Paradigm Academic Press
Research and Advances in Education
ISSN 2788-7057
SEP. 2026 VOL.5, NO.3

Awareness, Perceived Benefits, and Skill-Erosion Concerns Regarding Generative AI Among Higher Education Students

Mark Kevin Astrero[1], Saranay I. Doyaoen[1], Perfecto M. Garcia[1], Darius King Galo[1] & Aileen Melody de Vera-Galo[1]

[1] College of Education, Northeastern College, Philippines

Correspondence: Mark Kevin Astrero, College of Education, Northeastern College, Philippines.

doi:10.63593/RAE.2788-7057.2026.09.006

Abstract

This research aimed to assess awareness of generative AI limitations, highlight its advantages, and identify potential threats from misuse, focusing on differences in opinions between future educators and students from non-education fields. The data were collected through surveying 931 undergraduate students of a private university in the Philippines. The researchers used a comparative-correlational research design with a qualitative strand. The study found moderate awareness of generative AI's limitations (M = 3.31, SD = 0.94) and a moderate-to-cautious attitude toward its implementation (M = 3.15, SD = 0.87). Education students showed higher awareness than students from other majors (d = 0.40), while differences in usage, perception, and fear of skills deteriorating were negligible (d = 0.14–0.25) and, to some extent, neutralized in robustness checks. Differences between genders and years of study in perceptions were minor. The findings suggest that future teachers' curricula should include a course on AI literacy and emphasize developing critical thinking skills to prepare students for implementing generative AI solutions in teaching practices.

Keywords: generative artificial intelligence, ChatGPT, higher education, pre-service teacher education, academic integrity, critical thinking, technology acceptance

1. Introduction

After ChatGPT was launched to the general public in late 2022, generative AI's presence in higher education institutions increased from something brand new to something increasingly common among college students. Major publications now state that many students applied generative AI tools for at least one task: brainstorming, summarizing content, or working on complete projects (Ravšelj et al., 2025). The expansion outpaced adjustments to higher education policies, faculty training, and curricula to incorporate new technologies. As a result, students had to develop their own norms for proper tool use (Baek et al., 2023; Petricini et al., 2023).

The consequences of this gap differ widely across educational fields. For students preparing to become teachers in universities, the way they become acquainted with generative AI technologies has far-reaching implications: future teachers will soon become agents in decisions about the use of those technologies by their own students. Several studies on future teachers' readiness to integrate AI into their teaching show that attitudes formed during initial professional training significantly influence classroom decisions about integrating AI technologies into the educational process. However, few large-scale studies have compared education students' awareness and risk perceptions of using the technology with those of students in other fields.

This research addresses this gap by using survey data from 931 undergraduate students at a multidisciplinary private institution of higher learning in the Philippines, including the College of Education and other fields such as criminology, information technology, hospitality management, geodetic engineering, and business administration. It aims to answer three research questions:

(a) What is the level of awareness among students of the limitations, benefits, and concerns of generative AI?

(b) Are there any differences in these constructs by sex, school year, and, primarily, between students in education programs and those in programs other than education?

(c) What has this group of students identified as fair or unfair use of generative AI, in their opinion, and what do they suggest doing to secure skill enhancement?

The study combines Likert-scale measures adapted from the student survey of Chan and Hu (2023) with open-ended qualitative responses, following earlier survey work on generative AI in higher education (Damiano et al., 2024).

2. Literature Review

2.1 Student Awareness and Perceived Benefits of Generative AI

Across national contexts, survey results indicate a mix of excitement and trepidation among students about using generative AI. The global study by Ravšelj et al. (2025), covering more than 23,000 cases, found that the most popular uses of ChatGPT included finding research materials, brainstorming, and summarizing. However, participants thought the technology had shortcomings in supporting academic processes and called for regulation by authorities. Chan and Hu (2023) studied university students in Hong Kong and found that while students viewed generative AI as having potential for personalized learning, they also raised concerns about accuracy and privacy.

In studies conducted by Ngo (2023) and Yang (2024), it was noted that university students see time savings, personalized feedback, and the ability to access services 24/7 as key benefits, while also identifying the need to assess digital sources as one of the main limitations. In turn, Himang et al. (2026) studied college students in the Philippines using the Unified Theory of Acceptance and Use of Technology, focusing on performance-related expectations and social influence as strong factors affecting ChatGPT adoption among students. It offered evidence that these variables are moderated by participants' age, gender, and prior experience.

2.2 Concerns About Over-Reliance and Skill Erosion

Another line of research focuses on the possible consequences of overusing generative AI and its effects on education. Jose et al. (2025) called it the "cognitive paradox." Generative AI can reduce cognitive load, allowing students to engage in higher-order thinking. However, the authors explain that excessive use of AI-generated answers can affect the cognitive engagement needed for meaningful learning, which can be confirmed with empirical data. The article authored by Jayasuriya et al. (2025) describes an experiment that was conducted with undergraduate and postgraduate degree students who study agricultural sciences. The researchers used exploratory factor analysis and identified over-reliance on generative pre-trained transformers as a separate dimension that can be measured.

As per the findings of Cahyani et al. (2024), it was noted that there is a negative correlation between the frequency of AI tool use and the performance of university students taking critical reading tests. In contrast, in a study performed by Hading et al. (2024), where qualitative interviews were conducted with students of the English literature department, it was discovered that the students themselves believed that their critical thinking skills were declining and that they became increasingly dependent on technology, even though they praised the efficiency of AI use in obtaining information. Another point of concern is academic misconduct and reduced creativity (Baek et al., 2023; Petricini et al., 2023; Zafar et al., 2024).

2.3 Pre-Service Teachers and Generative AI Readiness

Because the success of teacher education programs affects who will shape how students interact with artificial intelligence (AI), several studies have examined generative AI and AI readiness among pre-service teachers. Salas-Pilco et al. (2022) carry out a systematic analysis of the use of AI in the educational process by analyzing studies on teacher education, highlighting that the majority of the research focuses on describing the actions, attitudes, and perceptions of pre-service teachers about artificial intelligence, while ethical practices and long-term effects of the use of AI are still at a very low level. According to Gatlin (2023), despite a clear understanding of the importance of incorporating AI technologies into teaching methodologies, U.S. pre-service teachers felt underprepared and untrained to approach the use of AI in education. Özüdoğru and Yıldız Durak (2025), in their questionnaire and 816 interviews, found that AI readiness is influenced by factors such as AI literacy, AI fear, and beliefs about AI technology innovation, and is significantly differentiated by gender.

Similarly, Eyüp and Kayhan (2023) found that pre-service language teachers' attitudes toward AI were moderated by anxiety about losing their jobs and a lack of control over their day-to-day responsibilities. This phenomenon has been analyzed by Zakaria and Hashim (2024) through the lenses of the Unified Theory of Acceptance and Use of Technology and the Technological Pedagogical Content Knowledge frameworks.

Further, the study conducted by Zaugg (2024) revealed that the experience of using AI teaching assistants contributed to a greater willingness of pre-service teachers to allow their students to use such technologies in the future, implying that, in the process of training, teachers should be presented with and come into contact with the technology they will use in their future professional practice. The results indicate that pre-service teachers occupy a unique place in this context, as they are both current users and future gatekeepers of AI technology, making their views and ideas more important than those of other stakeholders involved in the process.

Even with this increasing knowledge, there are still very few studies that have been able to statistically compare education majors to non-education majors at the same college or university by using a reliable multi-construct measuring tool, and very few studies at the same time have been conducted in the context of private higher education in the Philippines, where teacher education consists of one of the biggest clusters of educational degrees in the country (Damiano et al., 2024; Ravšelj et al., 2025). The current research addresses this gap.

3. Research Method

This study used a comparative-correlational design with a cross-sectional Likert-scale survey. The quantitative component answered the study's comparative and correlational research questions, while an embedded qualitative strand, consisting of three open-ended items, addressed the third research question. The study included 931 undergraduate students at a private higher education institution in the Philippines. The undergraduate student body included learners in education, criminology, information technology, hospitality management, engineering, and business, while fewer students pursued political science, mass communications, management accounting, and midwifery. A census approach to sending out invitations was adopted in this study, where all the undergraduate students who were enrolled during the first semester of the academic year 2025-2026 were requested to fill out an online questionnaire using Google Forms from 25th October to 2nd November 2025. Participation was voluntary, so the 931 students who consented and submitted the questionnaires were classified as a voluntary response sample derived from a census-type invitation rather than complete enumeration; the implications for representativeness are discussed in the Limitations chapter. All submissions were complete for closed-ended items, so no cases were dropped. Participants were 473 males (50.8%) and 458 females (49.2%). The demographic segmentation also included year level, with 267 (28.7%) first-year students, 272 (29.2%) sophomores, 282 (30.3%) juniors, and 110 (11.8%) seniors. The study also analyzed distinctions in academic majors. It stratified students into teaching programs (Bachelor of Elementary Education and Bachelor of Secondary Education in English, Filipino, Mathematics, and Science) and all other degrees (non-teaching programs).

Table 1. Demographic Characteristics of Respondents (N = 931)

Characteristic	n	%
Sex		
Male	473	50.8
Female	458	49.2
Year Level		
First Year	267	28.7
Second Year	272	29.2
Third Year	282	30.3
Fourth Year	110	11.8
Degree-Program Cluster		
Teacher Education (BEEd/BSEd)	264	28.4
Non-Education Programs	667	71.6

The questionnaire was divided into 4 sections. The first section included demographic details such as gender, year of study, and course studied, as well as a single question about previous experience incorporating generative AI. The second section included 7 questions assessing awareness of generative AI limitations and risks (for example, that generative AI can produce false, biased, and inappropriate information). The third part of the questionnaire included 8 more questions about recognizing benefits and intentions to use generative AI in studies or teaching processes. Lastly, the fourth section included 4 questions about concerns related to the loss of professional skills and becoming too dependent on generative AI (such as academic integrity and connection with others). The Likert scales used five items for agreement (1 = strongly disagree and 5 = strongly agree),

while generative AI was measured with a single five-point scale. Each of the three sections with attitude questions included a single open-ended question. The items were based on Chan and Hu's student survey from 2023, and descriptions of the questions can be found in Appendix A. When preparing the online survey, we administered the first two awareness questions in identical form (Appendix A). Therefore, a follow-up analysis used a six-item awareness composite.

Cronbach's alpha (Cronbach, 1951) was used to determine the internal consistency of each multi-measure construct. All three constructs met the conventional cut-off of .70: Awareness of Limitations, $\alpha = .93$; Perceived Benefits, $\alpha = .94$; and Skill-Erosion Concerns, $\alpha = .85$, indicating good internal consistency (Table 2). With the unnecessary awareness component eliminated, α for the six-item awareness composite stood at .92. The composite variables from each variable cluster were estimated as the mean of all the initial variables included in the cluster. Descriptive statistics, independent-samples t-tests (such as for sex and degree-program cluster), ANOVA (for year level), and Pearson correlations were carried out for the quantitative strand; the detailed information about the assumption tests and the effect sizes used is given below. As for the qualitative analysis, the responses to the questions were coded using the inductive thematic coding approach with respect to the important words. Two researchers separately created preliminary code lists from a small subset of responses and then adjusted them in discussion with each other. They coded every item in the entire dataset.

To illustrate how the score levels should be understood, the composite means have been allocated into five equal-interval bands taken from the answer options. Since the scale range (5 - 1 = 4) was divided by the number of options (5), the interval was 0.80. The means of 1.00-1.80 were defined to mean very low, 1.81-2.60 low, 2.61-3.40 moderate, 3.41-4.20 high, and 4.21-5.00 very high. These bands are an important way to describe the mean's location on the answer scale; they are not diagnostic cut-off mark, which is why they were not used in any statistical testing during this study.

The distributions of all composites were checked prior to the parametric analyses. The outputs of the Shapiro–Wilk and Kolmogorov–Smirnov tests were found to be significant for each of the four variables (all $p < .001$). However, with $N = 931$, these tests may detect even small deviations from normality. Hence, it was decided to rely on the skewness and kurtosis indices suggested for large samples (Kim, 2013). The skewness indices ranged from -0.33 to -0.14. The kurtosis indices ranged from -0.47 to 0.27, within the absolute thresholds for significant non-normality (Kim, 2013), indicating an almost normal distribution (see Table 2). Levene's test indicates unequal variances for some comparisons (for instance, awareness by sex, $F = 10.01$, $p = 0.002$, and others). We tested group differences using Welch's t-tests, compared with categorical grouping by year level, and post hoc comparisons using Games-Howell between sexes and degree programs, without assuming homogeneity of variance (Delacre et al., 2017). Mann-Whitney U and Kruskal-Wallis tests were also performed for robustness checks, leading to the same conclusions as the parametric tests in every comparison. Effect sizes were measured using Cohen's d for two-group comparisons and η^2 (ω^2 as a less biased estimator) for year-level comparisons and interpreted with reference to the conventional effect sizes of 0.20, 0.50 and 0.8 for d and .01, .06 and .14 for η^2 (Cohen, 1988). Pearson correlations were also calculated between the constructs and further validated via computation of Spearman's ρ.

Because all constructs were measured with the same instrument at a single point in time, it was also relevant to examine common method and response style effects (Kock et al., 2021; Podsakoff et al., 2003). Harman's single-factor test was applied to all 20 closed-ended items, and straight-lining (identical responses to all 19 Likert items) and response variability within respondents were calculated to analyze response style. Then it was possible to carry out the analysis of the comparison age groups again (a) not taking into account those respondents whose responses demonstrated straight-lining and (b) utilizing the data, which presupposes that the average value of each respondent for all the Likert items is subtracted from each response on a single question in order to demonstrate only the relative degree of emphasis by the respondent on the topic of the issue being studied.

The study adhered to the National Ethical Guidelines for Research Involving Human Participants (Philippine Health Research Ethics Board [PHREB], 2022) as well as to the Data Privacy Act of 2012 (Republic Act No. 10173). Administrative approval from the institution via its Office of the Dean was sought prior to administration of the survey process. Ethical review for the study was not sought from any research ethics committee; instead, the study was structured as a minimal-risk, anonymous, non-interventional survey where participation was voluntary and no sensitive data were collected. Respondents participated voluntarily and received information about the research objectives before completing the questionnaire. The first page of the online questionnaire included a consent form assuring that participation was voluntary, participants could withdraw at any time, and their answers would be considered only in aggregate, without linking them to any individual respondent. Only those who clicked on the "I agree" button proceeded to answer questions (Appendix A). All 931 respondents consented to participate. No identifying information of respondents apart from the

information about the program and academic year was considered or reported. The names and emails of the respondents that were collected for verification purposes were deleted prior to data analysis.

4. Results

4.1 Descriptive Overview

The descriptive statistics for the four constructs utilized in the study can be seen in Table 2. On a scale of 1 to 5, students indicated moderate frequency of use of generative AI (M=3.08, SD=1.09), moderate awareness of the limitations of generative AI (M=3.31, SD=0.94), moderate perceived benefits and behavioral intentions (M=3.15, SD=0.87), and moderate concern about skill loss and over-dependency (M=3.04, SD=0.86). The four constructs are clustered at the moderate level (2.61–3.40), which is noted. However, the distribution of means was wide: while 46.8% of students scored high or very high on awareness of shortcomings, only 35.3% did so for perceived benefits and 31.5% for concerns about skill erosion. On the other hand, 18.8%–25.7% of students scored low and very low in the three constructs (see Table 2). All of these constructs have been normally distributed. Thus, the generative AI attitude of the sample is better described as cautious involvement rather than enthusiastic acceptance or rejection.

Table 2. Descriptive Statistics and Internal Consistency of Study Constructs

Construct	Items	M	SD	Level	Skewness	Kurtosis	Cronbach's α
Generative AI Usage Frequency	1	3.08	1.09	Moderate	−0.13	−0.47	—
Awareness of Limitations/Risks	7	3.31	0.94	Moderate	−0.33	−0.01	.93
Perceived Benefits/Behavioral Intention	8	3.15	0.87	Moderate	−0.29	0.15	.94
Skill-Erosion/Over-Reliance Concerns	4	3.04	0.86	Moderate	−0.17	0.27	.85

Note. Level bands: 1.00–1.80 = very low; 1.81–2.60 = low; 2.61–3.40 = moderate; 3.41–4.20 = high; 4.21–5.00 = very high (descriptive equal-interval convention; see Research Method). Percentage of respondents in the low/very low, moderate, and high/very high bands, respectively: usage, 26.6%, 39.7%, 33.7%; awareness, 18.8%, 34.5%, 46.8%; benefits, 20.5%, 44.1%, 35.3%; concerns, 25.7%, 42.6%, 31.5%. Shapiro–Wilk tests were significant for all composites (W = .911–.973, all $p < .001$); skewness and kurtosis were within ±2 and ±7, respectively. α for the six-item awareness composite (duplicated item removed) = .92.

4.2 Differences by Sex

Results from Welch's independent-samples t-tests (Table 3) revealed that female students have significantly more awareness than male students when it comes to the limitations of generative AI (Female M = 3.40, SD = 0.84; Male M = 3.23, SD = 1.02; $t(905.5) = -2.71$, $p = .007$, $d = −0.18$). The same significant difference also occurred in the frequency of usage (Female M = 3.17, SD = 1.00; Male M = 2.99, SD = 1.16; $t(917.0) = -2.60$, $p = .009$, $d = −0.17$). Gender did not contribute to whether students benefited from AI ($t(893.1) = 0.28$, $p = .776$, $d = 0.02$) or worried about losing skills when using the technology ($t(925.5) = -1.55$, $p = .122$, $d = −0.10$). Both significant differences were negligible in magnitude ($|d| < 0.20$), meaning the score distributions of female and male students overlapped almost completely; with N = 931, statistical significance reflects sample size more than a practically meaningful gap.

Table 3. Comparison of Construct Means by Sex

Construct	Male M (SD)	Female M (SD)	Welch t (df)	p	d
Awareness	3.23 (1.02)	3.40 (0.84)	-2.71 (905.5)	.007	−0.18
Perceived Benefits	3.16 (0.97)	3.14 (0.76)	0.28 (893.1)	.776	0.02
Concerns	2.99 (0.89)	3.08 (0.81)	-1.55 (925.5)	.122	−0.10
Usage Frequency	2.99 (1.16)	3.17 (1.00)	-2.60 (917.0)	.009	−0.17

Note. Male n = 473; female n = 458. Welch's t-tests (equal variances not assumed). Cohen's d is computed as male minus female; negative values indicate higher scores among female students.

4.3 Differences by Degree-Program Cluster

When comparing Teacher Education and Non-Education students (Table 4), it was found that the education students scored significantly higher than their counterparts in all parameters measured, about Awareness (M = 3.58, SD = 0.78 vs. M = 3.21, SD = 0.98; t(597.3) = 6.04, p < .001, d = 0.40), Perceived Benefits (M = 3.23, SD = 0.75 vs. M = 3.11, SD = 0.91; t(578.0) = 2.04, p = .042, d = 0.14), Skill Erosion Concerns (M = 3.17, SD = 0.76 vs. M = 2.98, SD = 0.89; t(562.1) = 3.19, p = .002, d = 0.22) and Usage Frequency (M = 3.27, SD = 0.93 vs. M = 3.00, SD = 1.14; t(585.5) = 3.79, p < .001, d = 0.25). In accordance with the magnitude, only the awareness gap was considered to be moderate (d = 0.40, 95% CI [0.26, 0.54]); other aspects like usage (d = 0.25) and worries about losing skills (d = 0.22) were slight, and the disparity in perceived benefits (d = 0.14) was not significant and would not be considered significant according to a pooled-variance t-test (p = .060). Consequently, the distinguishing feature of educational students may be their awareness of generative AI's limitations, while their scores on other items are modest; we will analyze these in the following section on response styles.

Table 4. Comparison of Construct Means by Degree-Program Cluster

Construct	Education M (SD)	Non-Education M (SD)	Welch t (df)	p	d
Awareness	3.58 (0.78)	3.21 (0.98)	6.04 (597.3)	< .001	0.40
Perceived Benefits	3.23 (0.75)	3.11 (0.91)	2.04 (578.0)	.042	0.14
Concerns	3.17 (0.76)	2.98 (0.89)	3.19 (562.1)	.002	0.22
Usage Frequency	3.27 (0.93)	3.00 (1.14)	3.79 (585.5)	< .001	0.25

Note. Education n = 264; non-education n = 667. Welch's t-tests (equal variances not assumed). Positive d values indicate higher scores among education students.

4.4 Response-Style and Common-Method Sensitivity Analysis

Harman's single-factor analysis showed one unrotated factor explaining 54.8% of the variance across 20 closed-ended questions (three factors had eigenvalues greater than 1), meeting the 50% heuristic; thus, common-method variance cannot be excluded. In terms of response patterns, the one cluster contained some information about response patterns: 20.7% of non-education people gave identical answers to 19 Likert items (mainly to the midpoint of the scale) and only 5.3% of the respondents from the education group; $\chi^2(1) = 31.66$, $p < .001$. When respondents who were straight-lining were removed (n = 779), it was found that the education group scored higher in terms of awareness (d = 0.30, p < .001) and usage (d = 0.20, p = .007), and only that of skill-erosion concerns (d = 0.14, p = 0.056) was a noticeable number in that respect. The results in IPSAT scores showed that education respondents paid greater attention to limitation awareness than did their non-education peers (d = 0.35, p < 0.001), while placing less attention to perceived benefits (d = -0.30, p < 0.001), and having no difference in terms of their focus on skill erosion (d = -0.10, p = 0.271). Results from a six-item awareness validity measure were almost identical to those from a seven-item measure (e.g., education vs. non-education, d = 0.39, p < 0.001).

4.5 Differences by Year Level and Intercorrelations

Welch's one-way ANOVA tests (Refer to Table 5) performed for the study were able to find significant differences in all four aspects studied, such as Awareness (F (3, 384.3) = 3.52, p = 0.015, $\eta^2 = 0.012$); Perceived Benefits (F(3, 390.8) = 4.06, p = 0.007, $\eta^2 = 0.013$); etc. Fourth-year students had the highest average scores for these four factors (e.g., Usage Frequency M = 3.33), compared with first-year students, whose average was lowest (e.g., Usage Frequency M = 2.92). Nonetheless, the observable differences were limited ($\eta^2 = .012–.025$; $\omega^2 = .008–.022$), and the course of action varied, since second-year students outperformed their third-year classmates in each category measured. Tests used showed differences in awareness between first- and fourth-year students (Mdiff = 0.31, p = .038), in benefits between fourth-year and first-year (Mdiff = 0.30, p = .018) and third-year students (Mdiff = 0.29, p = .031), in concerns between first-year and second-year students (Mdiff = 0.18, p = .041) and in usage between first-year students and third-year students as well as between second-year and fourth-year students (all $p \leq .015$). As the design was cross-sectional, these comparisons yielded information about the differences in cohorts at one point in time instead of the changes in students in the process of their education and the sample of the fourth-year students was quite small (n = 110).

Table 5. Construct Means by Year Level and Welch's ANOVA Results

Construct	First Year M (SD)	Second Year M (SD)	Third Year M (SD)	Fourth Year M (SD)	Welch F	p	η^2
Awareness	3.22 (0.87)	3.38 (0.78)	3.25 (1.08)	3.52 (1.05)	3.52	.015	.012
Perceived Benefits	3.07 (0.82)	3.21 (0.79)	3.08 (0.96)	3.37 (0.91)	4.06	.007	.013
Concerns	2.95 (0.82)	3.13 (0.76)	2.96 (0.92)	3.20 (0.94)	4.26	.006	.014
Usage Frequency	2.92 (1.05)	3.26 (0.99)	2.95 (1.18)	3.33 (1.10)	8.30	< .001	.025

Note. First year n = 267; second year n = 272; third year n = 282; fourth year n = 110. Welch's F has 3 numerator degrees of freedom; denominator degrees of freedom = 384.3 (awareness), 390.8 (benefits), 387.4 (concerns), and 393.5 (usage). η^2 is from the conventional one-way ANOVA. Kruskal–Wallis tests yielded the same conclusions (all $p < .01$).

As shown in Table 6, all factors were significantly and positively intercorrelated (all $p < .001$). In particular, awareness correlated the most with benefit perception ($r = .67$) and skill-erosion concerns ($r = .65$). In contrast, benefit perception was also strongly correlated with the concerns ($r = .68$). Thus, we see the situation where awareness, benefit perception, and risk concern grow together rather than supplanting one another. The Spearman coefficients indicated similar results ($\rho = .42–.64$, all $p < .001$). Some of this correlation may be explained by using the same self-reported measures. Usage frequency and skill erosion concerns were highly correlated ($r = .44$). However, once awareness of concern and perceived benefits were accounted for, r dropped to almost zero ($r = .05$; $p = .102$). However, when awareness was controlled for perceived benefits, r remained high: $r = .37$; $p < .001$. It suggests that skill erosion concerns might be correlated with awareness of shortcomings rather than with the usage itself.

Table 6. Intercorrelations Among Study Constructs

Variable	1	2	3	4
1. Usage Frequency	—			
2. Awareness	.46*	—		
3. Perceived Benefits	.56*	.67*	—	
4. Concerns	.44*	.65*	.68*	—

Note. * $p < .001$.

4.6 Qualitative Findings

The students' responses to the three open-ended questions (n = 880–903) were grouped into themes that largely matched the quantitative results of the survey. When asked about the benefits and disadvantages of generative AI, students cited research and idea generation (28.7%) as its major use, followed by using generative AI as a study aid for clarification (25.7%), fact-checking (18.1%), and time-saving (12.1%); less than 1% of the respondents stated that they would make use of generative AI very rarely as a last resort.

Students responded about when generative AI use could be beneficial or harmful. Students frequently mentioned its usefulness in understanding complex topics (21.4%) and in helping generate ideas (20.3%). On the negative side, the most commonly mentioned concern was over-reliance that can lead to procrastination (12.5%), along with plagiarism (6.9%) and a diminished ability to think critically (4.3%).

When students were asked what changes instructors and the institution should implement to safeguard critical thinking, problem solving, and various transferable skills, most of them suggested that there should be more emphasis on group work/teamwork and activities with collaborative elements (22.6%), along with the necessity to include tasks with a requirement of critical thinking and independent problem solving (19.5%). The second suggestion was to use more practical, real-world activities (15.5%) and assess students either in class or using tools that would restrict AI (6.0%). Only a small number of students (3.0%) expressed their opinion that no changes were needed. Given the breadth of students' suggestions, we can surmise that student anxiety about losing critical thinking skills accompanies recognition of the benefits of generative AI.

5. Discussion and Conclusions

This research aimed to analyze higher education students' awareness of generative AI, the benefits it offers, and concerns about skill loss in the Philippines. It examined how pre-service teacher education students differ from other students. The findings are discussed below in relation to the three research questions.

Regarding the first research question, the sample's overall profile shows moderate values around the midpoint across all constructs, which aligns with global trends in "cautious engagement." Like those examined in the analysis of Ravšelj et al. (2025) and Chan and Hu (2023), in this case the students were divided, as they neither welcomed nor rejected generative AI as they realized its advantages, such as time saving, easy explanations, and ability to generate ideas, but understood the drawbacks, including, uncertain results or danger of dependence, which has been repeatedly confirmed by the studies by Ngo (2023) and Yang (2024). The results indicate that Filipino students' attitudes toward generative AI are not new and align with those developed internationally, and they also show that students accept the technology model mentioned earlier. However, the moderate means mask high heterogeneity: almost half of respondents are very conscious of the limitations, and about one-fourth are not very concerned about skill decline. It means that a uniform AI literacy intervention will find students at very different starting points.

For the second question, education students scored higher on each of the four measures than non-education students; however, the effect sizes and sensitivity analysis suggest a need for a more nuanced interpretation rather than a general one. The most significant difference was in awareness of restrictions: it showed the strongest effect ($d = .40$) and remained significant even after removing responses from non-differentiated respondents and in the ipsatized analysis, in which education students emphasized restrictions much more than their friends. This is in line not only with the research that the readiness of future teachers to utilize AI in practice is influenced by information literacy and feelings of danger, but not just by experience (Özüdoğru & Yıldız Durak, 2025), but also with the results of the study revealing biased attitudes among future language teachers (Eyüp & Kayhan, 2023). In comparison, the larger concern and benefit scores among education majors were narrow. After adjusting for response style, they minimized these scores even more or eliminated them completely. Something as simple as the engagement of education majors in the survey could be the main reason for their high scores, as they tended to avoid straight-lining and to distinguish between different questions in a better way, meaning that the difference between the groups could be due to the effects of response style or common method, and not to their more efficient approach. Moreover, this casts doubt on the claim that AI users should not be critical of it, because the case at hand shows that frequent users and those who are aware of AI to some degree belong to the same group of respondents. This, however, facilitates the emergence of the "cognitive paradox" (Jose et al., 2025, p. 1) where the same methods which relieve the cognitive burden from users compel them to measure the costs of using them.

One tentative explanation for the higher awareness of education majors, which requires direct testing in future work, is related to the concept of social foresight — future teachers may already be thinking about how generative AI will influence the skills development of their future students, which would imply a form of indirect risk perception that is missing in the case of more conventional students. This view aligns with Zaugg's (2024) finding that pre-service teachers' willingness to let their students work with AI assistants is strongly influenced by their prior experience with AI-based tools during their education and training.

With regard to the third research question, the qualitative conclusions offer a constructive idea generated by students for addressing concerns about overreliance on generative AI, without the need to eliminate the use of generative AI. Instead of calling for restrictions and prohibitions, students mainly advocated for a greater incorporation of teamwork activities and assignments that require some independent critical thinking — recommendations which are in line with research studies in instructional design that warn about the potential passive nature of using AI-generated materials, where students will not engage in active analysis of the received information (Cahyani et al., 2024; Hading et al., 2024; Jayasuriya et al., 2025). Regarding teacher education programs, it means that two approaches should be undertaken: first, it should be integrated into the program to ensure that future teachers receive the appropriate skills for separating effective from ineffective use of generative AI, consistent with studies that emphasize the importance of hands-on experience (Gatlin, 2023; Zaugg, 2024); and second, some assignments that generative AI cannot do should be preserved, as students have also suggested.

Overall, the findings indicate that the prevalent notion of distinguishing between "embracing" or "restricting" generative AI in institutions is likely not applicable, particularly from the point of view of students. Specifically, students who used generative AI more often also tended to be more aware of its limitations, and concern about skill erosion was tied more closely to this awareness than to frequency of use itself, which points to awareness, rather than avoidance, as the lever for protecting skill development. This finding prompts the College of Education and relevant institutions to rethink how they teach students about AI.

6. Limitations

It is important to consider certain limitations when interpreting these findings. Firstly, findings are from a cross-sectional design and do not allow for establishing any causal or developmental relationships; year-level differences are equated to the cohorts examined at a specific point in time instead of tracking how the students grow. Second, the data come from one private university in one region of the Philippines. Although all undergraduates at this university were invited to participate, the sample remains a voluntary response sample, as most participants are criminology, information technology, and hospitality management students. Those who opted not to respond may have significantly different engagement levels with generative AI, making it impossible to generalize to the entire university and other universities in the Philippines.

The third limitation was that everything was determined through self-reporting. Harman's formula indicated that more than half of the factors were driven by a single factor, and variability in response types was observed across study programs. Thus, the analysis may be biased, and some validity-assessment techniques were not applied (Kock et al., 2021; Podsakoff et al., 2003). Fourth, the frequency of use was determined by a single item, and the awareness scale was based on two identical items, which increases the reliability of the awareness composite. Fifth, evidence for construct validity relies on the instrument's prior use (Chan & Hu, 2023) and on internal consistency, since the adapted items have not undergone confirmatory factor analysis in this sample. The qualitative part used keyword-based inductive coding of written responses, and the coders reached agreement through deliberation rather than calculating an agreement coefficient. Although institutional permission and informed consent were obtained, the ethics review board did not approve the protocol.

The research should apply probability sampling in collaboration with several educational institutions, use longitudinal designs to observe students throughout their college years, evaluate their behaviors and effectiveness, and validate research instruments with Filipino students.

7. Implications

The study, which used both qualitative and quantitative methods and included 931 Filipino college students, yielded mixed results on awareness of generative AI tools, perceptions of their value, and concerns about skills that might be lost while using them. The outcomes of the research indicate that the teacher education curricula need to introduce the concept and teach students to utilize generative AI, while at the same time giving them the opportunities to accomplish sophisticated tasks without the use of technology; teachers and policymakers might forget the fact that the discussions of the generative AI application are not to be based merely on the question of whether to accept the technologies or to ban them.

Author Contributions

Conceptualization, M.K.A. and S.I.D.; methodology, M.K.A., S.I.D., and P.M.G.; software, M.K.A. and D.K.G.; validation, M.K.A., S.I.D., P.M.G., D.K.G., and A.M.V.-G.; formal analysis, M.K.A. and S.I.D.; investigation, M.K.A., S.I.D., P.M.G., D.K.G., and A.M.V.-G.; resources, M.K.A. and A.M.V.-G.; data curation, M.K.A., P.M.G., and D.K.G.; writing—original draft preparation, M.K.A. and S.I.D.; writing—review and editing, M.K.A., S.I.D., P.M.G., D.K.G., and A.M.V.-G.; visualization, M.K.A. and D.K.G.; supervision, M.K.A., S.I.D., and A.M.V.-G.; project administration, M.K.A. All authors have read and agreed to the published version of the manuscript.

Funding

This research received no external funding.

Institutional Review Board Statement

This research was done in compliance with the National Ethical Guidelines for Research Involving Human Subjects (PHREB, 2022) and the Data Privacy Act of 2012 (Republic Act No. 10173). The organization in question approved the survey through the Office of the Dean. Because this survey did not collect any personal data and posed minimal risk to respondents, it was not submitted for ethics review. The authors obtained informed consent electronically from participants before the survey began. Any personal details that were obtained for administrative purposes were removed prior to the analysis.

Acknowledgments

The authors would like to acknowledge the administrative and logistical support provided by the participating institution in facilitating data collection. During the preparation of this manuscript, the authors used Generative AI tools solely as a study topic and research object; no Generative AI tools were used to generate text, data, or graphics for the manuscript. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest

The authors declare no conflicts of interest

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 the publisher and/or the editor(s). The publisher 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.

References

Baek, C., Tate, T., & Warschauer, M. (2023). "ChatGPT seems too good to be true": College students' use and perceptions of generative AI. *OSF Preprints.* https://doi.org/10.31219/osf.io/6tjpk

Cahyani, R. Y. D., Dewi, O. C., & Darmawan, A. (2024). AI and critical reading skill among university students: The impact of artificial intelligence (AI) on Al-Ghifari University students' critical reading skills. *TheGIST, 6*(1), 1–10. https://doi.org/10.53675/gist.v6i1.1374

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*, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203771587

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297–334. https://doi.org/10.1007/BF02310555

Damiano, A. D., Lauría, E. J. M., & Sarmiento-Cano, C. (2024). Early perceptions of teaching and learning using generative AI in higher education. *Journal of Educational Technology Systems, 52*(3), 346–375. https://doi.org/10.1177/00472395241233290

Data Privacy Act of 2012, Republic Act No. 10173 (2012) (Phil.). https://privacy.gov.ph/data-privacy-act/

Delacre, M., Lakens, D., & Leys, C. (2017). Why psychologists should by default use Welch's t-test instead of Student's t-test. *International Review of Social Psychology, 30*(1), 92–101. https://doi.org/10.5334/irsp.82

Eyüp, B., & Kayhan, S. (2023). Pre-service Turkish language teachers' anxiety and attitudes toward artificial intelligence. *International Journal of Education and Literacy Studies, 11*(4), 43–56. https://doi.org/10.7575/aiac.ijels.v.11n.4p.43

Gatlin, M. (2023). Assessing pre-service teachers' attitudes and perceptions of using artificial intelligence in the classroom. *Texas Educator Preparation, 7*(2), 1–8. https://doi.org/10.59719/txep.v7i2.35

Hading, E. F., Rustan, D. R. H. P., & Ruing, F. H. (2024). EFL students' perceptions on the integration of AI in fostering critical thinking skills. *GLENS: Global English Insights Journal, 2*(1), 1–10. https://doi.org/10.61220/glens.v2i1.466

Himang, C. M., Villa, S. C., Jr., & Nolon, N. F. (2026). Adoption of ChatGPT among higher education students. *International Journal of Information and Communication Technology Education, 22*(1), 1–29. https://doi.org/10.4018/ijicte.411220

Jayasuriya, R., Selvanayaki, S., & Deepa, N. (2025). Assessing the impact of GPTs on critical thinking, creativity and independent problem solving in academic learning of students. *Research Square.* https://doi.org/10.21203/rs.3.rs-6367830/v1

Jose, B., Cherian, J., & Verghis, A. M. (2025). The cognitive paradox of AI in education: Between enhancement and erosion. *Frontiers in Psychology, 16*, Article 1550621. https://doi.org/10.3389/fpsyg.2025.1550621

Kim, H.-Y. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52–54. https://doi.org/10.5395/rde.2013.38.1.52

Kock, F., Berbekova, A., & Assaf, A. G. (2021). Understanding and managing the threat of common method bias: Detection, prevention and control. *Tourism Management, 86*, Article 104330. https://doi.org/10.1016/j.tourman.2021.104330

Ngo, T. T. A. (2023). The perception by university students of the use of ChatGPT in education. *International Journal of Emerging Technologies in Learning, 18*(17), 4–19. https://doi.org/10.3991/ijet.v18i17.39019

Özüdoğru, G., & Yıldız Durak, H. (2025). Conceptualizing pre-service teachers' artificial intelligence readiness and examining its relationship with various variables. *Information Development, 41*(3), 916–932. https://doi.org/10.1177/02666669251335657

Petricini, T., Wu, C., & Zipf, S. T. (2023). Perceptions about generative AI and ChatGPT use by faculty and

college students. *OSF Preprints.* https://doi.org/10.35542/osf.io/jyma4

Philippine Health Research Ethics Board. (2022). *National ethical guidelines for research involving human participants 2022*. Philippine Council for Health Research and Development.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Ravšelj, D., Keržič, D., & Tomaževič, N. (2025). Higher education students' perceptions of ChatGPT: A global study of early reactions. *PLOS ONE, 20*(2), Article e0315011. https://doi.org/10.1371/journal.pone.0315011

Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences, 12*(8), Article 569. https://doi.org/10.3390/educsci12080569

Yang, H. (2024). Towards responsible use: Student perspectives on ChatGPT in higher education. *Proceedings of the European Conference on E-Learning, 23*(1), 415–422. https://doi.org/10.34190/ecel.23.1.2790

Zafar, S., Shaheen, F., & Rehan, J. (2024). Use of ChatGPT and generative AI in higher education: Opportunities, obstacles and impact on student performance. *IRASD Journal of Educational Research, 5*(1), 1–12. https://doi.org/10.52131/jer.2024.v5i1.2463

Zakaria, N. Y. K., & Hashim, H. (2024). Shaping the future of education: Conceptualising pre-service teachers' perspectives on artificial intelligence (AI) integration. *International Journal of Academic Research in Business and Social Sciences, 14*(5), 2066–2076. https://doi.org/10.6007/ijarbss/v14-i5/21584

Zaugg, T. (2024). Embracing innovation in education: Investigating the use and impact of artificial intelligent assistants among preservice teachers. *Research Square.* https://doi.org/10.21203/rs.3.rs-5032655/v1

Appendix A

Survey Instrument

Informed-Consent Statement (first screen)

Confidentiality and Ethics. Note that if you wish to take part, your participation is completely voluntary. You can choose not to participate in this study or withdraw from the study at any time. Any information you provide will be treated with the utmost confidentiality and anonymity. The information obtained from this study will only be used for statistical purposes and will not be used to identify the participants. Please click "I agree" if you are willing to participate in this survey. [Response option: I agree.]

Section 1. Profile and Generative AI Use

Sex; Year Level; Program

Have you ever used generative AI technologies like ChatGPT? (5-point scale; higher values indicate more frequent use)

Section 2. Awareness of Limitations/Risks (7 items)

A1. I understand that generative AI technologies like ChatGPT have limitations when handling complex tasks.

A2. I understand that generative AI technologies like ChatGPT can generate factually inaccurate output.

A3. I understand generative AI technologies like ChatGPT can generate out-of-context or inappropriate output.

A4. I understand generative AI technologies like ChatGPT can exhibit biases and unfairness in their output.

A5 I understand generative AI technologies like ChatGPT may rely too heavily on statistics, which can limit their usefulness in certain contexts.

A7. I understand generative AI technologies like ChatGPT have limited emotional intelligence and empathy, which can lead to insensitive or inappropriate output.

A6. I understand that generative AI technologies like ChatGPT can pose security and data privacy risks, as sensitive information shared in prompts might be stored or processed improperly.

Open-ended question 1: Considering what you know about the pros (benefits) and cons (risks) of tools like ChatGPT, how do you plan to use them in your work or studies to get the most help while avoiding the problems?

Section 3. Perceived Benefits/Behavioral Intention (8 items)

B1. I envision integrating generative AI technologies like ChatGPT into my teaching and learning practices in the future.

B2. Students must learn how to use generative AI technologies well for their careers.

B3. I believe generative AI technologies such as ChatGPT can improve my digital competence.

B4. I believe generative AI technologies such as ChatGPT can help me save time.

B5. I believe AI technologies such as ChatGPT can provide me with unique insights and perspectives I may not have thought of.

B6. I think AI technologies such as ChatGPT can provide me with personalized and immediate feedback and suggestions for my assignments.

B7. I think AI technologies such as ChatGPT are a great tool as they are available 24/7.

B8. I think AI technologies such as ChatGPT are a great tool for student support services because they offer anonymity.

Open-ended question 2: When is using ChatGPT for a school or work task helpful, and when is it harmful? Why?

Section 4. Skill-Erosion/Over-Reliance Concerns (4 items)

C1. Using generative AI technologies such as ChatGPT to complete assignments undermines the value of university education.

C2. Generative AI technologies such as ChatGPT will limit my opportunities to interact with others and socialize while completing coursework.

C3. Generative AI technologies such as ChatGPT will hinder my development of generic or transferable skills such as teamwork, problem-solving, and leadership skills.

C4. I can become over-reliant on generative AI technologies.

Open-ended question 3: Given your concerns that generative AI could hinder the development of skills (like problem-solving or teamwork) or lower the value of a university education, what is one specific change you think instructors or the university should make to assignments or course design to protect or even strengthen these skills?

Author Profile

MARK KEVIN ASTRERO is a college instructor at Northeastern College in Santiago City, Philippines, and a PhD student at St. Paul University Philippines. His research focuses on pre-service teacher development, English language competence, and entrepreneurial education. He actively contributes to educational research to enhance instruction and communicative skills.

SARANAY I. DOYAOEN serves as Dean and faculty member at Northeastern College, Philippines. With expertise in teacher education and mathematics, her research centers on student retention rates, academic policy formation, and pre-service teacher competencies. She is dedicated to institutional improvement and fostering academic success in higher education.

PERFECTO M. GARCIA III, LPT, is an Instructor at Northeastern College in the Philippines. Specializing in physical education, health, and fitness, he develops curriculum modules designed to promote physical literacy and active lifestyles among students. His work emphasizes holistic student development through physical activity and wellness education.

DARIUS KING A. GALO, LPT, is an educator and researcher affiliated with Northeastern College, Philippines. His scholarly work focuses on teacher education, English language proficiency, and overcoming public speaking anxiety among pre-service educators. He is committed to advancing pedagogical practices and enhancing communicative competence in classroom environments.

AILEEN MELODY M. DE VERA, LPT, is an educator and faculty member at Northeastern College, Philippines. She teaches courses in Art Appreciation and language studies (Masining na Pagpapahayag). Her work focuses on fostering creative literacy, aesthetic awareness, and effective linguistic communication among higher education students.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).