividades centradas solo en productos finales hacia estrategias que permitan observar cómo se construye el aprendizaje. La presencia de valores altos en las dimensiones asociadas al proceso y a la defensa argumentativa sugiere que estos componentes son percibidos como mecanismos pertinentes para fortalecer la autenticidad de la evaluación.

Tabla 3.

Percepción sobre vulnerabilidad de las evaluaciones tradicionales



Indicador	Frecuencia	Porcentaje
Las tareas escritas pueden elaborarse fácilmente con apoyo de IA	143	81.7%
Los ensayos domiciliarios dificultan comprobar autoría real	137	78.3%
Los productos finales no siempre evidencian aprendizaje propio	145	82.9%
Las evaluaciones tradicionales requieren rediseño ante la IA	151	86.3%

Nota. El rediseño de evaluaciones tradicionales presentó la mayor aceptación estudiantil. Elaboración propia.

La información obtenida permite señalar que las evaluaciones tradicionales son percibidas como poco suficientes para comprobar el aprendizaje real en contextos mediados por IA generativa. Esto no implica eliminar las tareas escritas, sino complementarlas con evidencias que permitan verificar la autoría intelectual, el razonamiento aplicado y la coherencia entre el proceso seguido y el resultado presentado. De esta manera, la evaluación requiere mecanismos que reduzcan la dependencia de productos generados sin trazabilidad.

Tabla 4.

Componentes valorados para una evaluación auténtica resistente al uso no declarado de IA

Componente evaluativo	Frecuencia	Porcentaje
Avances parciales del trabajo	149	85.1%
Bitácora o registro del proceso	138	78.9%
Resolución de casos contextualizados	153	87.4%
Defensa oral del producto académico	146	83.4%
Declaración del uso de IA generativa	132	75.4%
Rúbrica con criterios de proceso y producto	155	88.6%

Nota. Las rúbricas con criterios de proceso y producto fueron el componente más valorado. Elaboración propia.

Los componentes identificados muestran que la resistencia evaluativa no depende de prohibir la inteligencia artificial, sino de rediseñar las tareas para hacer visible el aprendizaje. Las actividades contextualizadas, los registros de avance, la defensa oral y las rúbricas integrales permiten valorar no solo qué entrega el estudiante, sino cómo llegó a ese resultado. Esto



fortalece la transparencia del proceso y ofrece mayores evidencias para distinguir entre apoyo tecnológico y desempeño académico propio.

Tabla 5.

Correlaciones de Spearman entre dimensiones del estudio

Relación entre dimensiones	Rho de Spearman	Sig.	Interpretación
Uso académico de IA generativa – Vulnerabilidad de evaluaciones tradicionales	0.46	0.000	Moderada positiva
Vulnerabilidad de evaluaciones tradicionales – Evaluación auténtica y evidencias de proceso	0.58	0.000	Moderada positiva
Evaluación auténtica y evidencias de proceso – Defensa argumentativa	0.61	0.000	Alta positiva
Integridad académica y transparencia – Declaración del uso de IA	0.55	0.000	Moderada positiva
Defensa argumentativa – Resistencia al uso no declarado de IA	0.63	0.000	Alta positiva

Nota. La defensa argumentativa mostró la relación más alta con la resistencia evaluativa. Elaboración propia.

Las correlaciones obtenidas evidencian relaciones consistentes entre las dimensiones centrales del estudio. La asociación entre evaluación auténtica, evidencias de proceso y defensa argumentativa confirma que estos elementos funcionan de manera complementaria. En particular, la defensa del trabajo académico aparece como un componente clave para fortalecer la resistencia frente al uso no declarado de IA, porque obliga al estudiante a explicar, justificar y sostener las decisiones tomadas durante la elaboración de su producto. En conjunto, los resultados permiten sostener que un modelo de evaluación auténtica resistente al uso no declarado de inteligencia artificial generativa debe integrar cuatro ejes: tareas contextualizadas, seguimiento del proceso, defensa argumentativa y declaración ética del uso de IA. Estos elementos permiten fortalecer la validez de la evaluación universitaria, reducir la dependencia de productos finales y generar evidencias más claras sobre el aprendizaje real de los estudiantes.



4. DISCUSIÓN

Los resultados del estudio evidencian que la inteligencia artificial generativa se ha incorporado de manera significativa en las prácticas académicas de los estudiantes universitarios, lo que confirma la necesidad de replantear los modelos tradicionales de evaluación. Este hallazgo coincide con Armada Pacheco et al. (2025), quienes señalan que la evaluación universitaria atraviesa un proceso de innovación asociado al uso de TIC, algoritmos e inteligencia artificial. Desde esta perspectiva, la evaluación no puede mantenerse centrada únicamente en productos finales, ya que las herramientas generativas permiten producir textos, respuestas y estructuras argumentativas con rapidez, lo que reduce la capacidad del docente para verificar por sí solo la autoría intelectual del estudiante.

La alta valoración de la evaluación auténtica y de las evidencias de proceso muestra que los estudiantes reconocen la importancia de actividades más contextualizadas y verificables. Este resultado se relaciona con Seclén Medina et al. (2025), quienes destacan que los modelos híbridos y las metodologías activas fortalecen la participación estudiantil y la autonomía académica en educación superior. En este sentido, las tareas basadas en casos, proyectos aplicados, avances parciales y registros de aprendizaje permiten que el estudiante no solo entregue un producto, sino que demuestre cómo lo construyó, qué decisiones tomó y qué criterios utilizó durante el proceso.

Otro aspecto relevante fue la percepción de vulnerabilidad de las evaluaciones tradicionales. Los estudiantes identificaron que las tareas escritas y los trabajos domiciliarios pueden ser desarrollados con apoyo de IA sin dejar suficiente evidencia del aprendizaje real. Este resultado dialoga con Revesai (2025), quien advierte sobre la dependencia académica hacia la inteligencia artificial generativa y sus posibles efectos en el rendimiento estudiantil. Por ello, el problema no se limita al uso de la tecnología, sino a la posibilidad de que el estudiante delegue procesos cognitivos esenciales, como analizar, argumentar, sintetizar y tomar decisiones académicas propias.

La defensa argumentativa apareció como uno de los componentes más relevantes dentro del modelo evaluativo propuesto. Este resultado permite sostener que la explicación oral, la justificación de decisiones y la capacidad de responder preguntas sobre el trabajo elaborado fortalecen la autenticidad de la evaluación. En concordancia con Quinde-Zambrano et al.



(2025), el uso de inteligencia artificial puede contribuir al desarrollo del pensamiento crítico cuando se integra con una orientación pedagógica adecuada. Sin embargo, para que ello ocurra, la evaluación debe exigir al estudiante un posicionamiento propio, capacidad de análisis y dominio real del contenido trabajado.

La declaración del uso de IA generativa obtuvo una valoración positiva, aunque menor en comparación con otros componentes evaluativos. Esto sugiere que todavía existe una cultura académica en construcción respecto al uso transparente de estas herramientas. Lucero-Baldevenites et al. (2025) señalan que el empleo de ChatGPT en actividades académicas por parte de docentes universitarios refleja una transformación progresiva de las prácticas educativas. No obstante, esta transformación requiere normas claras, criterios institucionales y acuerdos pedagógicos que orienten tanto a docentes como a estudiantes sobre cuándo y cómo puede utilizarse la IA en una actividad evaluativa.

Los resultados también permiten comprender que la resistencia evaluativa no debe interpretarse como prohibición absoluta del uso de IA generativa. Por el contrario, implica diseñar experiencias de evaluación donde el apoyo tecnológico pueda ser declarado, regulado y contrastado con evidencias del aprendizaje humano. Brunton et al. (2025) muestran que la inteligencia artificial generativa puede mejorar el desempeño de estudiantes en condiciones de desventaja, lo que permite reconocer su potencial inclusivo y formativo. Sin embargo, dicho potencial solo se aprovecha cuando la IA se utiliza como apoyo para aprender y no como sustituto del razonamiento estudiantil.

En conjunto, los hallazgos refuerzan la necesidad de avanzar hacia modelos de evaluación auténtica basados en cuatro componentes articulados: producto contextualizado, proceso visible, defensa argumentativa y uso ético declarado de IA. Estos componentes permiten responder a los desafíos actuales de la educación superior, porque reducen la dependencia de tareas genéricas, aumentan la trazabilidad del aprendizaje y fortalecen la validez de la evaluación. Así, el modelo propuesto no busca eliminar la inteligencia artificial del proceso académico, sino integrarla de manera crítica, ética y pedagógicamente responsable.

La discusión permite afirmar que la evaluación universitaria requiere una transición desde esquemas centrados en la entrega hacia modelos centrados en la demostración del aprendizaje. Esta transición implica que el estudiante evidencie no solo qué produjo, sino cómo lo pensó, cómo lo construyó, cómo utilizó la IA y cómo puede defender



académicamente sus decisiones. De esta manera, la evaluación auténtica se convierte en una alternativa pertinente para preservar la integridad académica y, al mismo tiempo, responder a las transformaciones generadas por la inteligencia artificial generativa en la educación superior.

5. CONCLUSIÓN

Los resultados permitieron evidenciar que la inteligencia artificial generativa se encuentra integrada en las prácticas académicas de los estudiantes universitarios, por lo que la evaluación tradicional centrada únicamente en productos finales resulta insuficiente para verificar el aprendizaje real, la autoría intelectual y el razonamiento propio del estudiante.

La evaluación auténtica se consolidó como una alternativa pertinente para enfrentar los desafíos generados por el uso no declarado de inteligencia artificial, debido a que permite incorporar tareas contextualizadas, resolución de casos, avances parciales, registros de proceso y criterios de valoración más vinculados con el desempeño académico real.

La defensa argumentativa del producto académico constituyó uno de los componentes más relevantes del modelo propuesto, ya que permite comprobar la comprensión del estudiante, la coherencia de sus decisiones, el dominio del contenido y la capacidad para justificar el proceso seguido durante la elaboración de una actividad evaluativa.

La declaración ética del uso de inteligencia artificial generativa debe incorporarse como un criterio permanente dentro de la evaluación universitaria, no desde una lógica punitiva, sino como una práctica de transparencia académica que permita diferenciar el apoyo tecnológico del aporte intelectual del estudiante.

El modelo de evaluación auténtica propuesto, basado en producto contextualizado, proceso visible, defensa argumentativa y uso ético declarado de IA, contribuye a fortalecer la validez, la trazabilidad y la resistencia de las evaluaciones universitarias frente al uso no declarado de herramientas generativas.

REFERENCIAS BIBLIOGRÁFICAS

Ahangama, N. (2026). Designing assessments in the generative AI era: A tailored assessment framework for ICT tertiary education. *International Journal of Educational Technology in Higher Education*, 23, 9. <https://doi.org/10.1186/s41239-026-00582-0>



- Armada Pacheco, J. M., Infante Rivera, L. de J., Landeo Quispe, A. S., Mencia Sánchez, N. G., & Franco Medina, J. L. (2025). Evaluación inteligente: Innovaciones en métodos de evaluación universitaria a través de TIC y algoritmos de IA. *e-Revista Multidisciplinaria del Saber*, 3, e-RMS09042025. <https://doi.org/10.61286/e-rms.v3i.120>
- Brunton, R. J., Rhazzafe, S., Moodley, R., Kuhn, S., Caraffini, F., Wilford, S., Higginbottom, R., Colreavy-Donnelly, S., & Gongora, M. (2025). Using generative artificial intelligence to enhance the performance of disadvantaged students in secondary education. *Social Sciences & Humanities Open*, 12, 102110. <https://doi.org/10.1016/j.ssaho.2025.102110>
- Cardozo, S. (2025). Inteligencia artificial desde la perspectiva de estudiantes universitarios paraguayos: un análisis segmentado sobre conocimiento, acceso y prácticas éticas. *Revista Científica en Ciencias Sociales*, 7, e701123. <https://doi.org/10.53732/rccsociales/e701123>
- Elshall, A. S., & Badir, A. (2025). Balancing AI-assisted learning and traditional assessment: The FACT assessment in environmental data science education. *Frontiers in Education*, 10, 1596462. <https://doi.org/10.3389/educ.2025.1596462>
- Fischer, I., Sweeney, S., Lucas, M., & Gupta, N. (2024). Making sense of generative AI for assessments: Contrasting student claims and assessor evaluations. *The International Journal of Management Education*, 22(3), 101081. <https://doi.org/10.1016/j.ijme.2024.101081>
- Kickbusch, S., Ashford-Rowe, K., Kemp, A., Boreland, J., & Huijser, H. (2025). Beyond detection: Redesigning authentic assessment in an AI-mediated world. *Education Sciences*, 15(11), 1537. <https://doi.org/10.3390/educsci15111537>
- Lucero-Baldevenites, E. V., Bueno-Fernández, M. M., Caballero-de Lamarque, C. P., & Bracho-Fuenmayor, P. L. (2025). El uso del ChatGPT en actividades académicas por parte de docentes universitarios. *Episteme Koinonía. Revista Electrónica de Ciencias de la Educación, Humanidades, Artes y Bellas Artes*, 8(16). <https://doi.org/10.35381/e.k.v8i16.4591>
- Mendoza Zaragoza, N. E. (2026). Rediseño de la evaluación universitaria en la era de la inteligencia artificial generativa: Del producto terminado a la visibilidad del proceso intelectual. *Estudios y Perspectivas Revista Científica y Académica*, 6(2), 1155–1175. <https://doi.org/10.61384/r.c.a.v6i2.1891>



- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Ncube, P. D. N., Dzvapatsva, G. P., Matobobo, C., & Ranga, M. M. (2026). Redefining student assessment in AI-infused learning environments: A systematic review of challenges and strategies for academic integrity. *AI and Ethics*, 6, Article 68. <https://doi.org/10.1007/s43681-025-00871-w>
- Ojeda, A. D., Solano-Barliza, A. D., Ortega Alvarez, D., & Boom Cárcamo, E. (2023). Análisis del impacto de la inteligencia artificial ChatGPT en los procesos de enseñanza y aprendizaje en la educación universitaria. *Formación Universitaria*, 16(6). <http://dx.doi.org/10.4067/S0718-50062023000600061>
- Quinde-Zambrano, L. F., Franco-Arroyo, P. P., & Reyes-Palau, N. C. (2025). Uso de la inteligencia artificial para el desarrollo del pensamiento crítico en Ciencias Sociales en la básica superior. *Sociedad & Tecnología*, 8(3), 438–456. <https://doi.org/10.51247/st.v8i3.596>
- Real González, P. V., Ruperti Muentes, E. R., Sumba Arévalo, V. M., & Boderó Arizaga, L. del C. (2026). Ética académica y mediación pedagógica ante la inteligencia artificial generativa en educación superior. *Sinergia Académica*, 9(4), 1–25. <https://doi.org/10.51736/sa1012>
- Revesai, Z. (2025). Generative AI dependency: The emerging academic crisis and its impact on student performance—A case study of a university in Zimbabwe. *Cogent Education*, 12(1), 2549787. <https://doi.org/10.1080/2331186X.2025.2549787>
- Seclén Medina, A., Valladares Cisneros, M. G., Ordóñez Zea, A. E., & Gamarra Rojas, S. (2025). Modelos híbridos y metodologías de aprendizaje activo en la educación superior: Enfoques innovadores para potenciar la participación estudiantil y el desarrollo de la autonomía académica. *Revista Social Fronteriza*, 5(6), e958. [https://doi.org/10.59814/resofro.2025.5\(6\)958](https://doi.org/10.59814/resofro.2025.5(6)958)
- Yang, H. (2026). Advancements in artificial intelligence and the reframing of student assessment in geography education. *Journal of Geography in Higher Education*. Advance online publication. <https://doi.org/10.1080/03098265.2026.2628342>



Yllescas Tandazo, L. C., Román Román, D. G., Varela Jácome, M. I., Aguilar Verdezoto, G. E., & Benalcázar Arroba, C. L. (2025). Percepción docente sobre la integración de la inteligencia artificial en los procesos de enseñanza-aprendizaje en universidades ecuatorianas: Teachers' perception of the integration of artificial intelligence in teaching and learning processes at Ecuadorian universities. *Revista Científica Multidisciplinar G-Nerando*, 6(2), 4031. <https://doi.org/10.60100/rcmg.v6i2.892>

Conflicto de Intereses: Los autores afirman que no existen conflictos de intereses en este estudio y que se han seguido éticamente los procesos establecidos por esta revista. Además, aseguran que este trabajo no ha sido publicado parcial ni totalmente en ninguna otra revista.

FINANCIAMIENTO

Los autores no recibieron financiamiento para el desarrollo de esta investigación.

CONTRIBUCIÓN DE AUTORÍA:

Nombres de autores e iniciales: Ginna Maricela Suárez Barberán (GMSB), Luis Germán Sánchez Méndez (LGSM), Cristian Raúl Quintana Serrano (CRQS), Patricia Jeanneth Miranda Ruiz (PJMR).

1. Conceptualización: (GMSB) (LGSM)
2. Curación de datos: (CRQS) (PJMR)
3. Análisis formal: (LGSM) (CRQS)
4. Adquisición de fondos: (GMSB)
5. Investigación: (GMSB) (PJMR)
6. Metodología: (LGSM) (CRQS)
7. Administración del proyecto: (GMSB)
8. Recursos: (PJMR) (CRQS)
9. Software: (CRQS)
10. Supervisión: (LGSM)
11. Validación: (LGSM) (PJMR)
12. Visualización: (CRQS)
13. Redacción – Borrador original: (GMSB) (CRQS)
14. Redacción – Revisión y edición: (LGSM) (PJMR)

