

Article

Academic Integrity and Generative Artificial Intelligence in Higher Education: Psychometric Validation of a Short Questionnaire

Luzmila Lourdes Garro-Aburto ¹, Jorge Miguel Chávez-Díaz ^{2,*}, Carlos Alberto Hinojosa Salazar ³,
Edith María Llerena Espinoza ⁴ and Sara Pamela Sánchez Sandoval ¹

¹ Postgraduate School, Universidad César Vallejo, Lima 15487, Peru; luzgarro@ucvvirtual.edu.pe (L.L.G.-A.); pasanchez@ucvvirtual.edu.pe (S.P.S.S.)

² Faculty of Law and Economics, Universidad Científica del Sur, Lima 15836, Peru

³ Faculty of Economic and Administrative Sciences, Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas, Chachapoyas 01001, Peru; carlos.hinojosa@untram.edu.pe

⁴ Department of Chemistry, Universidad Nacional de Educación Enrique Guzmán y Valle, Lima 15472, Peru; elllerena@une.edu.pe

* Correspondence: jchavezd@cientifica.edu.pe

Abstract

Generative artificial intelligence has reshaped writing, learning, and assessment practices in higher education, posing new challenges for academic integrity. In this context, the objective of this study was to analyze the psychometric properties of an academic integrity questionnaire among college students. An instrumental study with a quantitative approach and cross-sectional design was conducted with a sample of 419 higher education students. The original instrument included 17 items distributed across three dimensions: ethical use of AI, awareness of risks of misuse, and support for academic writing. Descriptive analyses, reliability estimates, exploratory factor analysis, and confirmatory factor analysis were performed using a robust approach based on DWLS, considering the ordinal nature of the items. The results supported a refined version of 15 items organized into two correlated factors: normative integrity in AI use and support for academic writing. The final model showed adequate fit, high internal consistency, convergent validity, acceptable but borderline discriminant validity, and preliminary evidence of measurement invariance across gender, although the latter should be interpreted with caution because the configural model did not show optimal absolute fit.

Keywords: academic integrity; generative artificial intelligence; higher education; psychometric validation; academic writing



Academic Editor: Michael DeBraga

Received: 31 May 2026

Revised: 10 July 2026

Accepted: 12 July 2026

Published: 15 July 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Generative artificial intelligence (GAI) has substantially transformed teaching, learning, writing, and assessment practices in higher education. Tools based on large language models, such as ChatGPT and similar systems, enable the generation of text, the reorganization of ideas, the synthesis of information, and the provision of immediate feedback. These capabilities have been valued for their potential to support learning and expand access to academic resources; however, they have also heightened concerns about misinformation, bias, algorithmic opacity, technological dependence, the replacement of intellectual work, and the undermining of academic integrity (Dwivedi et al., 2023; Lim et al., 2023; Tlili et al., 2023).

Consequently, the debate over AI in higher education is no longer limited to its technical feasibility but has shifted toward its pedagogical, ethical, and institutional implications. Recent evidence suggests that these systems do not function merely as instrumental tools but as socio-technical technologies that reshape the relationship between students, faculty, knowledge, and assessment (Bearman et al., 2023; Bearman & Ajjawi, 2023; Chiu, 2024). In this context, universities face the challenge of integrating AI in an educational and regulated manner, avoiding approaches that rely solely on prohibition, surveillance, or automated detection (López Gómez et al., 2024).

One of the areas where these tensions are most prominent is academic integrity. Traditionally, this concept has been associated with honesty, responsibility, authorship, transparency, and the rejection of practices such as plagiarism, collusion, or fraud. However, GAI challenges the conventional boundaries of these categories by intervening in different phases of academic work, such as idea generation, planning, writing, revision, and editing. Consequently, distinguishing between legitimate support and undue substitution of intellectual effort becomes more complex, especially when students lack clear criteria for declaring, delimiting, or justifying the use of generative tools (Farrokhnia et al., 2024; Nam & Bai, 2023; Yusuf et al., 2024).

Recent evidence suggests that academic integrity in AI contexts should not be reduced to the possibility of “cheating,” but rather understood as a broader issue of academic responsibility. In higher education, the use of ChatGPT and other generative tools poses risks that go beyond plagiarism or the unauthorized production of texts, including uncritical acceptance of automated responses, difficulties in verifying sources, loss of writing agency, and ambiguity regarding authorship (Essel et al., 2024; Lee et al., 2024; van Niekerk et al., 2025). At the same time, these technologies can support academic learning and writing through argumentative organization, linguistic review, and feedback, provided their use is embedded within clear frameworks of transparency, pedagogical guidance, and process-centered assessment.

Consequently, assessing academic integrity in contexts mediated by generative AI requires distinguishing between the educational use of these tools and their use as a substitute for authorship (Banihashem et al., 2024; Barrett & Pack, 2023). This distinction is particularly relevant in a context still marked by inconsistent institutional responses: while some universities have opted for restrictive policies, others have moved toward strategies of gradual integration through guidelines, resources, and curriculum guidance, though without a clear consensus on how to fairly regulate their use in higher education, particularly given concerns about privacy, accuracy, equity, intellectual property, assessment, and academic integrity (Jin et al., 2025; Wang et al., 2024).

From the students’ perspective, the adoption of ChatGPT and similar tools is often associated with perceived usefulness, accessibility, motivation, and support for learning (Habibi et al., 2023; Strzelecki, 2024). However, these attitudes are not uniform, as they vary depending on cultural context, prior experience, and perceived benefits, underscoring the need for culturally relevant tools to empirically measure practices and attitudes related to academic integrity in university settings mediated by artificial intelligence (Abdaljaleel et al., 2024).

However, the growing body of literature on the acceptance, use, and regulation of academic integrity in higher education has not been matched by an equal volume of psychometric research aimed at assessing academic integrity in these new settings. Much of the available research has focused on perceptions, attitudes, opportunities, risks, or institutional responses, while the availability of brief, valid, and reliable instruments that allow for examining how students understand and practice academic integrity when interacting with generative systems remains limited. This gap is significant because higher education

institutions require empirical evidence to design policies, educational interventions, and support strategies that do not rely solely on normative assumptions or faculty perceptions.

Furthermore, measuring academic integrity in contexts mediated by AI requires an updated conceptualization of the construct. In this context, traditional academic integrity tools require conceptual and metric adaptations that can capture emerging practices, such as automated writing assistance, awareness of the risks of misuse, and ethical decision-making in the face of intelligent systems (Chan, 2025; Torres-Diaz et al., 2025).

In this vein, the present study aimed to analyze the psychometric properties of an academic integrity questionnaire in contexts mediated by artificial intelligence, as applied to college students. Specifically, the study examined the instrument's factor structure, reliability, convergent validity, discriminant validity, and measurement invariance. In doing so, the study sought to provide a brief, empirically grounded tool, as well as evidence regarding the latent structure of the construct, which can serve as a basis for future comparative research and interventions aimed at promoting the ethical, critical, and transparent use of AI in academic life.

2. Materials and Methods

2.1. Study Design

A quantitative, cross-sectional study was conducted to analyze the psychometric properties of an academic integrity questionnaire in artificial intelligence-mediated contexts among college students. To this end, the study evaluated the instrument's factor structure, internal consistency, convergent validity, discriminant validity, and measurement invariance.

2.2. Participants

The target population comprised students enrolled in higher education programs. A non-probabilistic purposive sampling strategy was used because the objective of the study was to obtain initial psychometric evidence for the questionnaire rather than to estimate population parameters. The inclusion criteria were that participants be enrolled in a higher education program, provide informed consent, and complete the questionnaire in its entirety. Incomplete records or those with inconsistent response patterns were excluded.

The final analytic sample consisted of 419 volunteer students. To contextualize the educational scope of the validation, the sample was composed of students from public universities (73.75%) and private universities (26.25%). Most participants were enrolled in undergraduate programs (81.62%), whereas 18.38% were graduate students. Regarding mode of study, 85.92% were enrolled in on-campus programs, 10.02% in hybrid programs, and 4.06% in online programs. By gender, 54.65% identified as female and 45.35% as male. The predominant age group was 18 to 21 years old (49.40%).

For the factor analyses, the total sample was randomly divided into two independent subsamples. The first subsample ($n = 216$; 51.55%) was used for exploratory factor analysis (EFA), while the second subsample ($n = 203$; 48.45%) was used for confirmatory factor analysis (CFA). This procedure separated the exploratory and confirmatory phases, reduced the risk of capitalizing on chance, and strengthened the evaluation of the questionnaire's internal structure.

2.3. Instrument

An academic integrity scale for AI-mediated contexts, adapted from a previously validated instrument, was used to assess the ethical use of AI, awareness of the risks of misuse, and support for academic writing in higher education. The original version consisted of 17 items distributed across three dimensions: ethical use of AI, awareness of

the risks of misuse, and support for academic writing. Responses were recorded using a five-point Likert scale, ranging from 1 = strongly disagree to 5 = strongly agree. The questionnaire is included in Appendix A. The original instrument reported initial evidence of reliability and validity in a graduate student population (Maharmah et al., 2025).

2.4. Linguistic and Contextual Adaptation

The instrument underwent a process of linguistic and contextual adaptation to the Spanish-speaking university setting. First, the translated items were reviewed by two specialists with experience in academic writing, research methodology, and the educational use of digital technologies. This review focused on semantic equivalence, conceptual clarity, and contextual relevance. Subsequently, a pilot test was conducted with 20 students to evaluate the clarity, comprehensibility, and relevance of the statements. Based on this process, minor editorial adjustments were made without altering the theoretical content of the items. A formal back-translation procedure was not conducted because the aim of the present study was not to establish strict equivalence between parallel language versions, but rather to validate a culturally and contextually appropriate Spanish version for university students. This procedure was considered suitable for the present validation because semantic equivalence was reviewed by specialists, the items were evaluated with members of the target population, and the final version was subsequently examined through exploratory and confirmatory psychometric analyses. Nevertheless, future studies should assess the cross-linguistic equivalence of the instrument through formal translation, back-translation, and invariance analyses in English-speaking or bilingual samples.

2.5. Data Collection Procedure

Data collection was conducted using an online questionnaire. Before responding, participants received information about the study's objectives, the voluntary nature of their participation, data confidentiality, and the exclusive academic use of the information. Informed consent was provided online prior to completing the questionnaire. Only complete records with valid response patterns were included in the final dataset. The study was conducted in accordance with the ethical principles applicable to research involving human subjects. Additionally, the Research Ethics Committee of César Vallejo University issued Report No. 00156-2025/CEI-PDE, in which it determined that formal ethical review was waived due to the nature and methodological characteristics of the study.

2.6. Data Analysis

The psychometric analysis was conducted in sequential stages. First, descriptive statistics were examined for all items, including mean, standard deviation, median, range, skewness, and kurtosis. Preliminary internal consistency was then estimated using Cronbach's alpha and McDonald's omega.

The total sample was randomly divided into two independent subsamples. The first subsample was used for EFA and the second for CFA. In the EFA, the suitability of the matrix was assessed using the Kaiser–Meyer–Olkin index and Bartlett's sphericity test. Given the ordinal nature of the items, factor extraction was conducted using the minimum residual method with oblique rotation. The number of factors was evaluated using complementary criteria, including eigenvalues, inspection of the scree plot, parallel analysis, factor loadings, theoretical interpretability, and empirical definition of the factors.

CFA was subsequently conducted in SEMLj in jamovi v.2.6.44 using a robust approach based on DWLS, consistent with the ordinal nature of the items. Alternative models were compared using χ^2 , CFI, TLI, RMSEA with 95% confidence intervals, and SRMR. Reliability, convergent validity, discriminant validity, and measurement invariance across gender

were also examined. Residual covariances were considered only as restricted, theoretically justified, and partially data-driven model refinements.

3. Results

3.1. Descriptive Analysis of the Items

The 17 items of the academic integrity questionnaire in AI-mediated contexts were analyzed. No missing data were observed, so the analytical sample consisted of 419 valid responses. As shown in Table 1, the means ranged from 3.68 to 4.05, while the standard deviations ranged from 0.94 to 1.11. Furthermore, all items utilized the full response range of the Likert scale (1–5), and the median was 4 in all cases. The skewness coefficients were negative for all items, suggesting a tendency toward favorable responses, while the kurtosis values did not show severe deviations. Taken together, these results indicate an adequate distribution of responses and sufficient variability for subsequent analyses.

Table 1. Descriptive statistics for the 17 initial items.

Item	N	Mean	SD	Median	Minimum	Max	Skewness	Kurtosis
EUA1	419	4.05	0.98	4.00	1.00	5.00	−1.17	1.26
EUA2	419	3.83	1.03	4.00	1.00	5.00	−0.88	0.45
EUA3	419	3.87	0.98	4.00	1.00	5.00	−0.81	0.53
EUA4	419	3.99	0.94	4.00	1.00	5.00	−0.91	0.59
EUA5	419	3.77	1.07	4.00	1.00	5.00	−0.58	−0.24
EUA6	419	4.00	1.05	4.00	1.00	5.00	−1.12	0.93
ARM1	419	3.97	1.06	4.00	1.00	5.00	−1.07	0.74
ARM2	419	3.99	0.96	4.00	1.00	5.00	−0.94	0.67
ARM3	419	3.68	1.11	4.00	1.00	5.00	−0.59	−0.44
ARM4	419	3.97	1.03	4.00	1.00	5.00	−0.96	0.45
ARM5	419	4.02	1.00	4.00	1.00	5.00	−0.97	0.46
ARM6	419	3.87	0.96	4.00	1.00	5.00	−0.71	0.25
SAW1	419	3.87	1.06	4.00	1.00	5.00	−0.81	0.10
SAW2	419	3.93	0.97	4.00	1.00	5.00	−0.82	0.31
SAW3	419	3.89	1.00	4.00	1.00	5.00	−0.82	0.29
SAW4	419	3.85	1.01	4.00	1.00	5.00	−0.86	0.38
SAW5	419	3.82	1.04	4.00	1.00	5.00	−0.78	0.23

3.2. Preliminary Internal Consistency

The full 17-item scale showed high internal consistency, with a Cronbach's alpha of 0.959 and a McDonald's omega of 0.960. When analyzing the original dimensions separately, adequate coefficients were also observed: ethical use of AI ($\alpha = 0.901$, $\omega = 0.904$), awareness of risks of misuse ($\alpha = 0.903$, $\omega = 0.907$), and support for academic writing ($\alpha = 0.921$, $\omega = 0.922$). Taken together, these results demonstrate high internal homogeneity of the items, although the high coefficients warranted a more detailed examination of the instrument's internal structure.

3.3. Exploratory Factor Analysis

Exploratory factor analysis (EFA) was conducted on the first random subsample ($n = 216$) to examine the initial latent structure of the questionnaire's 17 items. Prior to this,

the suitability of the correlation matrix was assessed. The overall KMO index was excellent ($KMO = 0.954$), and Bartlett's sphericity test was statistically significant, $\chi^2(136) = 3020$, $p < 0.001$, confirming the appropriateness of factor analysis. Factor retention was evaluated using multiple complementary criteria, including eigenvalues, inspection of the scree plot, parallel analysis, the magnitude and pattern of factor loadings, theoretical interpretability, and empirical definition of the factors. The EFA results are presented in Table 2. The solution showed that items related to ethical AI use and awareness of risks of misuse tended to cluster mainly in Factor 1, whereas items related to support for academic writing loaded on Factor 2. Although a third factor emerged, it showed weak empirical definition and was mainly associated with EUA3 and, to a lesser extent, EUA2. The factor correlation matrix also indicated a substantial association between the first two factors. Overall, these findings did not provide robust support for the original three-dimensional structure and suggested the need to test alternative confirmatory models.

Table 2. Exploratory factor analysis of the initial 17-item version.

Item	Factor 1	Factor 2	Factor 3	Uniqueness
EUA1	0.465	—	—	0.524
EUA2	0.459	—	0.445	0.375
EUA3	—	—	0.619	0.243
EUA4	0.589	—	—	0.273
EUA5	0.499	—	—	0.567
EUA6	0.802	—	—	0.244
ARM1	0.726	—	—	0.349
ARM2	0.79	—	—	0.241
ARM3	0.559	—	—	0.503
ARM4	0.819	—	—	0.335
ARM5	0.892	—	—	0.254
ARM6	0.733	—	—	0.343
SAW1	—	0.673	—	0.312
SAW2	—	0.658	—	0.209
SAW3	—	0.767	—	0.258
SAW4	—	0.954	—	0.256
SAW5	—	0.697	—	0.438

Note. Extraction method: minimum residual; rotation method: oblimin. Loadings below 0.30 were suppressed. Factor 1 mainly grouped items related to ethical AI use and awareness of misuse risks; Factor 2 grouped items related to support for academic writing; Factor 3 showed weak empirical definition, mainly associated with EUA3 and EUA2.

3.4. Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) was conducted on the second random subsample ($n = 203$) to test alternative measurement models and select the solution that best balanced statistical fit, parsimony, and theoretical interpretability. As shown in Table 3, four models were compared: a unidimensional model, the original three-factor theoretical model, a two-factor model without residual covariances, and an adjusted version of the latter with three theoretically justified residual covariances between items of the same factor. The comparison was conducted by considering overall fit, parsimony, conceptual interpretability, and discriminant validity between factors.

Table 3. Comparison of confirmatory models for the Academic Integrity Questionnaire.

Model	Structure Evaluated	Items	χ^2	df	χ^2/df	CFI	TLI	RMSEA	SRMR	HTMT	Decision
Model 1	One-dimensional	17	632	119	5.31	0.990	0.988	0.146	0.077	—	Rejected
Model 2	Three original factors	17	337	116	2.91	0.996	0.995	0.097	0.056	—	Rejected
Model 3	Two factors	15	267	89	3.00	0.996	0.995	0.099	0.052	0.842	Acceptable
Model 4	Two factors with residual covariances	15	161	86	1.87	0.998	0.998	0.066	0.043	0.842	Retained

Note. The three-factor model corresponded to ethical use of AI, awareness of the risks of misuse, and support for academic writing. The fitted model included the residual covariances EUA2 ~ EUA3, SAW3 ~ SAW4, and EUA6 ~ ARM1. The HTMT index was reported only in the two-factor models, as its assessment was relevant for examining the discriminant validity between conceptually similar dimensions. The final model retained was Model 4, as it presented the best balance between overall fit, parsimony, and theoretical interpretability.

As shown in Table 3, the one-factor model exhibited the poorest fit, primarily due to its high RMSEA value. The original three-factor model showed improved fit indices; however, the exploratory results and the conceptual overlap between ethical AI use and awareness of risks of misuse suggested that these two dimensions were not clearly differentiated in the present sample. Therefore, a more parsimonious two-factor solution was tested and compared with the original theoretical model. The two-factor model offered a conceptually interpretable solution, although its RMSEA remained above the recommended criterion. Finally, an adjusted two-factor model was estimated with three residual covariances between item pairs from the same substantive factor. This specification showed the best balance between overall fit, parsimony, and theoretical interpretability, with CFI = 0.998, TLI = 0.998, RMSEA = 0.066, and SRMR = 0.043. Therefore, it was retained as the final model for the questionnaire.

It is important to note that the retained model represents an empirically refined solution rather than a purely confirmatory replication of the original structure. The inclusion of residual covariances was treated as a restricted and partially data-driven modeling decision. These covariances were considered only after comparing competing models, and only when they involved items within the same factor and reflected clear semantic proximity. Correlated residuals were preferred over alternative specifications such as cross-loadings, item reassignment, or additional factors because the associated item pairs appeared to reflect local wording overlap and conceptual proximity rather than a distinct latent construct. Thus, the final model should be interpreted as a theoretically plausible and parsimonious solution for the present sample, while recognizing the need to replicate its stability in independent samples.

Based on the exploratory and confirmatory results, a two-factor model comprising 15 items was retained. The first factor, normative integrity in the use of AI, included items related to the ethical use of technology, academic responsibility, and awareness of the risks of misuse. The second factor, support for academic writing, grouped items referring to the use of AI as a resource to support the drafting, organization, and improvement of academic texts.

As shown in Table 4, EUA5 and ARM3 presented comparatively weaker psychometric performance than most retained items, with moderate factor loadings and relatively high uniqueness values. Therefore, these items were excluded from the refined version to improve parsimony and strengthen the interpretability of the final factor solution. The final model was estimated using a robust approach based on DWLS, consistent with the ordinal nature of the items, and incorporated three residual covariances between pairs of items from the same factor (EUA2 ~ EUA3, SAW3 ~ SAW4, and EUA6 ~ ARM1).

Table 4. Evidence supporting the exclusion of items EUA5 and ARM3.

Item	Main EFA Loading	Secondary Loading	Uniqueness	Item–Other Correlation in Original Dimension	Decision
EUA5	0.499	—	0.567	0.594	Removed
ARM3	0.559	—	0.503	0.627	Removed

Note. EUA5 and ARM3 showed comparatively weaker psychometric contribution than most retained items, as reflected in their moderate factor loadings and relatively high uniqueness values. Their exclusion contributed to a more parsimonious 15-item solution while preserving the theoretical coverage of the construct.

The model showed an adequate overall fit: $\chi^2(86) = 161$, $p < 0.001$; CFI = 0.998; TLI = 0.998; SRMR = 0.043; RMSEA = 0.066, 95% CI [0.050, 0.081], $p = 0.052$. Although the χ^2 test was significant, the incremental and residual fit indices supported a satisfactory solution. Table 5 presents the standardized factor loadings for the final model.

Table 5. Standardized factor loadings and confidence intervals for the final model.

Factor	Item	β	95% CI for β
Normative Integrity in AI Use	EUA1	0.900	[0.866, 0.933]
	EUA2	0.782	[0.732, 0.832]
	EUA3	0.782	[0.730, 0.833]
	EUA4	0.876	[0.842, 0.911]
	EUA6	0.855	[0.813, 0.897]
	ARM1	0.868	[0.830, 0.905]
	ARM2	0.847	[0.805, 0.888]
	ARM4	0.747	[0.686, 0.808]
	ARM5	0.892	[0.857, 0.927]
	ARM6	0.827	[0.781, 0.873]
Support for Academic Writing	SAW1	0.861	[0.818, 0.904]
	SAW2	0.927	[0.900, 0.954]
	SAW3	0.871	[0.830, 0.911]
	SAW4	0.867	[0.826, 0.908]
	SAW5	0.842	[0.795, 0.889]

As shown in Table 5, all standardized factor loadings were greater than 0.70. The lowest loading was for ARM4 ($\beta = 0.747$), while the highest was for SAW2 ($\beta = 0.927$). Taken together, these results support an adequate representation of the items in their respective latent factors.

3.5. Included Residual Covariances

Table 6 presents the three residual covariances included in the final model, all of which were statistically significant and conceptually justifiable. EUA2 and EUA3 showed a standardized residual covariance of 0.479; SAW3 and SAW4, of 0.473; and EUA6 and ARM1, of 0.551. Since each pair of items belonged to the same factor and shared content proximity, these associations were interpreted as local dependencies attributable to semantic similarity, rather than as evidence of an additional dimension.

Table 6. Residual covariances included in the final model.

Residual Relationship	Estimate	SE	β	95% CI for β
EUA2 ~ EUA3	0.186	0.025	0.479	[0.395, 0.564]
SAW3 ~ SAW4	0.116	0.029	0.473	[0.334, 0.611]
EUA6 ~ ARM1	0.142	0.023	0.551	[0.475, 0.628]

These residual covariances were retained because they were theoretically interpretable and involved item pairs with close semantic overlap within the same latent dimension. Specifically, EUA2 and EUA3 both refer to reflective and norm-oriented regulation of AI use; SAW3 and SAW4 capture closely related forms of AI-assisted writing support during idea formulation and expression; and EUA6 and ARM1 refer to responsible academic use of AI through acknowledgement and verification of AI-generated content. The decision to correlate these residuals was preferred over specifying additional factors or allowing cross-loadings because the associations appeared to reflect local dependence derived from item wording and conceptual proximity, rather than evidence of a substantively distinct dimension. However, given the moderate-to-high magnitude of these residual associations, future users of the instrument should not assume that these covariances will be universally stable. If these local dependencies are not replicated in new samples, researchers should report the fit of the unmodified two-factor model, examine item wording, and consider whether specific items require refinement or removal before retaining correlated residuals.

3.6. Reliability, Convergent Validity, and Discriminant Validity

Table 7 presents the reliability of the final model was adequate for both factors. The normative integrity dimension in the use of AI yielded $\alpha = 0.949$, ordinal $\alpha = 0.959$, and $\omega = 0.942$, while the academic writing support dimension yielded $\alpha = 0.926$, ordinal $\alpha = 0.945$, and $\omega = 0.908$. Taken together, these coefficients indicate high internal consistency.

Table 7. Reliability and convergent validity of the final model.

Factor	α	α Ordinal	ω	AVE
Normative Integrity in AI Use	0.949	0.959	0.942	0.704
Support for Academic Writing	0.926	0.945	0.908	0.764

Note: Discriminant validity was assessed in the text using the HTMT index between the two factors (HTMT = 0.842).

Convergent validity was supported by the average variance extracted (AVE) values. The normative integrity dimension in AI use yielded an AVE of 0.704, and the academic writing support dimension yielded an AVE of 0.764. In both cases, the values exceeded the recommended threshold of 0.50, indicating that the items explain a sufficient proportion of the variance in their respective factors.

Discriminant validity was assessed using the HTMT index. The value observed between the two factors was 0.842, which was below the strict cutoff of 0.85 and the flexible criterion of 0.90. However, the latent correlation between the factors was high, $r = 0.870$, 95% CI [0.835, 0.904], indicating substantial conceptual overlap. Thus, discriminant validity was acceptable but close to the strict threshold. These results suggest that the two dimensions are empirically distinguishable but strongly related. Accordingly, normative integrity in AI use and support for academic writing should not be interpreted as entirely independent constructs, but rather as closely connected facets of responsible AI use in academic contexts.

3.7. Measurement Invariance

To assess the equivalence of the final model across groups, a measurement invariance analysis was conducted across gender groups. Configural, metric, scalar, and strict models were estimated sequentially using WLSMV, taking into account the ordinal nature of the items.

As shown in Table 8, the configural model exhibited adequate incremental and residual indices, with CFI = 0.976, TLI = 0.971, and SRMR = 0.041. However, the RMSEA value was high, RMSEA = 0.114, indicating that the baseline model did not show optimal absolute fit. Therefore, the invariance results should be interpreted cautiously. The progressive imposition of restrictions did not result in substantial deterioration in fit. Specifically, metric invariance yielded $\Delta\text{CFI} = 0.005$, $\Delta\text{RMSEA} = -0.016$, and $\Delta\text{SRMR} = 0.005$; scalar invariance yielded $\Delta\text{CFI} = 0.000$, $\Delta\text{RMSEA} = -0.010$, and $\Delta\text{SRMR} = -0.005$; and strict invariance yielded $\Delta\text{CFI} = 0.001$, $\Delta\text{RMSEA} = -0.004$, and $\Delta\text{SRMR} = 0.006$. Taken together, these results provide preliminary and supportive evidence of measurement invariance across gender. Nevertheless, because the configural model did not show optimal absolute fit, these findings should not be interpreted as definitive evidence of full gender comparability. Further replication in larger, independent, and more balanced samples is recommended.

Table 8. Measurement invariance of the final model by gender.

Model	χ^2	df	CFI	TLI	RMSEA	SRMR	ΔCFI	ΔRMSEA	ΔSRMR	Decision
Configural	640.00	172	0.976	0.971	0.114	0.041	—	—	—	Acceptable with reservations
Metric	554.60	185	0.981	0.979	0.098	0.046	0.005	−0.016	0.005	Accepted
Scalar	591.90	228	0.982	0.983	0.087	0.041	0.000	−0.010	−0.005	Accepted
Strict	594.20	243	0.982	0.985	0.083	0.046	0.001	−0.004	0.006	Accepted

4. Discussion

The objective of the study was to analyze the psychometric properties of an academic integrity questionnaire in artificial intelligence-mediated contexts among college students. Overall, the results support a refined 15-item version organized into two correlated factors: normative integrity in the use of AI and support for academic writing. This solution showed favorable indicators of fit, internal consistency, and convergent validity, suggesting that the instrument constitutes a brief and empirically supported alternative for assessing academic integrity in university settings characterized by the growing use of generative tools. Nevertheless, the findings should be interpreted with appropriate caution because the final CFA model included three residual covariances, discriminant validity was acceptable but close to the strict HTMT threshold, and the evidence of measurement invariance across gender was preliminary due to the elevated RMSEA of the configural model.

A key initial finding is that the empirical structure did not fully replicate the original three-dimensional framework. Items related to the ethical use of AI and awareness of the risks of misuse tended to behave as expressions of a common dimension. This convergence is theoretically meaningful because, in the context of generative AI, ethical academic behavior is not limited to compliance with institutional rules. Rather, it also requires recognizing the risks associated with misuse, anticipating their consequences, verifying AI-generated content, acknowledging technological assistance, and regulating one's own interaction with these systems. From this perspective, ethical use and awareness of misuse risks may operate as complementary components of a broader normative orientation toward responsible AI use.

This interpretation is consistent with recent frameworks for AI literacy, which argue that ethical engagement with artificial intelligence is closely linked to critical evaluation, autonomy, and an understanding of the social and educational implications of intelligent systems (Allen & Kendeou, 2024; Ng et al., 2024; Visedsun et al., 2026). It also aligns with current debates on academic integrity in the age of generative artificial intelligence, which emphasize that integrity should not be reduced to the avoidance of plagiarism, but should be understood as a situated academic practice involving transparency, critical judgment, responsible decision-making, and the preservation of student authorship (Chan, 2025; Cotton et al., 2024; Huang et al., 2025; Torres-Diaz et al., 2025). Therefore, the merger of ethical AI use and awareness of misuse risks into a single factor may reflect the way students cognitively organize these issues in practice: not as separate domains, but as interdependent aspects of responsible participation in AI-mediated academic work.

The second factor, support for academic writing, remained empirically distinct and exhibited high factor loadings. This result suggests that students distinguish between the normative dimension of AI use and its use as a support resource for planning, revising, organizing, or improving academic texts. This distinction is particularly relevant because it avoids a reductionist interpretation according to which any use of AI in writing constitutes a threat to academic integrity. Rather, the findings suggest that the use of these tools can be understood as a legitimate practice when it is oriented toward educational support rather than the replacement of authorship. This interpretation aligns with recent reviews on AI-assisted writing, which suggest that generative tools can facilitate revision processes, discursive organization, and metacognitive reflection when used critically and with transparency criteria (Arseven & Bal, 2025; Lancaster, 2023). Similarly, studies on student perceptions show that college students recognize the benefits of generative AI in supporting learning, writing, and feedback, although they simultaneously express concerns about accuracy, ethics, privacy, and independent academic development (Chan & Hu, 2023). Consequently, the questionnaire makes it possible to distinguish between regulatory integrity practices and instrumental uses of AI related to writing—a distinction that is particularly useful for institutions seeking to regulate these technologies without hindering their educational potential.

At the same time, the high latent correlation between the two factors suggests that normative integrity in AI use and support for academic writing should be understood as closely related facets of a broader construct: responsible AI use in academic contexts. Students may distinguish between using AI as a writing support tool and using it in accordance with integrity principles, but these practices are not independent. For instance, the legitimate use of AI to improve grammar, organize ideas, or revise academic texts still requires transparency, verification, authorship awareness, and alignment with institutional expectations. Consequently, the two-factor structure should not be interpreted as evidence of fully separate constructs, but as a psychometrically useful differentiation within a broader responsible-use framework.

Compared with previous instrument validation studies, the results obtained show favorable psychometric indicators. The original scale reported a structure of 17 items and three dimensions among graduate students (Maharmah et al., 2025). However, in the university sample analyzed in this study, the three-factor solution was neither the most parsimonious nor the most stable. This difference may be associated with contextual, cultural, academic, and sampling variations, as well as with the rapid transformation of student practices related to generative AI. In particular, the predominance of undergraduate students and on-campus programs may have influenced the way participants interpreted academic integrity, risk awareness, and AI-assisted writing practices. Therefore, local validation should not be understood as a mechanical replication of a previous structure,

but rather as a process of conceptual and metric adjustment that allows examination of how the construct is organized in specific educational populations.

From a psychometric perspective, the inclusion of three residual covariances in the final model should be interpreted with caution. Although these associations improved model fit and were theoretically justified by the semantic proximity between item pairs, their moderate-to-high magnitude suggests the presence of local dependence among specific items. This does not necessarily compromise the two-factor structure, since the covariances occurred within the same substantive dimension and did not require cross-loadings or additional factors. However, it does indicate that some items may share residual variance beyond the variance explained by the latent construct. Therefore, the final model should be understood as a restricted and partially data-driven refinement rather than as definitive evidence that these residual associations will appear in all populations. Future studies should test whether these covariances remain stable in independent samples and should examine whether item wording can be refined to reduce local dependence without weakening the theoretical coverage of the scale.

An additional finding of the study was the preliminary evidence of measurement invariance across gender. The configural, metric, scalar, and strict models showed acceptable incremental and residual fit indices, and the changes in CFI, RMSEA, and SRMR across nested models did not indicate substantial deterioration in model fit. Nevertheless, the elevated RMSEA in the configural model indicates that the baseline solution did not achieve optimal absolute fit. Since configural invariance constitutes the foundation for subsequent levels of equivalence, the results should be understood as initial support for gender comparability rather than definitive evidence of full invariance. Consequently, any practical use of the questionnaire to compare gender groups should be made cautiously until these findings are replicated in larger, independent, and more balanced samples.

From an applied perspective, the instrument can support diagnostic processes, training, and institutional decision-making. Universities need brief, psychometrically sound tools that can help identify how students understand academic integrity in AI-mediated writing, assessment, and knowledge-production scenarios. Within this framework, the questionnaire could be used to assess training needs, design interventions on the responsible use of AI, and monitor changes following digital literacy programs. However, comparisons across gender or other academic groups should be made cautiously until the invariance evidence is replicated in larger and more balanced samples. The instrument may also inform curricular development, academic integrity policies, and AI literacy initiatives, particularly in contexts where institutional responses to generative AI cannot be limited to surveillance or automated detection, but must incorporate ethical guidelines, authentic assessment, and transparent policies regarding the permitted use of these tools (Foltynek et al., 2023; Moorhouse et al., 2023; Perkins, 2026). In this regard, the validated questionnaire can help assess a complementary dimension of AI literacy in higher education: the responsible, critical, and transparent integration of generative tools into academic work, a dimension that is particularly relevant in Latin American contexts where the institutional integration of AI is progressing unevenly (Acosta-Enriquez et al., 2025; Garro-Aburto et al., 2025).

The findings should be interpreted with certain limitations in mind. First, the study relied on self-reports, so responses may have been influenced by social desirability, especially given that academic integrity is a strongly normative construct. Second, the use of non-probabilistic purposive sampling limits the external validity and generalizability of the findings. The sample was composed predominantly of undergraduate students enrolled in on-campus programs; therefore, the results may not fully apply to graduate education, fully online programs, or more institutionally heterogeneous settings. Third, although the final model showed favorable indicators, the residual covariances identified in the CFA need to

be tested in new samples to confirm their stability and to determine whether item wording should be further refined. Fourth, the two-factor solution may be associated with the cultural and academic context of the sample; therefore, future research should examine the 15-item, two-factor version in other educational populations, including different academic levels, institutional types, disciplinary areas, and degrees of exposure to generative AI. Fifth, temporal stability was not assessed, so future studies should estimate test–retest reliability. Finally, the validity of the questionnaire should be examined in relation to more objective or behavioral indicators of academic integrity and AI use, such as declared use of AI, citation and acknowledgement practices, task-level AI disclosure data, transparency statements in assignments, performance on authentic academic writing tasks, and, when ethically and institutionally available, records of academic misconduct. Given that the questionnaire was administered and psychometrically validated in Spanish, it will also be necessary to evaluate the measurement properties and cross-linguistic equivalence of an English version.

5. Conclusions

The study provided favorable psychometric evidence regarding a refined version of the academic integrity questionnaire in AI-mediated contexts, as applied to college students. The results support a two-factor structure of normative integrity in AI use and support for academic writing, suggesting that academic integrity in generative technology-mediated scenarios is organized not only around normative compliance but also in relation to the legitimate use of AI as an academic support resource.

From a methodological perspective, the instrument constitutes a brief and conceptually interpretable tool for future research on academic integrity, AI literacy, and university practices mediated by generative technologies. From an applied perspective, it can support diagnostic processes, training, and institutional decision-making oriented toward the responsible, critical, and transparent use of AI in higher education.

However, the findings should be interpreted with caution due to the use of self-reports, the non-probabilistic sampling strategy, the predominance of undergraduate and on-campus students, the acceptable but borderline evidence of discriminant validity, the need to verify the stability of residual covariances in new samples, and the possible contextual nature of the observed factor solution. Likewise, although the invariance analysis provided favorable preliminary evidence for gender comparability, the elevated RMSEA of the configural model indicates that this result should be interpreted cautiously and replicated before the instrument is used for definitive comparisons between gender groups. Consequently, future research should examine test–retest reliability, validate the 15-item two-factor version in more diverse samples, extend the invariance evaluation to other academic and institutional groups, assess validity in relation to external behavioral indicators of academic integrity and AI use, and explore the cross-linguistic equivalence of the questionnaire in English-speaking samples.

Author Contributions: Conceptualization, L.L.G.-A. and J.M.C.-D.; methodology, J.M.C.-D. and C.A.H.S.; software, J.M.C.-D.; validation, J.M.C.-D., C.A.H.S. and E.M.L.E.; formal analysis, J.M.C.-D. and C.A.H.S.; investigation, L.L.G.-A., E.M.L.E. and S.P.S.S.; resources, L.L.G.-A., E.M.L.E. and S.P.S.S.; data curation, J.M.C.-D. and S.P.S.S.; writing—original draft preparation, L.L.G.-A., J.M.C.-D. and C.A.H.S.; writing—review and editing, L.L.G.-A., J.M.C.-D., C.A.H.S., E.M.L.E. and S.P.S.S.; visualization, J.M.C.-D. and C.A.H.S.; supervision, L.L.G.-A. and J.M.C.-D.; project administration, L.L.G.-A. and S.P.S.S.; funding acquisition, not applicable. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki. Formal ethical review was waived by the Research Ethics Committee of César Vallejo University, according to Report No. 00156-2025/CEI-PDE, due to the nature and methodological characteristics of the study.

Informed Consent Statement: Informed consent for participation was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author due to privacy and ethical restrictions.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

The final version of the questionnaire consists of 15 items divided into two dimensions: “Normative Integrity in AI Use” and “Support for Academic Writing.” Responses are recorded using a five-point Likert scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, and 5 = strongly agree.

Table A1. Original Spanish version and English translation of the final questionnaire.

Dimensión	Código	Redacción del Ítem	Escala de Respuesta
Integridad normativa en el uso de IA	EUA1	Utilizo herramientas de IA para apoyar mi pensamiento y no para reemplazarlo.	1–5
Integridad normativa en el uso de IA	EUA2	Consulto las normas o guías institucionales antes de usar IA en mis trabajos académicos.	1–5
Integridad normativa en el uso de IA	EUA3	Reflexiono regularmente sobre si el uso que hago de la IA respeta los principios de integridad académica.	1–5
Integridad normativa en el uso de IA	EUA4	Me aseguro de que la ayuda de la IA no afecte la autenticidad de mi estilo o voz académica.	1–5
Integridad normativa en el uso de IA	EUA6	Reconozco explícitamente el uso de herramientas de IA cuando estas han contribuido de manera significativa a la elaboración de mis trabajos académicos.	1–5
Integridad normativa en el uso de IA	ARM1	Verifico que el contenido generado por IA coincida con fuentes académicas confiables antes de incluirlo en mis escritos.	1–5
Integridad normativa en el uso de IA	ARM2	Entiendo cómo el uso inadecuado de la IA puede reducir los objetivos de aprendizaje de una asignatura.	1–5
Integridad normativa en el uso de IA	ARM4	Reconozco que presentar un texto generado por IA como propio constituye una falta académica.	1–5
Integridad normativa en el uso de IA	ARM5	Sé que la IA puede generar información convincente pero incorrecta.	1–5
Integridad normativa en el uso de IA	ARM6	Soy consciente de que algunas herramientas de IA pueden eludir los sistemas de detección de plagio.	1–5
Apoyo a la escritura académica	SAW1	Utilizo la IA para mejorar la gramática y la redacción de mis trabajos.	1–5
Apoyo a la escritura académica	SAW2	Empleo la IA para aclarar conceptos complejos al elaborar mis informes o ensayos.	1–5

Table A1. Cont.

Dimensión	Código	Redacción del Ítem	Escala de Respuesta
Apoyo a la escritura académica	SAW3	Uso herramientas de IA cuando experimento bloqueo al escribir o necesito inspiración para redactar.	1–5
Apoyo a la escritura académica	SAW4	Solicito a la IA alternativas de redacción para expresar mejor mis propias ideas.	1–5
Apoyo a la escritura académica	SAW5	Utilizo la IA para revisar mi trabajo y evitar plagio involuntario.	1–5
Dimension	Code	Wording of the Item	Response Scale
Normative Integrity in AI Use	EUA1	I use AI tools to support my thinking, not to replace it.	1–5
Normative Integrity in AI Use	EUA2	I consult institutional policies or guidelines before using AI in my academic work.	1–5
Normative Integrity in AI Use	EUA3	I regularly reflect on whether my use of AI adheres to the principles of academic integrity.	1–5
Normative Integrity in AI Use	EUA4	I make sure that the AI's assistance does not compromise the authenticity of my academic style or voice.	1–5
Normative Integrity in AI Use	EUA6	I explicitly acknowledge the use of AI tools when they have made a significant contribution to the development of my academic work.	1–5
Normative Integrity in AI Use	ARM1	I verify that AI-generated content aligns with reliable academic sources before including it in my writing.	1–5
Normative Integrity in AI Use	ARM2	I understand how the inappropriate use of AI can undermine the learning objectives of a course.	1–5
Normative Integrity in AI Use	ARM4	I acknowledge that submitting AI-generated text as my own constitutes academic misconduct.	1–5
Normative Integrity in AI Use	ARM5	I know that AI can generate information that is convincing but incorrect.	1–5
Normative Integrity in AI Use	ARM6	I am aware that some AI tools can bypass plagiarism detection systems.	1–5
Support for Academic Writing	SAW1	I use AI to improve the grammar and writing in my work.	1–5
Support for Academic Writing	SAW2	I use AI to explain complex concepts when writing my reports or essays.	1–5
Support for Academic Writing	SAW3	I use AI tools when I am stuck while writing or need inspiration to write.	1–5
Support for Academic Writing	SAW4	I ask AI for alternative phrasings to better express my own ideas.	1–5
Support for Academic Writing	SAW5	I use AI to review my work and avoid unintentional plagiarism.	1–5

References

- Abdaljaleel, M., Barakat, M., Alsanafi, M., Salim, N. A., Abazid, H., Malaeb, D., Mohammed, A. H., Hassan, B. A. R., Wayyes, A. M., & Sallam, M. (2024). A multinational study on the factors influencing university students' attitudes and usage of ChatGPT. *Scientific Reports*, 14(1), 1983. [\[CrossRef\]](#) [\[PubMed\]](#)
- Acosta-Enriquez, B. G., Humaní Jordan, O., Morales-Angaspilco, J. E., Campoverde Ventura, G., Ruiz Carrillo, J. A., Blanco-García, L. E., & Huaranga Rivera, L. (2025). Influence of perceived ethics, prejudice, and teacher concerns on artificial intelligence literacy and implementation: A cross-sectional study using path analysis. *Computers in Human Behavior Reports*, 20, 100829. [\[CrossRef\]](#)
- Allen, L. K., & Kendeou, P. (2024). ED-AI Lit: An interdisciplinary framework for AI literacy in education. *Policy Insights from the Behavioral and Brain Sciences*, 11(1), 3–10. [\[CrossRef\]](#)
- Arseven, T., & Bal, M. (2025). Critical literacy in artificial intelligence assisted writing instruction: A systematic review. *Thinking Skills and Creativity*, 57, 101850. [\[CrossRef\]](#)
- Banihashem, S. K., Taghizadeh Kerman, N., Noroozi, O., Moon, J., & Drachsler, H. (2024). Feedback sources in essay writing: Peer-generated or AI-generated feedback? *International Journal of Educational Technology in Higher Education*, 21, 23. [\[CrossRef\]](#)
- Barrett, A., & Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20, 59. [\[CrossRef\]](#)
- Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, 54(5), 1160–1173. [\[CrossRef\]](#)
- Bearman, M., Ryan, J., & Ajjawi, R. (2023). Discourses of artificial intelligence in higher education: A critical literature review. *Higher Education*, 86, 369–385. [\[CrossRef\]](#)
- Chan, C. K. Y. (2025). Students' perceptions of 'AI-giarism': Investigating changes in understandings of academic misconduct. *Education and Information Technologies*, 30(6), 8087–8108. [\[CrossRef\]](#)
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43. [\[CrossRef\]](#)
- Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, 6, 100197. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., & Ahuja, M. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. [\[CrossRef\]](#)
- Essel, H. B., Vlachopoulos, D., Essuman, A. B., & Amankwa, J. O. (2024). ChatGPT effects on cognitive skills of undergraduate students: Receiving instant responses from AI-based conversational large language models (LLMs). *Computers and Education: Artificial Intelligence*, 6, 100198. [\[CrossRef\]](#)
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474. [\[CrossRef\]](#)
- Foltynek, T., Bjelobaba, S., Glendinning, I., Khan, Z. R., Santos, R., Pavletic, P., & Kravjar, J. (2023). ENAI recommendations on the ethical use of artificial intelligence in education. *International Journal for Educational Integrity*, 19(1), 12. [\[CrossRef\]](#)
- Garro-Aburto, L. L., Rivera-Arellano, E. G., Chávez-Díaz, J. M., & Ochoa-Guevara, S. P. (2025). Motivation for scientific publication at the university level: Analyses in Peru, Ecuador, and Colombia. *Education Sciences*, 15(11), 1468. [\[CrossRef\]](#)
- Habibi, A., Muhaimin, M., Danibao, B. K., Wibowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence*, 5, 100190. [\[CrossRef\]](#)
- Huang, C., Alhur, A. A., Azam, M., & Naeem, S. B. (2025). The challenge of academic integrity in the age of generative artificial intelligence. *Libri*, 75(4), 339–354. [\[CrossRef\]](#)
- Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, 8, 100348. [\[CrossRef\]](#)
- Lancaster, T. (2023). Artificial intelligence, text generation tools and ChatGPT: Does digital watermarking offer a solution? *International Journal for Educational Integrity*, 19(1), 10. [\[CrossRef\]](#)
- Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploekl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, 6, 100221. [\[CrossRef\]](#)
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 100790. [\[CrossRef\]](#)

- López Gómez, H. E., Sanchez Soto, J. M., Aguado Riveros, U. I., Romero Nuñez, L. I., Castillo-Sáenz, R. A., Chávez-Díaz, J. M., Lévano Zegarra, C. I., & Huamani Palomino, W. (2024). Herramientas digitales y su papel para optimizar la detección de plagio académico. *Iberoamerican Journal of Science Measurement and Communication*, 4(3), 1–14. [\[CrossRef\]](#)
- Maharmah, A. A., Elfeky, A., Yacoub, R., Ibrahim, A., & Nemt-allah, M. (2025). Measuring ethical AI Use in higher education: Reliability and validity of the AI academic integrity scale for postgraduate students. *International Journal of Innovative Research and Scientific Studies*, 8(4), 707–715. [\[CrossRef\]](#)
- 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. [\[CrossRef\]](#)
- Nam, B. H., & Bai, Q. (2023). ChatGPT and its ethical implications for STEM research and higher education: A media discourse analysis. *International Journal of STEM Education*, 10, 66. [\[CrossRef\]](#)
- Ng, D. T. K., Wu, W., Leung, J. K. L., Chiu, T. K. F., & Chu, S. K. W. (2024). Design and validation of the AI literacy questionnaire: The affective, behavioural, cognitive and ethical approach. *British Journal of Educational Technology*, 55(3), 1082–1104. [\[CrossRef\]](#)
- Perkins, G. (2026). Resilient assessment in the age of AI: Authentic design and the case for verbal examinations in business education. *Assessment & Evaluation in Higher Education*, 1–16. [\[CrossRef\]](#)
- Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*, 49, 223–245. [\[CrossRef\]](#)
- 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(1), 15. [\[CrossRef\]](#)
- Torres-Díaz, J. C., Duarte, J., Rivera, D., & Flandoli, A. B. (2025). Artificial intelligence and academic integrity: Exploring plagiarism in Ecuadorian universities. *International Journal for Educational Integrity*, 21(1), 35. [\[CrossRef\]](#)
- van Niekerk, J., Delport, P. M. J., & Sutherland, I. (2025). Addressing the use of generative AI in academic writing. *Computers and Education: Artificial Intelligence*, 8, 100342. [\[CrossRef\]](#)
- Visedsun, N., Nupuang, S., Imjai, N., Chávez-Díaz, J. M., & Aujirapongpan, S. (2026). Developing FinTech adoption capability among Thai Gen Z finance students: Influencing AI literacy, self-directed learning and innovative thinking skill. *The International Journal of Management Education*, 24(3), 101437. [\[CrossRef\]](#)
- Wang, H., Dang, A., Wu, Z., & Mac, S. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence*, 7, 100326. [\[CrossRef\]](#)
- Yusuf, A., Pervin, N., & Román-González, M. (2024). Generative AI and the future of higher education: A threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21, 21. [\[CrossRef\]](#)

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.