EDUR • Educação em Revista.2026;42;e67812
DOI: http://dx.doi.org/10.1590/0102-469867812T
Preprint: https://doi.org/10.1590/SciELOPreprints.17329

https://creativecommons.org/licenses/by/4.0/

ARTIGO

ALFABETIZACIÓN EN INTELIGENCIA ARTIFICIAL GENERATIVA Y EVALUACIÓN AUTÉNTICA EN EDUCACIÓN SUPERIOR: INTEGRIDAD ACADÉMICA[1]

GERMÁN VINICIO ROBLEZ TORRES[1]
ORCID: https://orcid.org/0009-0003-5061-860X
gvroblez@utpl.edu.ec
ALEXANDRA MARICRUZ ANDRADE APOLO[1]
ORCID: https://orcid.org/0009-0009-4525-1114
amandrade11@utpl.edu.ec
KATHERINE ELIZABETH GONZAGA AGUILAR[1]
ORCID: https://orcid.org/0009-0005-3541-3133
kegonzaga@utpl.edu.ec

[1] Universidad Técnica Particular de Loja. Loja, Loja, Ecuador.

RESUMEN: La inteligencia artificial (IA) generativa plantea desafíos para la autoría, la trazabilidad del aprendizaje y la honestidad académica en el ámbito universitario. Esta investigación examina las asociaciones entre alfabetización en inteligencia artificial generativa, evaluación auténtica e integridad académica en educación superior. Se adoptó un diseño cuantitativo, exploratorio, descriptivo y correlacional, con una muestra no probabilística por conveniencia de 603 participantes: 482 estudiantes y 121 docentes de una universidad ecuatoriana. Los datos se recogieron mediante un cuestionario de 30 ítems, organizado en cinco dominios: alfabetización en IA generativa, uso académico de herramientas generativas, evaluación auténtica, integridad académica y trazabilidad del aprendizaje. El instrumento fue revisado mediante juicio de expertos respecto de su claridad y contenido, y presentó consistencia interna adecuada por dominio y alta en el conjunto de los ítems. Los resultados mostraron medias superiores al punto medio de la escala en todos los dominios. La integridad académica obtuvo la media más alta, seguida de la alfabetización en IA generativa, mientras que la trazabilidad del aprendizaje presentó el valor más bajo. Las correlaciones fueron positivas y estadísticamente significativas, especialmente entre evaluación auténtica, trazabilidad del aprendizaje e integridad académica. Se concluye que estas dimensiones presentan asociaciones en la muestra estudiada, lo que sugiere la pertinencia de considerarlas en el diseño de evaluaciones universitarias más situadas, transparentes y verificables, sin establecer relaciones causales.

[1] Artículo publicado con financiamiento de la *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* - CAPES/Brasil para los servicios de edición, diagramación y conversión de XML.

Palabras clave: Inteligencia artificial; Enseñanza superior; Evaluación del estudiante; Ética; Tecnología educacional.

ALFABETIZAÇÃO EM INTELIGÊNCIA ARTIFICIAL GENERATIVA E AVALIAÇÃO AUTÊNTICA NO ENSINO SUPERIOR: INTEGRIDADE ACADÊMICA

RESUMO: A inteligência artificial (IA) generativa apresenta desafios para a autoria, a rastreabilidade da aprendizagem e a honestidade acadêmica no âmbito universitário. Esta investigação examina as associações entre alfabetização em inteligência artificial generativa, avaliação autêntica e integridade acadêmica no ensino superior. Adotou-se um delineamento quantitativo, exploratório, descritivo e correlacional, com uma amostra não probabilística por conveniência de 603 participantes: 482 estudantes e 121 docentes de uma universidade equatoriana. Os dados foram coletados por meio de um questionário de 30 itens, organizado em cinco domínios: alfabetização em IA generativa, uso acadêmico de ferramentas generativas, avaliação autêntica, integridade acadêmica e rastreabilidade da aprendizagem. O instrumento foi revisado mediante julgamento de especialistas quanto à clareza e ao conteúdo e apresentou consistência interna adequada por domínio e elevada no conjunto dos itens. Os resultados mostraram médias superiores ao ponto médio da escala em todos os domínios. A integridade acadêmica apresentou a maior média, seguida da alfabetização em IA generativa, enquanto a rastreabilidade da aprendizagem apresentou o menor valor. As correlações foram positivas e estatisticamente significativas, especialmente entre avaliação autêntica, rastreabilidade da aprendizagem e integridade acadêmica. Conclui-se que essas dimensões apresentam associações na amostra estudada, o que sugere a pertinência de considerá-las no planejamento de avaliações universitárias mais situadas, transparentes e verificáveis, sem estabelecer relações causai.

Palavras-chave: Inteligência artificial; Ensino superior; Avaliação do estudante; Ética; Tecnologia educacional.

GENERATIVE ARTIFICIAL INTELLIGENCE LITERACY AND AUTHENTIC ASSESSMENT IN HIGHER EDUCATION: ACADEMIC INTEGRITY

ABSTRACT: Generative artificial intelligence (AI) poses challenges to authorship, learning traceability, and academic honesty in university settings. This study examines the associations among generative artificial intelligence literacy, authentic assessment, and academic integrity in higher education. A quantitative, exploratory, descriptive, and correlational design was adopted, using a nonprobability convenience sample of 603 participants: 482 students and 121 faculty members from an Ecuadorian university. Data were collected through a 30-item questionnaire organized into five domains: generative AI literacy, academic use of generative tools, authentic assessment, academic integrity, and learning traceability. The instrument was reviewed through expert judgment for clarity and content and demonstrated adequate internal consistency within each domain and high internal consistency across the full set of items. Mean scores exceeded the scale midpoint in all domains. Academic integrity had the highest mean score, followed by generative AI literacy, whereas learning traceability had the lowest value. Correlations were positive and statistically significant, particularly among authentic assessment, learning traceability, and academic integrity. These dimensions showed associations within the study sample, suggesting the relevance of considering them when designing more situated, transparent, and verifiable university assessments, without establishing causal relationships.

Keywords: Artificial intelligence; Higher education; Student assessment; Ethics; Educational technology.

INTRODUCCIÓN

La inteligencia artificial generativa se ha incorporado con rapidez a las prácticas de aprendizaje, escritura académica, búsqueda de información y producción de evidencias en la educación superior. Su expansión ha modificado la relación entre estudiante, conocimiento y evaluación, debido a que permite generar textos, sintetizar fuentes, resolver tareas y producir argumentos con niveles crecientes de fluidez. Este escenario no elimina la necesidad de formar competencias académicas; por el contrario, exige que las universidades definan criterios pedagógicos, éticos y evaluativos para orientar su uso (Crompton & Burke, 2023; Chen et al., 2020).

La discusión actual ya no se limita a permitir o prohibir estas herramientas. El problema central consiste en determinar cómo integrarlas sin debilitar la autoría, la trazabilidad del aprendizaje y la validez de la evaluación. La UNESCO (2025) sostiene que "la inteligencia artificial en educación debe responder a un enfoque centrado en el ser humano, con atención a la equidad, la protección de datos, la agencia humana y el desarrollo de competencias para estudiantes y docentes". En este marco, la alfabetización en inteligencia artificial generativa se convierte en una condición necesaria para que los actores universitarios comprendan sus posibilidades, límites y riesgos.

En educación superior, la evaluación constituye uno de los ámbitos más tensionados por la inteligencia artificial generativa (Perkins, 2023). Las tareas escritas tradicionales, los ensayos domiciliarios y las actividades centradas solo en productos finales presentan mayor vulnerabilidad cuando no incorporan procesos verificables de elaboración, retroalimentación y defensa del aprendizaje. Frente a esta situación, la evaluación auténtica adquiere relevancia porque vincula el desempeño académico con problemas, contextos y tareas próximas a la práctica profesional. Su aporte no radica únicamente en cambiar el formato de las actividades, sino en fortalecer la relación entre desempeño, criterio evaluativo y evidencia de aprendizaje (Herrington & Oliver, 2000; Villarroel et al., 2017).

La literatura reciente identifica la evaluación auténtica como una alternativa pertinente para responder a los desafíos de la inteligencia artificial generativa, especialmente cuando las actividades exigen análisis contextual, juicio profesional, toma de decisiones, reflexión crítica y aplicación situada del conocimiento. Sin embargo, este enfoque no resuelve por sí solo los problemas de integridad académica. La autenticidad de una tarea puede debilitarse si el estudiante delega procesos cognitivos relevantes en sistemas automatizados sin declarar su uso, sin comprender los resultados generados o sin demostrar dominio del proceso seguido (Kasneci et al., 2023).

La intersección entre la alfabetización en inteligencia artificial generativa, la evaluación auténtica y la integridad académica permanece inexplorada en la literatura actual. Si bien abundan diagnósticos sobre el uso de herramientas generativas en el ámbito universitario y sus amenazas asociadas al fraude, se carece de evidencia empírica que vincule la competencia tecnológica del estudiante con el diseño de evaluaciones formativas y transparentes (Long & Magerko, 2020). Abordar este vacío resulta crítico para las universidades que intentan reformular sus marcos normativos, superando el enfoque punitivo tradicional basado estrictamente en software de detección y vigilancia.

Por tanto, el objetivo de este estudio fue analizar la relación entre alfabetización en inteligencia artificial generativa, evaluación auténtica e integridad académica en educación superior, con el fin de aportar criterios pedagógicos para el diseño de evaluaciones universitarias más válidas, éticas y alineadas con los desafíos actuales de la formación profesional.

METODOLOGÍA

Caracterización de la investigación

Mediante un abordaje empírico-cuantitativo de diseño correlacional, se evaluó la convergencia entre alfabetización en inteligencia artificial generativa, integridad académica y evaluación auténtica (Bretag et al., 2019). El análisis se fundamentó en datos de primera mano obtenidos de la comunidad docente y estudiantil de educación superior.

El diseño exploratorio permitió abordar un fenómeno educativo en consolidación. El alcance descriptivo caracterizó usos, percepciones y prácticas evaluativas vinculadas con herramientas de inteligencia artificial generativa (Crompton & Burke, 2023). El alcance correlacional examinó la asociación entre los dominios del instrumento.

En primer lugar, se realizó una revisión bibliográfica sobre inteligencia artificial generativa en educación superior, alfabetización en inteligencia artificial, evaluación auténtica e integridad académica (Gulikers et al., 2004). Esta revisión permitió delimitar el problema de investigación, definir los dominios del instrumento y formular los ítems del cuestionario.

La definición de los dominios del cuestionario se realizó a partir de una revisión bibliográfica sobre inteligencia artificial en educación superior, alfabetización en IA, evaluación auténtica e integridad académica (Ng et al., 2021). Esta revisión permitió transformar el problema de investigación en dimensiones observables: alfabetización en IA generativa, uso académico de herramientas generativas, evaluación auténtica, integridad académica y trazabilidad del aprendizaje. Los ítems iniciales se formularon con base en estos dominios y fueron revisados posteriormente por especialistas.

Participantes y escenario de investigación

El estudio se desarrolló en una institución de educación superior ecuatoriana, identificada mediante el código IES-A para preservar su confidencialidad en la presentación de los resultados. Se omitieron el nombre institucional y los detalles de ubicación, carreras, asignaturas y autoridades que pudieran facilitar su identificación.

La población accesible estuvo conformada por estudiantes de pregrado y docentes vinculados con actividades evaluativas mediadas por recursos digitales. La muestra incluyó 603 participantes, distribuidos entre 482 estudiantes y 121 docentes. La inclusión de ambos grupos respondió al interés de examinar perspectivas complementarias del proceso evaluativo: los estudiantes participan en la realización de tareas y la producción de evidencias de aprendizaje, mientras que los docentes intervienen en el diseño, la aplicación y la retroalimentación de las evaluaciones. Esta complementariedad justifica su participación en el estudio, pero no supone que sus respuestas sean intercambiables ni que deban interpretarse como una medida homogénea de la comunidad universitaria.

El cuestionario incorporó instrucciones diferenciadas para contextualizar las respuestas según el rol institucional. Los estudiantes respondieron considerando las evaluaciones que realizan en sus asignaturas, mientras que los docentes consideraron aquellas que diseñan, aplican o retroalimentan. No obstante, algunos ítems referidos a la elaboración de trabajos, el aprovechamiento de la retroalimentación y la defensa del proceso seguido pueden adquirir significados distintos entre grupos. Por tanto, las instrucciones diferenciadas no se consideraron evidencia suficiente de equivalencia de medición.

En atención a estas diferencias, se realizaron análisis descriptivos y correlaciones entre dominios por separado para estudiantes y docentes, así como comparaciones exploratorias de las puntuaciones observadas entre ambos grupos. Estas comparaciones no se interpretaron como diferencias concluyentes en los constructos subyacentes, dado que no se evaluó la invariancia de medición del instrumento. Las asociaciones obtenidas en la muestra conjunta se consideraron resultados agregados complementarios, cuya interpretación requiere contrastarlos con los análisis específicos de cada grupo.

La selección de participantes se realizó mediante muestreo no probabilístico por conveniencia. Se consideraron elegibles los estudiantes mayores de edad matriculados en pregrado y los docentes activos con experiencia en evaluación universitaria, previa aceptación del consentimiento informado. El protocolo estableció como criterios de exclusión la ausencia de consentimiento, el incumplimiento de los requisitos de participación, las respuestas incompletas en más del 20 % de los ítems, la duplicación evidente y los patrones de respuesta inconsistentes. Las respuestas parciales dentro del límite permitido se conservaron; para calcular la puntuación de cada dominio se requirieron al menos cuatro respuestas válidas de sus seis ítems. Debido al procedimiento de selección y a la participación de una única institución, los resultados se circunscriben a la muestra analizada y no permiten generalizaciones estadísticas a la educación superior ecuatoriana.

Declaración de aspectos éticos

La investigación se desarrolló con observancia de los principios de autonomía, voluntariedad, privacidad, confidencialidad y minimización del riesgo aplicables a la investigación con personas. Se utilizó una encuesta anónima, voluntaria y autoadministrada, dirigida exclusivamente a participantes mayores de edad, sin realizar intervenciones experimentales, manipulación de variables, procedimientos invasivos, recolección de muestras biológicas ni otras actuaciones que pudieran afectar su integridad física o psicológica. Antes de acceder al cuestionario, los participantes recibieron información sobre la finalidad académica de la investigación, el carácter voluntario de su participación, el tratamiento de la información, las medidas destinadas a preservar la confidencialidad y su facultad de abandonar el formulario antes de enviarlo, tras lo cual manifestaron electrónicamente su consentimiento para participar.

Estas salvaguardas se establecieron en concordancia con el artículo 66, numeral 19, de la Constitución de la República del Ecuador, que reconoce el derecho a la protección de los datos de carácter personal y dispone que su recolección y tratamiento requieren autorización del titular o mandato legal, así como con los principios previstos en la Ley Orgánica de Protección de Datos Personales, particularmente los relativos al consentimiento libre, específico, informado e inequívoco, la finalidad, la pertinencia y la minimización de los datos, la confidencialidad y la seguridad de la información (Asamblea Nacional del Ecuador, 2008, 2021). Asimismo, la base utilizada para el análisis estuvo constituida exclusivamente por respuestas anónimas y los resultados fueron procesados y comunicados de manera agregada, sin conservar elementos que permitieran identificar o reidentificar individualmente a los participantes, en consonancia con las disposiciones ecuatorianas relativas a la protección y la anonimización de los datos personales.

En cuanto a la evaluación ética institucional, el estudio no fue sometido previamente a un Comité de Ética de Investigación en Seres Humanos (CEISH). Esta circunstancia se declara expresamente y no se presenta como equivalente a una aprobación, dispensa o exención formal otorgada por un comité. No obstante, por sus características metodológicas, la investigación correspondió a un diseño observacional, no intervencional y de bajo riesgo, basado en la recopilación anónima de información mediante un cuestionario, sin datos identificadores, datos sensibles, muestras biológicas, participación de población vulnerable ni indagación sobre aspectos sensibles de la conducta. Tales características son consistentes con los criterios previstos en la normativa ecuatoriana para identificar investigaciones sin riesgo

susceptibles de evaluación exenta, conforme al Reglamento para la aprobación, desarrollo, vigilancia y control de investigaciones observacionales y estudios de intervención en seres humanos y su reforma mediante el Acuerdo Ministerial No. 00038-2021 (Ministerio de Salud Pública del Ecuador, 2021), así como con los criterios operativos establecidos para este tipo de investigaciones por los CEISH autorizados por la autoridad sanitaria nacional.

La ausencia de evaluación previa por un CEISH constituye, por tanto, una condición procedimental que se informa de manera transparente, sin atribuir al estudio una exención ética que no fue formalmente emitida. De acuerdo con el artículo 47 del Reglamento Sustitutivo para la Aprobación y Seguimiento de los Comités de Ética de Investigación en Seres Humanos, expedido mediante Acuerdo Ministerial No. 00005-2022, los CEISH no evalúan protocolos cuya ejecución ya se encuentre en desarrollo o haya concluido, por lo que no corresponde gestionar retrospectivamente una aprobación ética para la investigación ya ejecutada (Ministerio de Salud Pública del Ecuador, 2022). En consecuencia, la protección de los participantes se sustentó en la participación voluntaria, el consentimiento informado electrónico, la ausencia de intervención, la recopilación anónima de los datos, la imposibilidad de identificar individualmente a los participantes y la presentación exclusivamente agregada de los resultados.

Instrumento de investigación

El instrumento de investigación fue un cuestionario estructurado de 30 ítems, organizado en cinco dominios con seis ítems cada uno. Los dominios fueron: alfabetización en inteligencia artificial generativa, uso académico de herramientas generativas, evaluación auténtica, integridad académica y trazabilidad del aprendizaje.

Tabla 1. *Dominios del cuestionario*

Dominio	Propósito	Ítems orientadores
Alfabetización en inteligencia artificial generativa	Evaluar conocimientos y criterios de uso crítico.	- Comprendo las funciones básicas de la IA generativa. - Verifico la información generada. - Reconozco límites, sesgos y riesgos académicos.
Uso académico de herramientas generativas	Identificar finalidades de uso en tareas universitarias.	- Uso IA generativa para organizar ideas. - La utilizo para revisar textos. - La utilizo como apoyo para estudiar o sintetizar información.
Evaluación auténtica	Valorar tareas centradas en desempeño aplicado.	- Las evaluaciones exigen resolver problemas contextualizados. - Las tareas requieren argumentación propia. - Los criterios se comunican mediante rúbricas.
Integridad académica	Examinar autoría, transparencia y responsabilidad académica.	- Declaro el uso de IA en tareas académicas. - Cito las fuentes utilizadas. - Asumo responsabilidad sobre el contenido entregado.

Trazabilidad del aprendizaje	Identificar evidencias del proceso seguido por el estudiante.	- Presento avances o borradores. - Justifico decisiones durante la tarea. - Puedo defender oralmente el proceso seguido.

Nota. Elaborado por los autores

Los 30 ítems se respondieron mediante una escala de frecuencia de cinco categorías ordenadas: nunca (1), rara vez (2), algunas veces (3), frecuentemente (4) y siempre (5). Adicionalmente, se incluyó la alternativa «no puedo responder», registrada en la base como valor perdido y excluida de la puntuación numérica. Esta alternativa no constituyó un nivel adicional de la escala ni se interpretó como una respuesta intermedia.

La alternativa «no puedo responder» se codificó como NA y se trató como dato ausente, sin sustituirla por cero, por el punto medio de la escala ni por un valor imputado. Para cada participante, la puntuación de cada dominio se calculó mediante la media aritmética de las respuestas válidas a sus seis ítems, siempre que hubiera respondido al menos cuatro ítems, conforme al protocolo de codificación. El denominador del cálculo correspondió al número de ítems efectivamente respondidos: cuatro, cinco o seis, según el caso. Posteriormente, la media de cada dominio para la muestra se obtuvo promediando las puntuaciones individuales disponibles.

En la base analizada, el número de respuestas válidas por ítem osciló entre 582 y 597, debido a la presencia de valores ausentes. Sin embargo, los 603 participantes respondieron al menos cuatro ítems en cada uno de los cinco dominios, por lo que todos dispusieron de una puntuación válida en cada dominio. Esto explica que los análisis por ítem presenten tamaños válidos diferentes, mientras que los análisis por dominio mantengan $n = 603$, sin que haya sido necesario imputar los valores ausentes.

El cuestionario incluyó una sección final de respuesta abierta. Esta sección recogió observaciones sobre el uso de inteligencia artificial generativa en la evaluación universitaria. Las respuestas abiertas se utilizaron como insumo complementario para interpretar los resultados descriptivos.

Antes de su aplicación, el instrumento fue sometido a validación aparente y de contenido. La validación aparente evaluó claridad, comprensión y ausencia de ambigüedad. La validación de contenido examinó pertinencia, representatividad y suficiencia de los ítems respecto de cada dominio. Este procedimiento siguió criterios para el desarrollo y la validación de instrumentos en investigación educativa y social (Polit & Beck, 2006; Boateng et al., 2018).

La validez de contenido del cuestionario se estableció mediante el juicio de tres expertos especializados en educación superior, evaluación, tecnología educativa e integridad académica, quienes examinaron la correspondencia teórica entre dominios, indicadores e ítems. A partir de sus dictámenes cualitativos, se optimizó la precisión conceptual de los reactivos y se eliminaron redundancias sintácticas.

Aplicación del cuestionario

Previa autorización institucional, el cuestionario se administró en línea mediante Google Forms entre abril y mayo de 2026. La invitación se remitió a docentes y estudiantes por canales institucionales autorizados. El enlace de acceso presentó primero el consentimiento informado; solo los participantes que aceptaron sus condiciones pudieron responder el instrumento.

El cuestionario se organizó en tres secciones. La primera recogió datos generales: rol institucional, edad, sexo, área académica, experiencia docente o administrativa, frecuencia de uso de herramientas

digitales y experiencia previa con inteligencia artificial generativa. La segunda incluyó los 30 ítems distribuidos en los cinco dominios del estudio. La tercera incorporó una pregunta abierta para registrar observaciones sobre el uso de inteligencia artificial generativa en la evaluación universitaria.

Con el fin de mitigar riesgos éticos y asegurar la confidencialidad, la recolección de datos prescindió de identificadores personales explícitos, operando bajo un esquema autoadministrado, voluntario y gratuito. El tratamiento de las respuestas incompletas siguió los criterios del protocolo de codificación: se conservaron los registros dentro del límite permitido de datos ausentes y se calcularon las puntuaciones de los dominios cuando existían al menos cuatro respuestas válidas.

La autorización institucional permitió distribuir el cuestionario por canales formales y delimitó el uso académico de la información. Los datos fueron tratados de forma agregada para proteger la identidad de los participantes y de la institución codificada como IES-A.

Análisis estadístico

Los datos fueron organizados en una matriz y depurados antes del análisis. La caracterización de los participantes se realizó mediante frecuencias y porcentajes. Los ítems y dominios se describieron mediante la media, la desviación estándar y la varianza.

La consistencia interna se examinó mediante el alfa de Cronbach por dominio, utilizando los casos con respuestas completas en los seis ítems correspondientes. De manera complementaria, se calculó el coeficiente para los 30 ítems con los casos completos del cuestionario. No se imputaron valores ausentes para estos cálculos. Los coeficientes se interpretaron como evidencia de consistencia interna en la muestra, sin atribuirles capacidad para demostrar la estructura factorial o la validez de constructo del instrumento.

Las asociaciones entre las puntuaciones de los dominios se examinaron mediante el coeficiente de correlación de Spearman, basado en los rangos, para describir relaciones monotónicas entre puntuaciones derivadas de ítems ordinales sin exigir normalidad bivariada. Se realizaron pruebas bilaterales con un nivel de significación de 0,05. Las correlaciones se calcularon separadamente para estudiantes y docentes; los resultados de la muestra conjunta se consideraron complementarios. Los coeficientes se interpretaron como medidas de asociación, sin atribuirles significado causal.

Las comparaciones exploratorias entre estudiantes y docentes se realizaron mediante la prueba U de Mann–Whitney bilateral, con aproximación asintótica y corrección por empates y continuidad. Se aplicó el procedimiento de Holm a las cinco comparaciones entre grupos y, separadamente, a las diez correlaciones de cada grupo. Se consideraron estadísticamente significativos los valores p ajustados inferiores a 0,05.

RESULTADOS

Caracterización de los participantes

La muestra estuvo conformada por 603 participantes: 482 estudiantes y 121 docentes. Los estudiantes representaron el 79,9 % de la muestra y los docentes, el 20,1 %. La mayor proporción de participantes se ubicó en el grupo de 18 a 24 años. Respecto al sexo, predominó la participación femenina.

Tabla 2. *Caracterización de los participantes*

Variable	Categoría	n	%
Rol institucional	Estudiante	482	79,9
	Docente	121	20,1
Sexo	Femenino	353	58,5
	Masculino	246	40,8
	Prefiero no responder	4	0,7
Edad	18-24 años	390	64,7
	25-34 años	99	16,4
	35-44 años	60	10,0
	45-54 años	42	7,0
	55 años o más	12	2,0
Área académica	Ciencias sociales, educación o humanidades	205	34,0
	Ciencias administrativas o económicas	151	25,0
	Ciencias de la salud	117	19,4
	Ingeniería, tecnología o ciencias exactas	95	15,8
	Otra	35	5,8
Frecuencia de uso de herramientas digitales	Rara vez	18	3,0
	Algunas veces	105	17,4
	Frecuentemente	281	46,6
	Siempre	199	33,0
Experiencia previa con IA generativa	No la he utilizado	39	6,5
	Una o dos veces	77	12,8
	Ocasional	181	30,0
	Frecuente	208	34,5
	Habitual	98	16,3

Nota. Elaborado por los autores

Análisis descriptivo por dominio

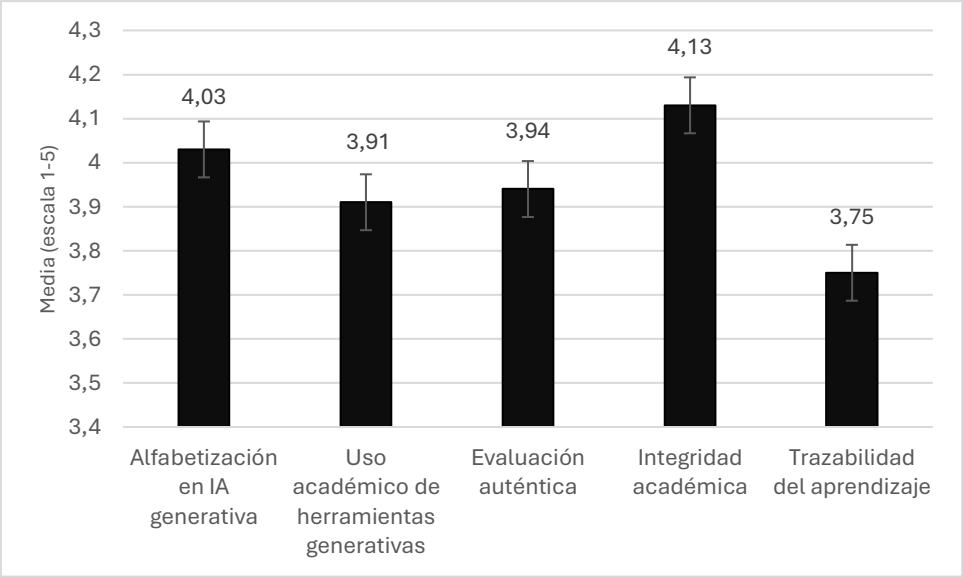
Los cinco dominios presentaron medias superiores al punto medio de la escala. El dominio con mayor media fue la integridad académica (M = 4,13; DE = 0,53), seguido de la alfabetización en IA generativa (M = 4,03; DE = 0,54). El dominio con menor media fue la trazabilidad del aprendizaje (M = 3,75; DE = 0,59). Este patrón sugiere una percepción favorable hacia la integridad académica y la alfabetización en IA, con puntuaciones autoinformadas menores en trazabilidad del aprendizaje.

Tabla 3. *Estadísticos descriptivos por dominio*

Dominio	n válido	Media	DE	Varianza
Alfabetización en IA generativa	603	4,03	0,54	0,29
Uso académico de herramientas generativas	603	3,91	0,58	0,34
Evaluación auténtica	603	3,94	0,55	0,30
Integridad académica	603	4,13	0,53	0,28
Trazabilidad del aprendizaje	603	3,75	0,59	0,35

Nota. Elaborado por los autores

Figura 1. *Promedio por dominio del cuestionario*



Nota. Elaborado por los autores

Tabla 4. *Descriptivos por dominio y comparación exploratoria entre estudiantes y docentes*

Dominio	Estudiantes: M (DE)	Docentes: M (DE)	U de Mann–Whitney	p sin ajustar	p ajustado
Alfabetización en IA generativa	4,03 (0,55)	4,03 (0,49)	29 364,0	0,906	0,906
Uso académico de herramientas generativas	3,96 (0,56)	3,75 (0,63)	34 242,5	0,003	0,015
Evaluación auténtica	3,93 (0,55)	3,99 (0,56)	27 270,5	0,269	0,537
Integridad académica	4,11 (0,54)	4,19 (0,49)	26 856,5	0,177	0,531
Trazabilidad del aprendizaje	3,73 (0,58)	3,85 (0,61)	25 480,0	0,031	0,125

Nota. Elaboración propia a partir de la base de datos del estudio. Estudiantes: n = 482; docentes: n = 121. M = media; DE = desviación estándar. Las puntuaciones se expresan en una escala de 1 a 5. Las comparaciones se realizaron mediante la prueba U de Mann–Whitney bilateral, con aproximación asintótica, corrección por empates y continuidad. Se informa el estadístico U correspondiente al grupo de estudiantes. Los valores p ajustados corresponden al procedimiento de Holm aplicado a las cinco comparaciones. La prueba contrasta las distribuciones de las puntuaciones mediante sus rangos; no constituye una prueba de diferencias entre medias.

La comparación exploratoria entre grupos mostró que, tras el ajuste de Holm, únicamente el uso académico de herramientas generativas presentó una diferencia estadísticamente significativa (p ajustado = 0,015), con puntuaciones más altas entre los estudiantes. En los restantes dominios no se identificaron diferencias estadísticamente significativas después del ajuste, aunque este resultado no demuestra equivalencia entre grupos. Las diferencias observadas se circunscriben a las puntuaciones de autoinforme de la muestra y deben interpretarse considerando las distintas funciones de estudiantes y docentes en el proceso evaluativo y la ausencia de evidencia sobre equivalencia de medición entre ambos grupos.

Análisis descriptivo por ítem

El análisis por ítem permitió describir las puntuaciones de los reactivos de cada dominio. En alfabetización en IA generativa, las medias oscilaron entre 3,96 y 4,09. En uso académico de herramientas generativas, los valores oscilaron entre 3,86 y 3,98. En evaluación auténtica, las medias fluctuaron entre 3,90 y 4,06. En integridad académica, todos los ítems se mantuvieron por encima de 4,08. En trazabilidad del aprendizaje, las medias fueron menores y oscilaron entre 3,70 y 3,82.

Tabla 5. *Codificación de ítems*

Dominio	Ítems
Alfabetización en IA generativa	AIG1-AIG6
Uso académico de herramientas generativas	UAG1-UAG6
Evaluación auténtica	EA1-EA6
Integridad académica	IA1-IA6
Trazabilidad del aprendizaje	TA1-TA6

Nota. Elaborado por los autores

Tabla 6. *Estadísticos descriptivos por ítem*

Código	Dominio	n válido	Media	DE	Varianza
AIG1	AIG	593	4,09	0,72	0,52
AIG2	AIG	595	4,06	0,71	0,50
AIG3	AIG	595	3,96	0,74	0,54
AIG4	AIG	593	4,06	0,69	0,48
AIG5	AIG	594	4,00	0,71	0,50
AIG6	AIG	593	4,02	0,72	0,52
UAG1	UAG	588	3,96	0,75	0,56
UAG2	UAG	588	3,98	0,76	0,57
UAG3	UAG	585	3,86	0,74	0,54
UAG4	UAG	586	3,88	0,76	0,57
UAG5	UAG	592	3,86	0,76	0,58
UAG6	UAG	591	3,95	0,73	0,54
EA1	EA	590	3,90	0,71	0,51
EA2	EA	584	3,93	0,75	0,56
EA3	EA	582	3,93	0,73	0,53

EA4	EA	582	4,06	0,72	0,51
EA5	EA	593	3,92	0,73	0,54
EA6	EA	588	3,90	0,73	0,53
IA1	IA	593	4,19	0,69	0,47
IA2	IA	597	4,11	0,71	0,51
IA3	IA	591	4,09	0,72	0,52
IA4	IA	596	4,19	0,68	0,46
IA5	IA	592	4,09	0,70	0,50
IA6	IA	587	4,09	0,73	0,53
TA1	TA	585	3,72	0,77	0,60
TA2	TA	591	3,76	0,76	0,58
TA3	TA	586	3,77	0,75	0,57
TA4	TA	593	3,72	0,75	0,56
TA5	TA	597	3,70	0,76	0,58
TA6	TA	594	3,82	0,76	0,58

Nota. Elaborado por los autores

Confiabilidad del instrumento

La consistencia interna se examinó mediante el coeficiente alfa de Cronbach, utilizando, para cada dominio, los participantes con respuestas completas en sus seis ítems, sin imputar valores ausentes. Los coeficientes oscilaron entre 0,846 y 0,867. Como análisis complementario, el alfa del conjunto de 30 ítems fue de 0,918 y se calculó con los 325 participantes que respondieron todos los ítems. Estos resultados aportan evidencia de consistencia interna en las puntuaciones de la muestra analizada, pero no demuestran la unidimensionalidad del cuestionario, la estructura de cinco dominios ni la equivalencia de medición entre estudiantes y docentes.

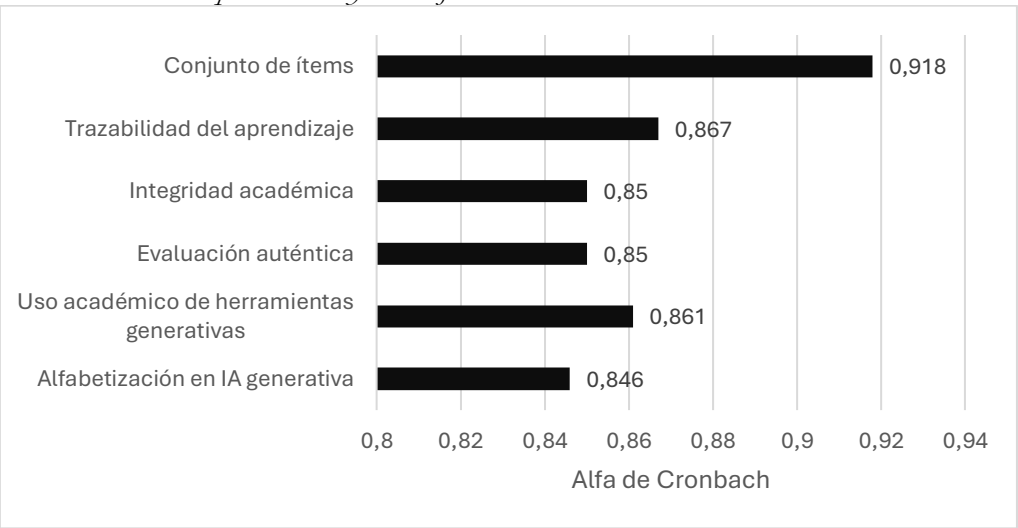
Tabla 7. *Consistencia interna por dominio y del conjunto de ítems*

Dominio o conjunto de ítems	Número de ítems	Participantes con respuestas completas	Alfa de Cronbach
Alfabetización en IA generativa	6	550	0,846
Uso académico de herramientas generativas	6	522	0,861
Evaluación auténtica	6	510	0,850
Integridad académica	6	542	0,850
Trazabilidad del aprendizaje	6	535	0,867
Conjunto de ítems	30	325	0,918

Nota. Elaboración propia a partir de la base de datos del estudio. Cada coeficiente se calculó excluyendo los casos con alguna respuesta ausente en el conjunto de ítems correspondiente, sin imputación. Los tamaños analíticos de esta tabla corresponden exclusivamente al cálculo del alfa y no modifican los 603 casos disponibles para los análisis de puntuaciones por dominio. El coeficiente del conjunto de ítems no demuestra la existencia de una única dimensión ni justifica, por sí solo, una puntuación global.

La revisión aparente y de contenido por especialistas y el análisis de consistencia interna corresponden a aspectos diferentes de la evaluación del instrumento. La primera examina la claridad y la correspondencia conceptual entre los ítems y los dominios propuestos, mientras que el alfa aporta información sobre su interrelación. En consecuencia, los cinco dominios se consideran una organización teórica del cuestionario cuya estructura interna requiere evaluación adicional mediante análisis factoriales; no se presentan como una estructura empíricamente validada a partir de los coeficientes obtenidos.

Figura 2. *Consistencia interna por dominio y del conjunto de ítems*



Nota. Elaboración propia. Coeficientes calculados con casos completos para cada dominio y para el conjunto de 30 ítems, sin imputación. Los valores no constituyen evidencia de validez de constructo.

Correlaciones entre dominios

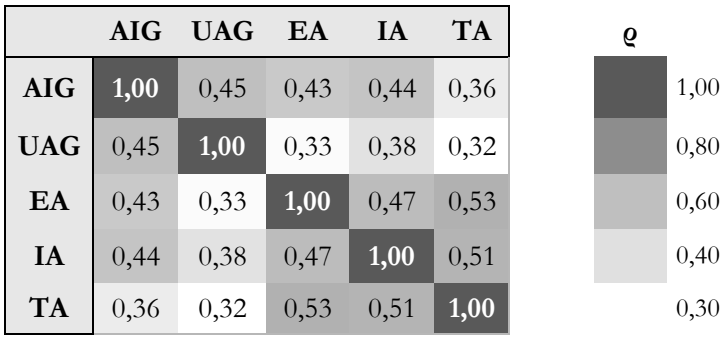
Todas las correlaciones entre dominios fueron positivas y estadísticamente significativas ($p < 0{,}001$). Las asociaciones más altas se observaron entre evaluación auténtica y trazabilidad del aprendizaje (rho = 0,53), integridad académica y trazabilidad del aprendizaje (rho = 0,51) y evaluación auténtica e integridad académica (rho = 0,47). En la muestra conjunta, estos resultados describen asociaciones positivas entre las puntuaciones autoinformadas de evaluación auténtica, trazabilidad del aprendizaje e integridad académica.

Tabla 8. *Correlaciones de Spearman entre dominios*

Dominio	AIG	UAG	EA	IA	TA
AIG	1,00	0,45	0,43	0,44	0,36
UAG	0,45	1,00	0,33	0,38	0,32
EA	0,43	0,33	1,00	0,47	0,53
IA	0,44	0,38	0,47	1,00	0,51
TA	0,36	0,32	0,53	0,51	1,00

Nota. Elaborado por los autores

Figura 3. *Matriz de correlaciones Spearman*



Nota. Elaborado por los autores

Las correlaciones de Spearman se calcularon separadamente para estudiantes y docentes, considerando las diez asociaciones posibles entre los cinco dominios. Todas fueron positivas y estadísticamente significativas en ambos grupos, incluso después del ajuste de Holm aplicado a las diez pruebas de cada grupo (p ajustado < 0,001). Los coeficientes oscilaron entre 0,333 y 0,530 en estudiantes y entre 0,299 y 0,516 en docentes.

Tabla 9. *Correlaciones de Spearman entre dominios, diferenciadas por rol institucional*

Asociación entre dominios	Estudiantes: ρ	Docentes: ρ
Alfabetización en IA generativa y uso académico de herramientas generativas	0,473	0,360
Alfabetización en IA generativa y evaluación auténtica	0,459	0,299
Alfabetización en IA generativa e integridad académica	0,468	0,343
Alfabetización en IA generativa y trazabilidad del aprendizaje	0,373	0,320
Uso académico de herramientas generativas y evaluación auténtica	0,334	0,357
Uso académico de herramientas generativas e integridad académica	0,366	0,492
Uso académico de herramientas generativas y trazabilidad del aprendizaje	0,333	0,361
Evaluación auténtica e integridad académica	0,466	0,477
Evaluación auténtica y trazabilidad del aprendizaje	0,530	0,516
Integridad académica y trazabilidad del aprendizaje	0,507	0,500

Nota. Elaboración propia a partir de la base de datos del estudio. ρ = coeficiente de correlación de Spearman. Estudiantes: n = 482; docentes: n = 121, en todas las asociaciones. Las pruebas fueron bilaterales; todos los valores p fueron inferiores a 0,001, tanto antes como después del ajuste de Holm realizado separadamente para las diez correlaciones de cada grupo. Los coeficientes describen asociaciones dentro de cada grupo y no constituyen una prueba de diferencias entre las correlaciones de estudiantes y docentes.

En ambos grupos, la asociación de mayor magnitud correspondió a la asociación entre evaluación auténtica y trazabilidad del aprendizaje (ρ = 0,530 en estudiantes y ρ = 0,516 en docentes), seguida de la asociación entre integridad académica y trazabilidad (ρ = 0,507 y ρ = 0,500, respectivamente). Este patrón indica que las puntuaciones superiores en evaluación auténtica e integridad académica tendieron a acompañarse de puntuaciones superiores en trazabilidad dentro de cada grupo. La coincidencia descriptiva entre ambos grupos no demuestra que sus correlaciones sean estadísticamente equivalentes ni permite establecer relaciones causales entre los dominios.

En conjunto, los datos obtenidos mediante la encuesta presentan puntuaciones medias superiores al punto medio de la escala en los cinco dominios, con el valor más alto en integridad académica y el más bajo en trazabilidad del aprendizaje. Las asociaciones entre evaluación auténtica, integridad académica y trazabilidad del aprendizaje fueron positivas tanto en estudiantes como en docentes. Estos resultados describen relaciones entre las puntuaciones de autoinforme dentro de cada grupo, sin establecer relaciones causales ni demostrar la validez de constructo del instrumento.

DISCUSIÓN

Los resultados muestran una valoración favorable de los cinco dominios analizados: alfabetización en inteligencia artificial generativa, uso académico de herramientas generativas, evaluación auténtica, integridad académica y trazabilidad del aprendizaje (Ashford et al., 2013). La media más alta correspondió a integridad académica, mientras que la más baja se observó en trazabilidad del aprendizaje.

Este patrón indica que los participantes reconocen la importancia de la responsabilidad académica en contextos mediados por inteligencia artificial generativa, pero perciben una menor consolidación de

prácticas orientadas a documentar, justificar y defender el proceso seguido en las tareas universitarias (Morris, 2018).

Este hallazgo se relaciona con los resultados planteados por Zawacki-Richter et al. (2019) en el artículo "Systematic review of research on artificial intelligence applications in higher education – where are the educators?". Los autores identificaron que la inteligencia artificial en educación superior se ha estudiado principalmente desde aplicaciones como evaluación, personalización, tutoría inteligente y servicios académicos; sin embargo, también señalaron la escasa reflexión crítica sobre sus riesgos éticos y pedagógicos. En este sentido, los resultados del presente estudio aportan evidencia desde la perspectiva de docentes y estudiantes, precisamente en un ámbito que la literatura ha considerado insuficientemente desarrollado: la relación entre uso de IA, evaluación e integridad académica (Chan, 2023).

La puntuación alta en alfabetización en IA generativa coincide con la necesidad de formar competencias específicas para comprender, usar y evaluar críticamente estos sistemas. Laupichler et al. (2022), en "Artificial intelligence literacy in higher and adult education: A scoping literature review", sostienen que la alfabetización en IA en educación superior todavía requiere mayor precisión conceptual, formación sistemática y escalas validadas para su medición. Los resultados del presente estudio son consistentes con esa línea, ya que muestran que la alfabetización en IA no puede reducirse al uso instrumental de herramientas generativas, sino que debe incluir verificación de información, reconocimiento de sesgos, comprensión de límites y responsabilidad en la declaración de uso (Bender et al., 2021; Floridi & Chiriatti, 2020).

El dominio de uso académico de herramientas generativas presentó una media favorable, aunque inferior a la alfabetización en IA y a la integridad académica. Esta diferencia es relevante. Sugiere que los participantes reconocen criterios éticos y conceptuales sobre la IA generativa, pero su incorporación práctica en tareas académicas aún no alcanza el mismo nivel de consolidación (Baidoo & Owusu, 2023; Lo, 2023). Este resultado se aproxima a lo señalado por Kasneci et al. (2023) en "ChatGPT for good? On opportunities and challenges of large language models for education", en el que se reconoce que los modelos de lenguaje pueden apoyar la escritura, la retroalimentación, la síntesis de información y la personalización del aprendizaje, pero requieren competencias para comprender sus limitaciones, verificar información y evitar usos acríticos.

La media más alta en integridad académica puede interpretarse como una señal positiva, pero no suficiente. Los participantes parecen valorar la autoría, la citación, la transparencia y la responsabilidad sobre el producto académico (McCabe et al., 2001; Newton & Lang, 2016). No obstante, la integridad académica en entornos de IA generativa no depende solo de la conciencia normativa. También exige condiciones pedagógicas que permitan diferenciar entre apoyo tecnológico legítimo y delegación indebida de procesos cognitivos. En esta línea, Cotton et al. (2023), en "Chatting and cheating: Ensuring academic integrity in the era of ChatGPT", advierten que la disponibilidad de ChatGPT obliga a replantear las estrategias de evaluación, ya que las respuestas institucionales centradas únicamente en detección o prohibición resultan limitadas frente a herramientas capaces de producir textos académicos con apariencia de autoría humana.

El resultado más bajo en trazabilidad del aprendizaje constituye uno de los aportes principales del estudio. Aunque la media se mantuvo por encima del punto medio de la escala, fue el dominio con menor puntuación media. Este resultado refleja una menor puntuación declarada en los ítems relativos a la documentación del proceso: borradores, avances, justificaciones, defensa oral, retroalimentación y reconstrucción de decisiones. Este hallazgo es importante porque la trazabilidad puede funcionar como puente entre evaluación auténtica e integridad académica. Cuando el proceso de aprendizaje queda documentado, se reduce la dependencia de mecanismos externos de vigilancia y se fortalece la evidencia pedagógica sobre cómo se construyó el producto final (Tlili et al., 2023).

La correlación más alta se observó entre evaluación auténtica y trazabilidad del aprendizaje. Este resultado es coherente con los fundamentos de la evaluación auténtica, que prioriza tareas contextualizadas, desempeño aplicado y evidencias de aprendizaje vinculadas con situaciones reales (Wiggins, 1990). La literatura sobre evaluación auténtica sostiene que este enfoque se orienta a medir desempeños significativos y situados, no solo respuestas finales o reproducción de contenidos. Desde esa perspectiva, la asociación encontrada indica que las puntuaciones más altas en evaluación auténtica se relacionan con puntuaciones más altas en trazabilidad del aprendizaje, especialmente cuando incorporan entregas parciales, retroalimentación y defensa argumentada.

La relación positiva entre evaluación auténtica e integridad académica también resulta consistente con estudios sobre diseño evaluativo y prevención de fraude académico. Bretag et al. (2019), en "Contract cheating and assessment design: Exploring the relationship", analizaron la conexión entre el diseño de la evaluación y las condiciones que pueden facilitar o reducir prácticas de deshonestidad académica. Los resultados del presente estudio se alinean con esa perspectiva: las evaluaciones contextualizadas, con criterios explícitos y centradas en el proceso, parecen asociarse con mejores condiciones para sostener la integridad académica. No obstante, esta asociación no implica que la evaluación auténtica elimine por sí sola el riesgo de uso indebido de IA generativa. A partir de esta asociación, puede plantearse, para investigaciones posteriores, la hipótesis de una posible función protectora del diseño evaluativo cuando se combina con alfabetización en IA y trazabilidad del aprendizaje, cuya eficacia no fue examinada en este estudio (Newton P., 2016).

La correlación positiva entre integridad académica y trazabilidad del aprendizaje refuerza esta interpretación. En contextos donde la IA generativa puede producir textos, esquemas, resúmenes o argumentos en pocos segundos, la evidencia del proceso adquiere valor metodológico y ético. No basta con evaluar el producto final. Es necesario observar cómo el estudiante formuló el problema, seleccionó fuentes, contrastó información, tomó decisiones, incorporó retroalimentación y justificó el resultado. Esta orientación desplaza la evaluación desde una lógica centrada en la sospecha hacia una lógica centrada en la demostración del aprendizaje.

A partir de estos hallazgos se proponen orientaciones pedagógicas cuya eficacia requiere evaluación posterior. Primero, las instituciones deben pasar de políticas restrictivas a políticas formativas sobre el uso de IA generativa. La alfabetización en IA debe incluir habilidades técnicas, criterios éticos, verificación de información, reconocimiento de sesgos y responsabilidad autoral. Segundo, el rediseño de la evaluación debe priorizar tareas auténticas, contextualizadas y difíciles de resolver mediante delegación automática. Tercero, la trazabilidad debe incorporarse como criterio evaluativo formal mediante borradores, bitácoras, defensa oral, rúbricas de proceso, reportes de uso de IA y evidencias de retroalimentación.

El análisis de consistencia interna aportó información sobre la interrelación de los ítems en la muestra estudiada. Por ello, los coeficientes obtenidos no demuestran que los cinco dominios constituyan una estructura empíricamente validada ni que el conjunto de los 30 ítems mida una única dimensión. La organización propuesta requiere evaluación adicional de su estructura interna y de su equivalencia de medición entre estudiantes y docentes.

Sin embargo, los resultados deben interpretarse con cautela. El estudio utilizó muestreo no probabilístico por conveniencia, por lo que los hallazgos no pueden generalizarse a toda la educación superior ecuatoriana. Además, los datos proceden de autoinforme, lo cual puede introducir sesgo de deseabilidad social, especialmente en ítems relacionados con integridad académica. La diferencia entre lo que los participantes declaran y lo que efectivamente realizan en actividades evaluativas debe considerarse

en estudios posteriores. También se requiere profundizar en diferencias por rol institucional, área académica, nivel de formación y frecuencia de uso de IA generativa.

Aunque las instrucciones contextualizaron las respuestas según el rol institucional, algunos ítems sobre elaboración de tareas, aprovechamiento de la retroalimentación y defensa del trabajo académico pueden adquirir significados distintos para estudiantes y docentes. No se evaluó la invariancia de medición entre ambos grupos, por lo que las diferencias observadas en sus puntuaciones no pueden interpretarse de manera concluyente como diferencias en los niveles de los constructos. El análisis separado por rol permite examinar las asociaciones dentro de cada grupo, pero no elimina esta limitación de comparabilidad.

Futuras investigaciones deberían incorporar diseños mixtos, análisis comparativos entre facultades o carreras, entrevistas a docentes y estudiantes y revisión documental de tareas evaluativas. También sería pertinente analizar evidencias reales de trazabilidad, como borradores, rúbricas, bitácoras, reportes de uso de IA y registros de retroalimentación. Esta línea permitiría pasar de la percepción declarada a la observación de prácticas evaluativas concretas.

CONCLUSIÓN

Este estudio analizó la relación entre alfabetización en inteligencia artificial generativa, evaluación auténtica e integridad académica en educación superior. Los hallazgos muestran asociaciones positivas entre las puntuaciones de estos dominios en la muestra estudiada. La participación de docentes y estudiantes permitió observar el fenómeno desde dos actores centrales del proceso formativo, con una muestra compuesta por 482 estudiantes y 121 docentes.

La integridad académica obtuvo la media más alta, seguida de la alfabetización en inteligencia artificial generativa. Este resultado refleja las valoraciones declaradas por los participantes respecto de la autoría, la transparencia y la responsabilidad académica ante el uso de herramientas generativas. Sin embargo, la trazabilidad del aprendizaje presentó la media más baja, lo que señala una menor puntuación autoinformada en los aspectos relacionados con la documentación del proceso de elaboración, justificar decisiones, presentar avances y defender el aprendizaje alcanzado.

Las correlaciones positivas y estadísticamente significativas entre las variables muestran asociaciones entre los dominios analizados, pero no confieren por sí mismas validez de constructo al modelo propuesto. De manera particular, la evaluación auténtica se asoció positivamente con la trazabilidad del aprendizaje y la integridad académica, sin que estas relaciones permitan establecer que las tareas situadas y orientadas al proceso produzcan mayor transparencia académica. Estos hallazgos sugieren la pertinencia de explorar respuestas institucionales que incorporen estrategias formativas y evaluativas, además de los mecanismos de vigilancia y detección.

La consistencia interna del instrumento fue adecuada en todos los dominios y alta en el conjunto de los 30 ítems. Este resultado aporta evidencia sobre la consistencia interna de las puntuaciones en la muestra analizada y permite considerar el cuestionario como una herramienta inicial de exploración, cuya estructura de cinco dominios requiere evaluación adicional. Su aplicación futura debería ampliarse a diferentes instituciones, áreas del conocimiento y modalidades de formación, junto con análisis de su estructura interna y de la equivalencia de medición entre estudiantes y docentes.

En términos pedagógicos, los resultados sugieren la pertinencia de considerar la alfabetización en inteligencia artificial generativa en las políticas de evaluación universitaria. Como orientación para su incorporación, se propone establecer criterios claros sobre uso permitido, declaración de herramientas, verificación de información, citación de fuentes y responsabilidad sobre el producto final. Asimismo, se

propone complementar la evaluación auténtica con evidencias de proceso, tales como borradores, bitácoras, rúbricas, defensas orales y reportes de uso de IA, cuya contribución a la integridad académica deberá evaluarse en estudios posteriores.

El estudio aporta evidencia empírica circunscrita a una muestra no probabilística de una institución para sustentar la reflexión sobre el rediseño de la evaluación en educación superior. Las evaluaciones situadas, transparentes y trazables constituyen una línea de trabajo pertinente para explorar la integridad académica en entornos mediados por inteligencia artificial generativa, sin que el presente estudio haya demostrado su eficacia. Futuras investigaciones deberían contrastar estos hallazgos mediante diseños longitudinales, análisis por áreas disciplinares, estudios comparativos entre instituciones y revisión directa de evidencias evaluativas producidas por estudiantes.

REFERENCIAS

Asamblea Nacional del Ecuador. (2008). Constitución de la República del Ecuador. Registro Oficial No. 449.

Asamblea Nacional del Ecuador. (2021). Ley Orgánica de Protección de Datos Personales. Quinto Suplemento del Registro Oficial No. 459.

Ashford-Rowe, K., Herrington, J., & Brown, C. (2014). Establishing the critical elements that determine authentic assessment. Assessment & Evaluation in Higher Education, 39(2), 205–222. https://doi.org/10.1080/02602938.2013.819566

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. Journal of AI, 7(1), 52–62. https://doi.org/10.61969/jai.1337500

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? In Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency (pp. 610–623). Association for Computing Machinery. https://doi.org/10.1145/3442188.3445922

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. Frontiers in Public Health, 6, Article 149. https://doi.org/10.3389/fpubh.2018.00149

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., van Haeringen, K., Saddiqui, S., & Rozenberg, P. (2019). Contract cheating and assessment design: Exploring the relationship. Assessment & Evaluation in Higher Education, 44(5), 676–691. https://doi.org/10.1080/02602938.2018.1527892

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., Rozenberg, P., Saddiqui, S., & van Haeringen, K. (2019). Contract cheating: A survey of Australian university students. Studies in Higher Education, 44(11), 1837–1856. https://doi.org/10.1080/03075079.2018.1462788

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20, Article 38. https://doi.org/10.1186/s41239-023-00408-3

Chen, L., Chen, P., & Lin, Z. (2020). Artificial intelligence in education: A review. IEEE Access, 8, 75264–75278. https://doi.org/10.1109/ACCESS.2020.2988510

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. Innovations in Education and Teaching International, 61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. International Journal of Educational Technology in Higher Education, 20, Article 22. https://doi.org/10.1186/s41239-023-00392-8

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. Minds and Machines, 30(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

Gulikers, J. T. M., Bastiaens, T. J., & Kirschner, P. A. (2004). A five-dimensional framework for authentic assessment. Educational Technology Research and Development, 52(3), 67–86. https://doi.org/10.1007/BF02504676

Herrington, J., & Oliver, R. (2000). An instructional design framework for authentic learning environments. Educational Technology Research and Development, 48(3), 23–48. https://doi.org/10.1007/BF02319856

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

Laupichler, M. C., Aster, A., Schirch, J., & Raupach, T. (2022). Artificial intelligence literacy in higher and adult education: A scoping literature review. Computers and Education: Artificial Intelligence, 3, Article 100101. https://doi.org/10.1016/j.caeai.2022.100101

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. Education Sciences, 13(4), Article 410. https://doi.org/10.3390/educsci13040410

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727

McCabe, D. L., Treviño, L. K., & Butterfield, K. D. (2001). Cheating in academic institutions: A decade of research. Ethics & Behavior, 11(3), 219–232. https://doi.org/10.1207/S15327019EB1103_2

Ministerio de Salud Pública del Ecuador. (2021). Acuerdo Ministerial No. 00038-2021: Reforma al Reglamento para la aprobación, desarrollo, vigilancia y control de investigaciones observacionales y estudios de intervención en seres humanos.

Ministerio de Salud Pública del Ecuador. (2022). Acuerdo Ministerial No. 00005-2022: Reglamento Sustitutivo del Reglamento para la Aprobación y Seguimiento de Comités de Ética de Investigación en Seres Humanos (CEISH) y Comités de Ética Asistenciales para la Salud (CEAS). Quinto Suplemento del Registro Oficial No. 118.

Morris, E. J. (2018). Academic integrity matters: Five considerations for addressing contract cheating. International Journal for Educational Integrity, 14, Article 15. https://doi.org/10.1007/s40979-018-0038-5

Newton, P. (2016). Academic integrity: A quantitative study of confidence and understanding in students at the start of their higher education. Assessment & Evaluation in Higher Education, 41(3), 482–497. https://doi.org/10.1080/02602938.2015.1024199

Newton, P. M., & Lang, C. (2016). Custom essay writers, freelancers, and other paid third parties. In T. Bretag (Ed.), Handbook of academic integrity (pp. 249–271). Springer. https://doi.org/10.1007/978-981-287-098-8_38

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. Computers and Education: Artificial Intelligence, 2, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era: ChatGPT and beyond. Journal of University Teaching and Learning Practice, 20(2). https://doi.org/10.53761/1.20.02.07

Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. Research in Nursing & Health, 29(5), 489–497. https://doi.org/10.1002/nur.20147

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. Smart Learning Environments, 10, Article 15. https://doi.org/10.1186/s40561-023-00237-x

UNESCO. (2025, 29 de septiembre). Qué debe saber sobre la IA y el derecho a la educación. https://www.unesco.org/es/articles/que-debe-saber-sobre-la-ia-y-el-derecho-la-educacion

Villarroel, V., Bloxham, S., Bruna, D., Bruna, C., & Herrera-Seda, C. (2018). Authentic assessment: Creating a blueprint for course design. Assessment & Evaluation in Higher Education, 43(5), 840–854. https://doi.org/10.1080/02602938.2017.1412396

Wiggins, G. (1990). The case for authentic assessment. Practical Assessment, Research, and Evaluation, 2, Article 2. https://doi.org/10.7275/ffb1-mm19

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, Article 39. https://doi.org/10.1186/s41239-019-0171-0

Submetido: 20/08/2026
Preprint: 10/08/2026
Aprobado: 19/09/2026
Editor de sección: Suzana dos Santos Gomes

DECLARACIÓN DE DISPONIBILIDAD DE LOS DATOS DE INVESTIGACIÓN

Los contenidos subyacentes a la investigación se encuentran disponibles.

DECLARACIÓN DE CONTRIBUCIÓN DE AUTORÍA

Autor 1 - Conceptualización, Curación de Datos, Redacción: revisión y edición.
Autora 2 - Investigación, Metodología, Administración del Proyecto.
Autora 3 - Software, Validación, Redacción: borrador original.

DECLARACIÓN DE CONFLICTO DE INTERÉS

Los autores declaran que no hay conflicto de interés con el presente artículo.

DECLARACIÓN SOBRE EL USO DE INTELIGENCIA ARTIFICIAL

Se utilizó ChatGPT (OpenAI) exclusivamente como apoyo para mejorar la redacción académica y organizar las referencias bibliográficas. La herramienta no fue utilizada en la definición del problema de investigación, el diseño metodológico, la elaboración del instrumento, la recolección o análisis de los datos, la interpretación de los resultados ni la formulación de las conclusiones. Los datos analizados proceden de las respuestas obtenidas mediante el cuestionario aplicado a los 603 participantes y no fueron generados ni simulados mediante inteligencia artificial. Todo contenido apoyado mediante esta herramienta fue revisado críticamente por los autores, quienes asumen plena responsabilidad por la exactitud, originalidad y contenido final del manuscrito.