

# Exploring Ethical Perceptions of AI Use in Academic Integrity Among Students in Teacher Education and Business Studies

Angelito M. Rivera, Ferdinand L. Osená

College of Teacher Education, ELJ Memorial College, Philippines

DOI: <https://dx.doi.org/10.47772/IJRISS.2025.903SEDU0283>

Received: 19 May 2025; Accepted: 23 May 2025; Published: 23 June 2025

## ABSTRACT

The paper investigates ethical perceptions toward the use of AI in terms of academic integrity among students in the teacher education and business programs of ELJ Memorial College, Philippines. Informed by the Ethical Risk Mitigation Framework, the study explores how students in the two departments understand and rank order risks associated with AI, focusing on bias and discrimination, privacy and security, transparency, over-reliance, and equity. The research utilized an exploratory mixed-methods design, combining a descriptive cross-sectional survey with qualitative data collection via interviews. Quantitative data were analyzed using descriptive statistics (mean, median, and IQR) and Mann-Whitney U, while qualitative data were thematically analyzed to enrich understanding of student perceptions. The results indicate that both groups of students, irrespective of their area of studies, acknowledge the ethical dilemmas introduced by AI in education; however, their concerns vary: while teacher education students are more oriented toward bias, transparency, over-reliance, and equity, business students concentrate on privacy and security. Significant differences were also found in the five ethical categories. The results underscore the impact of disciplinary training on moral reasoning and suggest the value of contextually situated, discipline-specific AI ethics training in higher education. The paper provides recommendations for designing institutional policies that facilitate AI's responsible and equitable uptake.

**Keywords:** Ethical perceptions, Artificial Intelligence (AI), Teacher education, Business education, AI risk dimensions

## INTRODUCTION

With the continuing integration of artificial intelligence (AI) into higher education, which is transforming teaching, learning, and administrative processes, come critical ethical challenges as well, such as algorithmic bias, privacy breaches, and threats to academic integrity (Al-Zahrani, 2024; United Nations Educational, Scientific and Cultural Organization [UNESCO], 2024c). AI systems such as ChatGPT and adaptive learning platforms offer personalized instruction and efficient operations, but also risk perpetuating systemic inequities and undermining human agency (U.S. Department of Education Office of Educational Technology, 2023). Internationally, guidelines like those set by UNESCO stress values such as transparency, accountability, and fairness in AI interactions, and push institutions to find a balance between hiring trends and ethical implementation of the technology (UNESCO, 2024b). The fast pace of implementation of AI technology has often surpassed policy implementation, which could result in a lag behind addressing ethical issues in academic areas and institutional accountability (Temper et al., 2025a).

While there is increasing research about the impact of AI in education, few studies investigate how perspectives on ethics differ by academic discipline. For example, teacher education programs focused on pedagogical equity may emphasize different aspects of AI grading than business education programs, where efficiency and data-driven decision-making are privileged (Zhou et al., 2024a). This disciplinary distinction is the key as pedagogical values informed prioritization of risk factors such as algorithmic discrimination or over-reliance on automated systems ([1]; Department of Science and Technology [DOST], 2024a). Furthermore, although some institutions in the Philippines, such as the University of the Philippines Open University (UPOU), have a transparency protocol for AI, there is disparity between rural and urban areas when it comes to the infrastructure and local government in the implementation of policy, which also aggravates inequity between both areas (DOST, 2024b).

As artificial intelligence becomes more integrated into governance and education, the development and enforcement of ethical frameworks and policies are increasingly critical to ensure responsible and equitable AI adoption (Taeihagh, 2021). Recent research at ELJ Memorial College has shown that students' awareness of global initiatives, such as the Sustainable Development Goals (SDGs), is shaped by both institutional context and discipline, highlighting the importance of localized studies in understanding educational challenges and ethical perspectives (Rivera, 2024). Thus, there is a need for more localized and discipline-specific empirical research to help inspire AI-integrating efforts and successful approaches.

Grounded in the Ethical Risk Mitigation Framework ([4]; Temper et al., 2025b), this highlighted five primary risks — bias, privacy, transparency gaps, over-reliance, and equity — and proposed mitigatory strategies such as algorithmic audits and human oversight. Based on the investigation of teacher education and business education students, the study examines the relationship between discipline-driven training and ethical reasoning/policy efficacy. The business education students may overestimate the efficiency AI brings to the analysis. In contrast, specialized students who have already chosen teacher education may exaggerate the value of AI-based approaches versus student-teachers' interactions (Zhou et al., 2024b). Such insights are crucial to make sure sound guidelines are available that align with the long-term perspective of the world, responding to the intricacy of actual institutions.

To guide this investigation, the study addresses the following research questions: How do students in teacher and business education programs perceive the ethical risks associated with AI use in academic integrity, specifically in terms of bias and discrimination, privacy and security, transparency, over-reliance, and equity? Are there significant differences between the ethical perceptions of teacher and business education students regarding these AI-related risks? By explicitly exploring these research questions, the study aimed to clarify its focus and provide a structured framework for analyzing discipline-specific perspectives on the ethical implications of AI in higher education.

Utilizing an exploratory mixed-methods framework, this study examined the ethical perceptions of students in higher education at ELJ Memorial College (ELJMC) and the potential for institutional policies to attenuate risks associated with AI. These findings are intended to provide actionable strategies for responsible AI adoption, such as using transparency, equity, and interdisciplinary collaboration (7). The study aims to promote academic honesty and ensure that AI is a beneficial tool for inclusive and equitable education by connecting global ethical concerns with discipline-specific considerations.

## **LITERATURE REVIEW ON ETHICAL RISKS OF AI IN EDUCATION ALGORITHMIC BIAS AND DISCRIMINATION**

Given the potential for biased training data and black box decision making, there is a strong possibility that education AI systems pass on systemic biases. Research globally highlights how voice recognition tools may not be able to recognize regional dialects, and how automated proctoring systems, which are often used in standardized testing, disproportionately penalize marginalized students (3). For instance, Yoder-Himes et al. (2022) found that students with darker skin tones, especially Black women, were much more likely to be flagged in AI proctoring software as potentially cheating, thereby deepening existing inequities in Science, Technology, Engineering, and Mathematics (STEM) education. Correspondingly, the study by Májovský et al. (2023) demonstrated that fake academic papers could be constructed through AI machines, providing believable, yet invalid references, highlighting concerns over integrity and research calls for mechanisms for verification. Some examples are how the UPOU (Cañas-Llamas, 2024) in the Philippines requires transparency in the use of AI to prevent bias, and how Far Eastern University [FEU] (2023) requires students to verify the AI-generated content so that discriminatory results would be avoided.

Separate contributors to AI bias include data bias from non-representational datasets and algorithmic bias from imperfectly designed algorithms (Chapman University, 2025), further complicating these challenges. For instance, Marian University Library (2023) explains how data influenced by the biases of the society in which it was created can cause systems of artificial intelligence to learn biased correlations, such as associating genders with certain professions. Addressing them requires technical solutions such as bias detection tools and

pedagogical reform to ensure equitable access and human oversight in AI-driven education systems.

### **Privacy and Data Security**

AI's dependency on large repositories of student data can usher in dire privacy risks, especially as institutions come to embrace AI-powered tools that promise personalized learning and efficiency on the administrative side. Finally, it should be noted that the Family Educational Rights and Privacy Act (FERPA) is a fundamental federal law that controls access to student data, but requires robust data governance frameworks (3). For example, Zawacki-Richter et al. (2025) propose a three-step process for higher education institutions to ensure compliance with FERPA, which includes discovering and classifying student data, monitoring risks, and remediating vulnerabilities using AI-enabled tools. The UPOU implements strict data anonymization and consent protocols (15), and Mindanao State University (MSU, 2024) includes accountability for breaches in its AI policy. However, few on-campus systems at these rural institutions exist to safeguard data, which creates inequities in access and security (8). These differences emphasize the need for standardized protocols to protect sensitive information from different educational environments.

AI systems that must process a lot of personal data with little transparency exacerbate data privacy challenges. As AI systems ingest more data about students, the challenges of meeting the de-identification requirements of FERPA become more difficult, and the re-identification risks greater (Sexton & Vance, 2024). In addition, Myla (2025) indicates that there is a crucial need to establish AI-powered data security solutions, including real-time threat detection and automated compliance audits, to comply with international standards. Without safeguards, schools would run afoul of privacy laws and undermine trust among students and families, they said. It follows that AI literacy programs and transparent data governance frameworks are needed measures to actively prevent the amplification of education inequality, while utilizing the potential of AI to render education fairly.

### **Transparency and Accountability**

A lot of AI systems in production are black boxes, making it impossible for any stakeholder to understand how the decision was reached, which makes transparency very difficult to achieve. As the development of Explainable AI (XAI) is international in scope and not confined to any single policy report, there is a clear need to embed explainability scores in AI tools and demand accountability to ensure alignment with educational goals. Recent systematic reviews and empirical studies emphasize the importance of transparency, trustworthiness, and accountability in educational AI systems by clearly explaining AI decision-making processes and evaluating explainability metrics to support stakeholders' understanding and confidence (Li et al., 2024; Holmes et al., 2024). Algorithmic bias and trust are interwoven with technology, law, and ethics (Singhal et al., 2024). As per the Commission on Higher Education (CHED) in the Philippines, human-mediated AI-based assessments should be utilized in an outcomes-based system (Teehankee, 2024). The remaining policies come from the MSU, which requires transparency reporting of AI tools, and the UPOU (15), which puts educators through the paces of auditing the output of algorithms and getting in compliance with guidelines for use.

This challenge is being addressed globally, such as through the AI competence framework, which is mainly focused on enhancing transparency and ethical practices in education. For example, the advice in one of the similar ways somehow that the minimum requirements for educators and students' literately in AI are introduced to inform AI users critically to better on the ways they interact with AI systems so as at least reduce over dependence on the systems but same time holding them accountable for their action. Crucial to guard against potential risks of bias in grading or biasedness in admission decisions, the importance of these measures, particularly in the case of language and/or cultural diversity of the test takers, cannot be underestimated. A combination of technical mitigations and policy reform can empower institutions to strike the correct balance between innovation and equity, enabling AI to be a force for equity in education rather than another way that inequity is perpetuated.

### **Academic Integrity and Misuse**

Generative AI tools, such as ChatGPT disrupt traditional integrity norms by overlapping original and AI-

generated work. As a study has shown (Currie et al., 2023), although ChatGPT may provide decent help with basic tasks, when applied to more complex medical education assignments, it generates errors that are seriously worthy of concern for anyone relying on AI knowledge for specialization. Similarly, according to Bin-Nashwan et al. (2023), students increasingly use ChatGPT to write essays, which can easily lead to academic dishonesty and decreased critical thinking. To illustrate, FEU in the Philippines prohibits the undisclosed and unapproved use of AI (16), while the UPOU (15) asks students to cite AI-generated content, where applicable, to ensure transparency. Detection tools, like Turnitin, spend their energies sifting through false positives, resulting in false accusations and tarnishing trust in assessment systems (Baker & Smith, 2024). For example, Eke (2023) has disclosed that even the most advanced detectors would classify human-AI collaborative work as generated by AI, making enforcement of integrity policies all the more complicated.

These challenges are compounded by the rapid evolution of AI tools, which outstrip institutional policy updates. Rejeb et al. (2024) indicated that 67% of educators do not know how to respond to the misuse of AI, and Kovari (2025) argued for AI-resistant assessments, for example, oral examinations or personal projects aimed at reducing the risks of plagiarism. However, according to Baker and Smith (2024), up to 30% of students can beat detection tools through paraphrasing, showing that current solutions are inadequate. Together, they will serve not only to keep academic challenges at bay but also to ensure that innovation and academic integrity go hand in hand in the ChatGPT era.

### **Equity and Access**

The gap in access means the existence of a digital divide in the context of technology and infrastructure. The scouting by AI will not help marginal communities. U.S. Department of Education (2023) echoes the need to advance equity in access, synchronization support, subsidies, and service workers to fill profound inequities in mountainous poor areas by increasing access to endless opportunities of AI. In the Philippines, the DOST is tackling gender inequities by upskilling women in AI literacy (8). However, rural schools still are not benefiting from AI's transformative power; data and resources limit connectivity (20).

For example, AI is both a solution to the limitations of remote education, such as through personalized learning platforms, and at the same time highlights the necessity for internet access, which remains limited in many areas, for instance, only about 37% of rural Philippine schools have reliable internet connectivity (Philippine Institute for Development Studies [PIDS], 2024; Department of Information and Communications Technology [DICT], 2025). In the diverse scope, it warns about the potential risk of discriminatory AI tools exacerbating existing inequalities, especially in linguistically diverse regions where translation tools hardly account for local dialects (Organization for Economic Cooperation and Development [OECD], 2024a).

Global initiatives stress that they must adopt inclusive AI policies and invest in infrastructure to mitigate these risks. An EDUCAUSE AI Landscape Study 2025 found that faculty training for equitable AI use is a priority for 63% of institutions (Robert & McCormack, 2025). However, only 2% of the institutions can gain new funding for such training; in short, there is a systemic underinvestment. UN Women Regional Office for Asia and the Pacific (2025) has also introduced its AI School, designed to ensure stakeholders can promote the need and rights for gender equality advocacy by recognizing and combating algorithmic bias, and ensure their advocacy for inclusive AI design. These actions are in response to the call for equity audits to illuminate and address AI-induced inequities in student outcomes (Williamson & Eynon, 2024). If we do not take coordinated, systemic efforts to make access, literacy, and culturally relevant artificial intelligence tools available, the digital divide will continue, and rural and marginalized communities will be disadvantaged three times.

### **Governance and Policy Frameworks**

International frameworks such as UNESCO's ethical guidelines rest on human rights, sustainability, and ensuring that AI is a companion to humans instead of a replacement. UNESCO encourages the integration of artificial intelligence into all areas of education, whether through competency frameworks for students and educators or by encouraging ethical design principles and universal access to AI tools (2). This aligns with the 2021 Recommendations on the Ethics of Artificial Intelligence and the 2019 Beijing Consensus, which highlight the importance of inclusion and cultural diversity in AI applications. In the local context, the CHED is already



deploying an outcomes-based education framework in the Philippines that aligns the use of AI with emerging competencies workers need to address real-world issues (24). In conjunction, MSU takes it a step further. It promotes the formation of ethics committees for the deployment of AI, and aims for transparency, fairness, privacy, and accountability in its system-wide policies (20).

In its National AI Roadmap, the Philippines is even more explicit about including gender-inclusive policies and establishing equitable access to AI technologies as part of its ambition to harness AI for national development (7). International initiatives, such as UNESCO's designation of the International Day of Education 2025, focus on the essential discourse of responsible integration of AI in schools worldwide (37). At the same time, Michigan State University indicates that embedding ethics in generative AI initiatives must be a core concern that ensures fairness, inclusivity, and academic integrity (Michigan State University Ethics Institute, 2023). By blending international guidelines with local approaches, nations have the power to tackle systemic challenges in education while making sure that AI acts as a force for empowerment rather than exclusion.

## Synthesis

Global and local frameworks showcase the opportunities and challenges surrounding students' learning with artificial intelligence (AI) in education. Although AI can enhance learning experiences, improve administrative tasks, and promote equity in education, its adoption entails serious ethical considerations. These include algorithmic bias, data privacy risks, transparency gaps, and threats to academic integrity. Global efforts highlight the importance of developing explainable AI systems, strong data governance, and inclusive policies that jointly address these risks. In the local context, institutions in the Philippines have already started to tackle these problems through ethics committees, AI literacy programs, and gender-inclusive policies. However, the breakneck pace of evolution in AI tools such as ChatGPT is already ahead of institutional policies, leading to gaps in enforcement and an ongoing debate over the extent to which innovation needs to be balanced with fairness and accountability for students within education.

Overall, several gaps in the research remain despite these efforts. First, the perceptions of students from various disciplines regarding the ethical considerations of AI use in education are not well understood. Most existing studies examine general trends or particular tools but do not address discipline-specific nuances that shape attitudes about AI-driven emerging technologies. Second, although there are global frameworks that serve as high-level inputs for ethical integration of AI technologies, the researchers did not find localized studies to identify the implementation of these principles. Third, the literature does not adequately explore the effectiveness of institutional policies to mitigate risks such as academic dishonesty or disproportionate access to AI tools, which may vary by context. Such gaps inform the necessity of further exploration into how students from diverse fields grapple with and understand the ethical implications of AI use within their academic spaces.

This study fills the aforementioned gaps by investigating the perspectives of teacher education and business studies students regarding the ethical perception of AI use in academic integrity at ELJMC. The research evaluated attitudes towards AI technologies by contrasting two diverse disciplines, as this should shed light on the influence of pedagogic practice or beliefs and professional training on attitudes toward AI technologies.

## Theoretical Framework

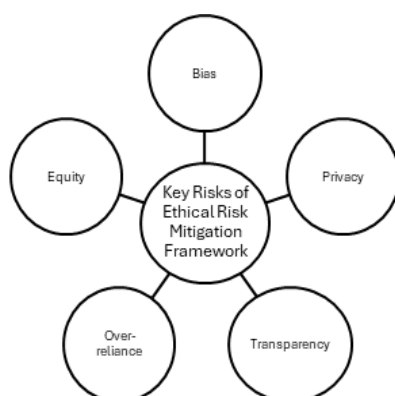


Figure 1. The Key Risks of Ethical Risk Mitigation Framework

The Ethical Risk Mitigation Framework provides a structured approach in addressing ethical risks of AI in education, focusing on bias, privacy, transparency, over-reliance, and equity. More specifically, that ERMF (4)(11) lists five key risks: bias and discrimination, privacy and security, the gap in transparency, dependency in AI, and equity and access. It raises the possibility that there are ways to mitigate the risks of algorithmic capture of AI in students, such as algorithmic audits, human supervision, and participatory educational policy processes that steer AI in line with educational goals. This ever-increasing view underlines the necessity of a Human Rights-Based Approach (HRBA) to equitable and sustainable ethical AI to encourage educators, policymakers, and technologists to work together (3).

This framework resonates with the present study examining the ethical perceptions of utilizing AI services in the context of teacher education and business studies. By exploring discipline-specific attitudes when it comes to risks such as algorithmic bias or data privacy, the study also takes up the framework's focus on contextual equity—ensuring that AI in tools can be contextualized to suit the varied natures of educational paths, so as not to perpetuate systemic disparities.

This framework fills an important research gap by connecting the global consensus on specific ethical standards with their local implementation challenges. The guidelines are provided by organizations such as UNESCO (2024) and the U.S. Department of Education, Office of Educational Technology (2023). AI integration in higher education is broad, and thus, this study aims to explore the unique interpretation of the principles in the context of higher education in the Philippines, where resource inequalities and cultural factors impact the adoption of AI. Using the framework for risk-mitigation strategies, including bias audits and equitable resource allocation, the study recommends tailored solutions for ethical AI literacy and governance in teacher and business education programs.

### **Statement of the Problem**

This study sought to answer questions by exploring the ethical perceptions of AI use in academic integrity among teacher education and business studies students at ELJMC, Philippines. Specifically, it aimed to investigate:

1. How do teacher education students perceive ethical risks of AI in education regarding bias and discrimination, privacy and security, transparency gaps, over-reliance on AI, and equity and access?
2. How do business education students perceive ethical risks of AI in education regarding bias and discrimination, privacy and security, transparency gaps, over-reliance on AI, and equity and access?
3. How do the ethical perceptions of Teacher Education students differ from those of Business Education students across the five dimensions?
4. What are the ethical perceptions of students regarding the use of AI in relation to academic integrity at ELJ Memorial College, particularly in terms of bias, privacy, transparency, over-reliance, and access?
5. How may the findings of this study be used in crafting an institutional policy on AI use?

## **METHODOLOGY**

### **Research Design**

An exploratory mixed-methods framework was used in this study, integrating quantitative survey and qualitative data to provide a comprehensive understanding of ethical perceptions to investigate ethical attitudes towards AI use on academic integrity of teacher education (TE) and business education (BE) students at ELJMC, Philippines.

The first phase of the study involved a quantitative survey conducted via stratified random samples of students from TE and BE using Likert scale instruments based on perceptions across five ethical dimensions. Descriptive statistics (median and interquartile range) and the Mann-Whitney U test were conducted across disciplines. Second, in-depth, semi-structured interviews with 20 discipline-representatives, who would make sense of the data from the survey, framed quantitative results, developing how pedagogical training or institutional policies impact ethical reasoning (Creswell & Creswell, 2023).

This mixed-methods methodology adopted the Ethical Risk Mitigation Framework (4), guided by core precepts, including transparency, accountability, and equity in the application of AI. The quantitative part traces trends in ethical risk perceptions through the years, presenting a superficial picture of students' attitudes. To this end, a qualitative method was used to situate students' lived perceptions and experiences by providing context-rich insights of how students make sense of and negotiate ethical issues connected to AI use in their academic disciplines. For example, a series of interviews sheds light on how students make sense of institutional policies and place faith in AI governance, which cannot be fully known from survey data (1). By cross-validating survey results with interview data, the method ensures the credibility of the results. It seeks to provide a broad insight into ethical problems, generalized concepts, and provide insight into specific situations.

Inter-coder reliability was established during the qualitative analysis to reduce possible bias further, and the list of census-validated survey instruments utilized was based on previous AI ethics research works (12). The participants were selected through stratified random sampling from the TE and BE departments. Ethical clearance was secured by getting consent, and the data was anonymized. The findings might inform the recommendations for AI policy by discipline, faculty training, and support allocation to ensure fair and equitable approaches and bridge global ethical imperatives and calls for action and local educational settings (2).

### Sample and Sampling Procedure

The sampling process involved three steps: first, secured a complete list of TE and BE students through the registrar's office; second, randomly selected participants within each stratum by way of a random number generator (randomizer website); and third, the sample included a 50% replacement sample to account for nonresponse. This means that all the subjects were eligible to be chosen and that the final sample was representative of the population.

Proportional representation of Teacher Education (TE) and Business Education (BE) students in ELJMC, combined with stratified random sampling, was employed. The sampling frame of the study consisted of all BE and TE students who had been admitted in the second semester of the college in the year 2024–2025 through the registrar's office. Stratification was arranged according to discipline, resulting in two strata: one for TE students (College of Teacher Education, CTE) and the other for BE students (College of Business Education, CBE). This approach reduces selection bias and ensures that subgroups are adequately reflected in the sample, an ethical consideration recommended for AI research (6).

For calculating the sample size to take from each stratum, the sample size was distributed proportionally according to the number of students enrolled in each department. The CTE students comprised 44% of the eligible population, and CBE students 56%. Correspondingly, 44% of the sample was selected from CTE and 56% from CBE. Sample size was determined by carrying out a power analysis for a Spearman's rho correlation and was 138 for  $\alpha = 0.05$ , power = 0.95, and effect size = .30 (39) through G\*power. For potential non-response, an added buffer of 50% was applied. Initially, there were 92 respondents from CTE and 118 from CBE for a total of 210. Following the distribution of the questionnaires, 88 students from CTE and 95 from CBE completed and returned them, resulting in a total population size of 183, surpassing the minimum required. Furthermore, 20 respondents were randomly selected from the survey participants for follow-up interviews to provide further depth to the qualitative component of the study.

TABLE I. Demographic Profile of the Respondents

Demographic Variable	Category	CTE Frequency	CTE Percentage (%)	CBE Frequency	CBE Percentage (%)
Sex	Male	18	20.5	22	23.2
	Female	70	79.5	73	76.8
	Total	88	100	95	100
Age	18 - 20	67	76.1	56	59.0

	21 - 23	19	21.6	38	40.0
	24 - 26	2	2.3	1	1.0
	Total	88	100	95	100

The demographic profile of the respondents shown in Table I reveals a predominance of female students in both the College of Teacher Education (CTE) and the College of Business Education (CBE). In the CTE group, females constitute 79.5% of the sample, while males represent 20.5%. This gender ratio is reduced in the CBE group with 76.8% females versus 23.2% males. More generally, such an unequal gender distribution in these fields aligns with the general developments of attendance in these courses. More women attend education, while business schools are typically at gender parity. However, in this case, women are prevalent. The knowledge of this gender balance is of interest since gender can impact experiences, concerns, and ethical views towards AI.

Regarding age distribution, most respondents in both groups are ages 18–23, comprising the typical undergraduate's age range. In the CTE sample, 76.1% are 18–20 years old, and 21.6% are 21–23 years old, with fractionally smaller proportions in older age brackets. Likewise, in the CBE group, 59.0% are in the 18–20 age group, and 40.0% belong to the 21–23 age group. The small number of older students (24 and above years old) participants in both groups represents some diversity in age among students, which might reflect different attitudes toward the use of AI and academic integrity. Overall, the demographic profile reflects a relatively young and overwhelmingly female sample.

### Data Gathering Instrument

Quantitative data were collected with a structured self-report questionnaire named "Questionnaire on Ethical Perceptions of AI in Higher Education". This instrument was modified and extended by the authors using the work of Zhou et al. (2024b), ensuring that ELJMC students would find the same relevance. The questions were developed to measure students' perceptions of AI in education across the five significant ethical dimensions identified in the literature and the Ethical Risk Mitigation Framework.

Part A of the questionnaire is the Demographic Profile of the respondents. This section gathers basic demographic information about the respondents, including their (Q1) college program (Teacher Education or Business Education), (Q2) sex (with inclusive options such as "Prefer not to say"), and (Q3) age. These variables would help contextualize the responses and allow for subgroup analysis based on demographic factors.

Ethical Perceptions is the coverage of the second part of the questionnaire, Part B. The questionnaire was adapted from the work of Zhou et al. (2024b). This section uses a 4-point Likert scale (1 = Strongly Disagree, 2 = Disagree, 3 = Agree, 4 = Strongly Agree) to measure respondents' agreement with statements related to the five ethical dimensions. Ethical perception covers five parts. First is the bias and discrimination, which focuses on algorithmic bias in grading or admissions (Q4-Q6). Second, the aspect of privacy and security involves worries about data collection and breach (Q7-Q9). Third, the transparency gaps (Q10-Q12), where students check if they actually get the reasoning for how AI makes decisions. Fourth, the dependence on AI examines whether students perceive the dependence on AI to be detrimental to critical thinking or human oversight. Finally, the equity and access (Q13-Q15) scrutinize the gap in access to AI tool resources between urban and rural students.

The questionnaire was created according to the Ethical Risk Mitigation Framework (4) and seeks to align with global standards of ethics and the relevant educational situation. It follows ethical research guidelines by indicating that users' participations were voluntary, that their responses would be anonymous, that they had the right to withdraw from the study at any point, and that they provided consent by participating. The tabulated structure of the instrument facilitates the administration and analysis of the quantitative and qualitative data.

The validity of the questionnaire was analyzed by content and face validity. Content validity was established by comparing the instrument to the Ethical Risk Mitigation Framework (4). Participant comments in the dry run were noted, such as some of the terms were ambiguous. Some questions in the survey were rewritten to make them more accessible for students. As a result of these modifications, it was confirmed that all subjects,



regardless of their AI knowledge, could understand the questions and answer them correctly. These alterations enhance the sensitivity of the questionnaire for subtle ethical sensitivities and maintain consistency with theory.

The qualitative data of the study were collected through semi-structured interviews that were based on open-ended questions to investigate students' perceptions and ethical issues concerning the use of AI for educational purposes. These questions were informed by a literature review and were adapted to align with the purpose of the study so that rich and meaningful descriptions would be received. The semi-structured nature of the interview allowed the interviewer to follow up on the participants' answers while ensuring that interviews followed a consistent pattern.

Meetings were conducted in person, after an agreement, and recorded and transcribed word for word. This method allowed for the development of a detailed, rich description of student attitudes and experiences, and articulated a nuanced understanding of the ethical and social implications of AI integration into academic settings, which would be challenging to find by relying on quantitative approaches alone.

## DATA ANALYSIS

Descriptive statistics were used to summarize the ethical perceptions of teacher education students based on AI risks for the first and second research questions. Mean and was computed for the five dimensions: bias and discrimination, privacy and security, transparency gaps, overreliance on AI, and equity and access, based on responses of teacher education students. This method yielded an overall picture of the central location of the perceptions among this group, with the ability to discern the subtleties of how teacher-education students perceive the ethical challenges of implementing AI in education (38).

The table below was used to interpret the weighted mean for each construct:

TABLE II. Verbal Description for the Ethical Perceptions

Scale Point	Range	Interpretation
1.00 – 1.74	Strongly Disagree	Very low ethical perception
1.75 – 2.49	Disagree	Low ethical perception
2.50 – 3.24	Agree	High ethical perception
3.25 – 4.00	Strongly Agree	Very high ethical perception

The Mann-Whitney U test was utilized to determine whether there were significant differences in ethical perceptions between teacher and business education students across the five dimensions. This nonparametric test was chosen after normality tests indicated that the data did not meet the assumptions required for parametric testing. Each dimension score distribution between the two independent groups was compared using the Mann-Whitney U test, and U statistics and p-values were provided to indicate statistical significance. This method enabled the drawing of sophisticated inferences about group differences in ethical risk perceptions even when the data were non-normal (Laerd Statistics, 2024).

For research question 3, the qualitative thematic analysis method was used to progressively search for, analyze, and interpret themes across the interview data. This process, of creating themes that identify and document the common or recurring thoughts and concerns related to bias, privacy, transparency, over-reliance, fairness and equity, opened up for the researchers a systematic opportunity to think about the ethical aspect of AI in education, in ways they would not otherwise have. Thematic analysis is flexible and can be used inductively to permit themes to be derived directly from the data, or deductively; to determine themes based on past theories or research questions (Braun & Clarke, 2006, as cited in Dawadi, 2020). The use of analytic protocols through six thorough steps, such as getting familiar with the data, coding, theme development, refinement, transparency, and validity in the analysis (Nowell et al., 2017 as cited in Dawadi, 2020). This approach is particularly

appropriate for educational research since it enables the study to gain a rich and deep understanding of the participants' experiences and the mental world views of the phenomenon, an aspect that quantitative methods may fail to reveal (Phan, 2025). Using thematic analysis, the project will generate fine-grained insights into ethical issues, which would be useful for policy recommendations relating to the deployment of AI in education.

### Ethical Considerations

This study follows proven ethical standards to protect all participants' rights, safety, and ethical safeguards. It is noticed that participation in the research is based on the consent of the individuals who were clearly informed that they are not obligated to participate in the research or answer the questions. They were also informed that they could participate and withdraw at any time without any negative consequence (American Psychological Association [APA], 2020). This principle of non-coercion means that respondents were placed under no pressure or duress to participate; it is, however, a requirement in all international ethical guidelines for human research (Penn LPS Online, 2024).

All responses were anonymized to keep data confidential during the study. No personally identifiable information was collected, and the data were kept secured in conformance with institutional regulations. This protects the privacy of respondents and prevents their data from being associated with their person, which, in turn, facilitates trust and honest answers. Anonymity also provides confidentiality because researchers cannot link participants to their responses (APA, 2023). Participants were also made aware that they have the right to stop the study at any time without needing to give a reason. This right was explicitly set out during the questionnaire's introduction and data collection to empower participants to make decisions about their participation at any point.

Informed consent was considered as given if subjects participated in the study. Respondents received a detailed explanation about the purpose of the study, the method of the study, and the intended use of their data. Participants indicated directly on the questionnaire their agreement regarding these details and their decision to participate (42). Finally, this study was evaluated and approved by the review committee of ELJMC. This endorsement confirms that the research is conducted by an ethical standard for human studies, including, but not limited to, minimizing the risk of harm, respecting autonomy, and treating people fairly. This article contains no studies with human participants or animals performed by any of the authors.

### Validation of the Instrument

Reliability coefficients (Cronbach alphas) for five constructs for the dry run of the questionnaire are presented in Table II for 38 respondents. This is important because these coefficients indicate the degree of hidden consistency, which is implicit in the scales employed to assess each dimension, offering a frame of reference to assess the consistency of the tools.

TABLE III Cronbach's Alpha for the Five Constructs during the Dry Run

Construct	Cronbach's Alpha	Interpretation
Bias and Discrimination	0.84	Good Internal consistency
Privacy and Security	0.87	Excellent internal consistency
Transparency	0.85	Good internal consistency
Over-reliance on AI	0.81	Good internal consistency
Equity and Access	0.86	Good Internal consistency

The reliability coefficient of the Bias and Discrimination structure was 0.84. This suggests that the items created to measure bias and discrimination are stable and reliable. For privacy and security, the developed scale showed good internal consistency, with values of 0.87, indicating that the scale is suitable for measuring concerns related to privacy and security that are becoming more prominent in technology-mediated education.

The Transparency Gap construct showed a reliability of 0.85, indicating a notable improvement in internal consistency. This difference may reflect a greater emphasis on transparency or a more cohesive understanding of the concept in competency-based educational environments. Over-reliance on AI exhibited good reliability, with CTE at 0.81. This high value suggests that the items measuring concerns about over-reliance on AI are remarkably consistent, possibly due to more uniform experiences with AI integration in these settings.

Lastly, the Equity and Access factor demonstrated excellent reliability, alpha at 0.86 and is supported by very good internal reliability. This indicates that constructs of equity and access are well articulated and widely accepted among respondents, making it a reliable base for further investigation and intervention.

In general, the results of the reliability analysis indicate that the five constructs have satisfactory acceptable levels of reliability, with one of the constructs at an excellent level.

A committee of 10 college experts with expertise in Research, English, and Information Technology was utilized to evaluate the content and face validity of the instrument. These experts responded with valuable critiques on pertinence, clarity, and calibration of the items, guaranteeing that the instrument captures the constructs for ethical perceptions around AI use in AI. This method is consistent with the best practices for instrument validation in that professional judgment is needed for the determination of its content validity and the refinement of the instrument for the intended population.

For quantitative content validity assessment, this study used the Face Validity Index (FVI) according to Yusoff (2019) and Marzuki et al. (2023). The FVI offers experts a structured way to score relevance and clarity on each item, and aggregated scores provide a quantitative indication of content validity. This procedure complements the expert qualitative feedback with quantifiable evidence that the instrument items reflect the intended constructs.

TABLE IV. Construct Validity Index for Questionnaire Items

Item No.	A	I-CVI	Pc	k*	Interpretation
4	8	0.8	0.044	0.79	I-CVI: Acceptable; Pc: Not by chance; k*: Excellent agreement
5	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
6	8	0.8	0.044	0.79	I-CVI: Acceptable; Pc: Not by chance; k*: Excellent agreement
7	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
8	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
9	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
10	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
11	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
12	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
13	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
14	8	0.8	0.044	0.79	I-CVI: Acceptable; Pc: Not by chance; k*: Excellent agreement
15	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
16	8	0.8	0.044	0.79	I-CVI: Acceptable; Pc: Not by chance; k*: Excellent agreement
17	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
18	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
19	8	0.8	0.044	0.79	I-CVI: Acceptable; Pc: Not by chance; k*: Excellent agreement

20	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
21	8	0.8	0.044	0.79	I-CVI: Acceptable; Pc: Not by chance; k*: Excellent agreement
22	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
23	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
24	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
25	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
26	10	1.0	0.001	1.00	I-CVI: Perfect; Pc: Not by chance; k*: Perfect agreement
27	9	0.9	0.010	0.90	I-CVI: High; Pc: Not by chance; k*: Excellent agreement
28	8	0.8	0.044	0.79	I-CVI: Acceptable; Pc: Not by chance; k*: Excellent agreement

The probability of chance agreement (Pc) values in the table ranges from 0.001 to 0.044, indicating that the observed content validity indices (I-CVI) and kappa statistics (k\*) are not attributable to random agreement among experts. Specifically, the very low Pc values (close to zero) for items with perfect I-CVI (1.0) and k\* values of 1.00 confirm that expert consensus is robust and statistically significant. Items with slightly higher Pc values (0.010 to 0.044) still demonstrate strong agreement beyond chance, corresponding to acceptable to high I-CVI and excellent k\* agreement (0.79–0.90). These findings collectively validate the instrument's content, suggesting that the items are well-constructed and reliably reflect the intended constructs. This aligns with recent methodological recommendations emphasizing the importance of adjusting for chance agreement to avoid overestimating content validity and ensuring rigorous instrument development (50).

The construct validity was measured by using the Item-level Content Validity Ratio (I-CVR), which determines the consensus of experts on the item's essentiality (Anuar et al., 2024). To avoid random agreement among experts, the obtained CVR values were compared to the critical values corresponding to the number of assessors, and the minimum acceptable CVR for 10 experts is 0.62 (48). The kappa statistic (K\*) was also calculated to offer more reliable ratings on item relevance (Polit & Beck, 2024). Such strict progressions ensured that the instrument items together perfectly represent the theoretical constructs of ethical perceptions towards AI usage, consequently establishing the construct validity.

Cronbach's alpha coefficient was used to test the internal consistency and reliability of the instrument. This statistic indicates the extent of average interrelationship of the set of items as a whole, showing scale consistency of the instrument in measuring the constructs. A high value of Cronbach's alpha indicates that the items are strongly correlated, and the instrument is reliable for the population of interest. This is a well-established method in survey-oriented research and was successfully employed in the current study to validate the instrument's reliability.

Qualitative review by experts was used in instrument validation, which was combined with quantitative indices (FVI and I-CVR) and reliability value (Cronbach's alpha statistics), ensuring that this is a solid evidence of validity and reliability data for the instrument. This rigorous validation confirms the theoretical underpinning and practical utility of the instrument developed in this study for measuring students' ethical perceptions of AI in academic integrity.

The k\* values that adjust for chance agreement are between 0.79 and 1.0, representing excellent to perfect agreement among experts. If items have k\* values around 1.0, then that would suggest perfect consensus; near 0.79, items will indicate strong agreement (McHugh, 2021). It is indicated that the content validity is acceptable, high, or perfect. The agreement achieved is excellent or perfect. Generally, this table illustrates a strong validation of the instrument and indicates that the items are well-constructed and reliable for their intended use.

Table V lists FVI for individual items, which indicate how well individual items seem to measure the associated construct from the subjective judgment. The scores of the FVI items are between 0.8 and 0.9, and the ones scoring 0.9 are treated as acceptable face validity, so that these items are obviously relevant and understandable



to the evaluators or respondents at face value. The items with an FVI value of 0.8 are rated to have marginal face validity, which implies that the items are generally acceptable. However, small changes in terms of clarity or relevance may be made in a few of them. Overall, most of the items showed good face validity and can be confidently retained for inclusion in the measure; however, a few of them, with marginal ratings, were reworked to improve the extent to which they are immediately understood and relevant.

TABLE V. Face Validity Index for Questionnaire Items

Item Number	FVI	Interpretation
4	0.9	Acceptable face validity
5	0.8	Marginal face validity
6	0.8	Marginal face validity
7	0.9	Acceptable face validity
8	0.9	Acceptable face validity
9	0.8	Marginal face validity
10	0.9	Acceptable face validity
11	0.9	Acceptable face validity
12	0.8	Marginal face validity
13	0.8	Marginal face validity
14	0.8	Marginal face validity
15	0.9	Acceptable face validity
16	0.9	Acceptable face validity
17	0.8	Marginal face validity
18	0.9	Acceptable face validity
19	0.9	Acceptable face validity
20	0.9	Acceptable face validity
21	0.9	Acceptable face validity
22	0.9	Acceptable face validity
23	0.9	Acceptable face validity
24	0.9	Acceptable face validity
25	0.9	Acceptable face validity
26	0.9	Acceptable face validity
27	0.9	Acceptable face validity
28	0.9	Acceptable face validity

## RESULTS AND DISCUSSION

### Internal Reliability

TABLE VI. CTE Cronbach's Alpha and Exploratory Factor Analysis for the Five Construct

Construct	Cronbach's Alpha	Interpretation	EFA	Interpretation
Bias and Discrimination	0.82	Good	1	Unidimensional

Privacy and Security	0.85	Good	1	Unidimensional
Transparency	0.81	Good	1	Unidimensional
Over-reliance on AI	0.80	Good	1	Unidimensional
Equity and Access	0.84	Good	1	Unidimensional

Table VI displays the results of the reliability and factor structure analysis for the five constructs of AI ethics and usage. The Cronbach's alpha of all the constructs varies between 0.80 and 0.85, which implies satisfactory internal consistency reliability. These are indications that the items of the constructs are strongly related to each other and consistently measure the theoretical variable, and in turn, this could justify our reliance on the scale to measure the related constructs in research and practice. Such strong reliability coefficients are essential for ensuring that the constructs are measured accurately and that the results are dependable.

The exploratory factor analysis (EFA) results show that each construct is unidimensional, with a factor loading of 1, indicating that all items within each construct load onto a single factor. This one-dimensionality attests that each item set assesses one coherent latent variable without much intrusion of other dimensions (Brown, 2022). Taken together, high reliability and unidimensional factor structure support the validity of these constructs as separate and homogenous dimensions to examine constructs. This strong psychometric support underlines the instrument's overall usefulness for academic research and applied purposes.

TABLE VII. CBE Cronbach's Alpha and Exploratory Factor Analysis for the Five Construct

Construct	Cronbach's Alpha	Interpretation	EFA	Factor Structure Interpretation
Bias and Discrimination	0.83	Good	1	Unidimensional
Privacy and Security	0.85	Good	1	Unidimensional
Transparency	0.81	Good	1	Unidimensional
Over-reliance on AI	0.80	Good	1	Unidimensional
Equity and Access	0.84	Good	1	Unidimensional

Table VII summarizes the reliability and factor structure of the five main constructs pertinent to AI ethics and governance. All constructs show good internal consistency, with Cronbach's alpha ranging between 0.80 and 0.85. These scores suggest that the items in each construct are closely related and consistently measure the respective concept. The high reliabilities are important in that they facilitate the constructs yield both reliable and consistent results across various sets of subjects, and that these results can be considered credible and robust for other analyses or applications using such measures.

The findings of the results of the EFA also provide some evidence for the psychometric performance of these constructs, as each one is one-dimensional, the components of which measure the same underlying factor. This apparent unidimensionality demonstrates that the items would form a single, underlying dimension, validating conceptual clarity and consistency (52). The clear structure of the factors across constructs confirms the theoretical model of ethical AI. It provides validation for their use as discrete, clearly defined units in assessing the ethical implications of an AI system. The strong reliability and unidimensional structure make this a robust measure for research and applied work in this field.

## AI Perceptions of the Respondents

Table VIII presents the mean scores and verbal interpretations of students' ethical perceptions regarding five key AI risks in academic settings. All constructs received mean scores between 2.69 and 2.98, corresponding to a "high ethical perception" across the board. This suggests that students are very cognizant of the ethical dangers of AI in education, and this falls in line with a heightened awareness of concerns such as algorithmic bias, privacy infringements, lack of transparency, over-reliance on AI, and the continued problem of unequal access.

TABLE VIII. CTE Respondents' Perceptions of Ethical Risks in AI

Ethical Risk	Mean	Verbal Interpretation
Bias and Discrimination	2.69	High ethical perception
Privacy and Security	2.93	High ethical perception
Transparency Gap	2.87	High ethical perception
Over-reliance on AI	2.81	High ethical perception
Equity and Access	2.98	High ethical perception

These results are consistent with emerging literature indicating the importance of attending to ethical risks as AI increasingly integrates into educational practices. AI's uptake in higher education can easily overtake policy formation, and institutions are exposed to ethical breaches. For instance, algorithm bias is still a significant concern given that well-audited AIs can reinforce current social unfair advantages ([3],[17]). Finally, privacy and security concerns are paramount, and appeals for comprehensive data and privacy protections, and compliance with laws such as FERPA and GDPR, to ensure student confidentiality and data security are made ([19],[18]). The high level of perceived transparency gaps emphasized the importance of explainable AI and clear accountability mechanisms, as proposed also by UNESCO (2024) and Lee et al. (2024). Overdependence on AI is a red flag that threatens academic integrity and critical thinking, reinforcing worries about uncontrolled automation becoming corrosive to vital human potentials ([25],[29]). Finally, the near-unanimous acknowledgment of equity and access challenges mirrors the persistent digital divide, particularly in low-resource settings, and the value of inclusive AI policy and infrastructure investment ([33],[34]).

The ubiquitous high ethical perception across all risk domains would imply that students not only know about these risks but are also amenable to following institutions in the attempts to counter these risks. Findings infer early signs that point to the necessity of considering ethical challenges, as AI continues to be incorporated into educational endeavors, complementing recent literature in the field. AI's rise within higher education can outpace policy development, and those institutions face risk as they move their strategies forward. In addition, the results endorse UNESCO's (2024b) recommendation to entrench transparency, accountability, and fairness within AI systems, emphasizing that continuous faculty and student education in AI literacy is crucial to enable all stakeholders to engage with AI technologies critically.

Students' perception of the ethical importance of the five key risks related to AI in educational applications is presented in Table IX, with the mean perceptions ranging from 2.57 to 2.78, all being "high ethical perception." This uniform pattern suggests that students are aware of and concerned about ethical issues with AI, such as bias and discrimination, privacy and security, transparency deficits, overdependence on AI, and equity and access. This is a very important awareness. This means that students are not just passive users of AIs but active subjects capable of evaluating their ethicality, and this is important for promoting a responsible use of AI in academic environments.

TABLE IX. CBE Respondents' Perceptions of Ethical Risks in AI

Ethical Risk	Mean	Verbal Interpretation
Bias and Discrimination	2.57	High ethical perception
Privacy and Security	2.78	High ethical perception
Transparency Gap	2.74	High ethical perception
Over-reliance on AI	2.67	High ethical perception
Equity and Access	2.72	High ethical perception

These findings are consistent with other research that emphasizes the need for students to develop ethical literacies in the wake of AI-infused education. For example, teacher and business education students show a strong concern for the ethical issues of fairness and privacy concerning AI. The US Department of Education (2023) also emphasizes the need for institutions to protect privacy and security as a matter of critical concern. The strongly perceived transparency gaps are consistent with recommendations for AI to be explainable in order to engender trust and accountability (4). Questions about overdependence on AI reflect those of Currie et al. (2023) who argue that too much reliance on the machines can detract from the process of critical thinking and academic honesty simply by being excessively dependent on AI. Finally, surfacing issues of equity and access reifies current discourses on digital divides and inclusive AI policies ([33],[34]).

The consistently high ethical perception in these domains highlights a need for education to provide such comprehensive AI governance frameworks that proactively anticipate these concerns. Institutions will need to value transparency in their AI systems, privacy protections, and efforts to minimize bias so that AI tools serve, rather than threaten, educational equity and quality. Furthermore, including ethical AI literacy within educational curricula can help give students what they need to understand, critique, and ethically engage with AI, contributing to a culture of ethical AI consciousness and responsibility. These initiatives demand the collective involvement of disciplines and stakeholders to address AI's ethical risks successfully.

### Difference Between the Ethical Risk Perception of CTE and CBE Students

Table X contains the results of the median, IQR, and Mann-Whitney U test. The median and interquartile range (IQR) for each construct reveal subtle but meaningful differences between Teacher Education (CTE) and Business Education (CBE) groups. Both groups share a median of 3.00 for Bias and Discrimination, but CTE has a slightly wider IQR (2.00–3.00) compared to CBE (2.00–3.00), indicating similar variability. In Privacy and Security, both groups again have a median of 3.00, but the IQR for CTE (2.00–4.00) is broader than that of CBE (3.00–4.00), suggesting that CBE responses are more consistently positive. Transparency shows the same pattern, with a median of 3.00 for both, but a wider IQR for CTE (2.00–4.00) versus CBE (3.00–4.00). Over-reliance on AI presents identical medians and IQRs (3.00, 2.00–4.00) for both groups, reflecting similar central tendencies and variability. Both groups have a median of 3.00 for Equity and Access, but CBE's IQR (3.00–4.00) is again narrower than CTE's (2.00–4.00), highlighting more concentrated responses among business students. In summary, while the median is the same across groups for all constructs, CBE students' responses tend to be more clustered at the higher end of the scale. In contrast, CTE students show greater variability in their perceptions.

TABLE X. CBE Respondents' Perceptions of Ethical Risks in AI

Construct	Teacher Education Median (IQR)	Business Education Median (IQR)	Mann-Whitney U	p-value	interpretation
Bias and Discrimination	2.4 (2.0-3.0)	2.8 (2.4-3.0)	2417.00	0.0002	Significant difference
Privacy and Security	3.0 (2.6-3.4)	3.2 (3.0-3.8)	2609.00	0.0010	Significant difference
Transparency	3.2 (2.8-3.6)	3.2 (3.0-3.8)	2664.50	0.0020	Significant difference
Over-reliance on AI	2.8 (2.2-3.2)	3.0 (2.6-3.4)	2542.50	0.0006	Significant difference
Equity and Access	3.0 (2.8-3.6)	3.2 (3.0-3.6)	2586.00	0.0008	Significant difference

Furthermore, the Mann-Whitney U Test compares the Ethical Perception of CTE students and CBE students along five different dimensions: Bias and Discrimination, Privacy and Security, Transparency, Over-reliance on AI, and Equity and Access. All ethical dimensions show a very large difference with p-values far below 0.05, which indicates that the two student samples do indeed have meaningful differences in their ethical perceptions



on all measured dimensions.

The significant Mann-Whitney U results provide evidence that contrary ideas may exist between CTE and CBE learners toward the ethical risks related to AI usage in education. This is consistent with previous studies demonstrating that disciplinary background is associated with understanding of ethics and concern about the ethics of AI technologies. For example, pre-service teachers might have a higher awareness about equity and access features as a result of their studies in inclusive pedagogy, as compared with that of business students and data governance and compliance frameworks (34). The persistent significance on all dimensions underscores the multidimensionality of ethical perceptions and the varying ways in which AI risks become salient for students across different educational contexts.

These results demonstrate the necessity of domain-specific ethical AI education. Schools and colleges should incorporate training on these concerns into field-specific curricula, establishing a broader ethical fluency and preparing students to evaluate AI technologies critically in their future professional roles. Furthermore, policymakers and educators should be aware of these perceptual differences when developing AI governance policies/curricula to raise awareness more effectively for a wide range of students. Prioritizing interdisciplinary conversation can help to narrow the gap in understanding and foster a more integrated approach to AI ethics in higher education ([4],[25]).

### **Ethical Perceptions of Students on AI Use in Academic Integrity**

A thematic analysis was conducted on student interview data from teacher and business education programs. The analysis revealed six major themes that characterize students' ethical perceptions and lived experiences regarding over-reliance on AI and its implications for academic integrity.

#### **1. Fairness and Accuracy of AI**

Students expressed mixed views regarding the fairness and accuracy of AI tools in academic settings. While some acknowledged the potential for AI to provide unbiased results, many raised concerns about technical errors, lack of contextual understanding, and the inability of AI to recognize individual effort or learning progress. As one participant noted,

“AI tools in education do not treat all students fairly. There are some instances that AI provides inaccurate information especially in grades computations. We experienced it first-hand that AI provided wrong information which was a disadvantage of using it.”

#### **2. Equity and Access**

A recurring theme was the issue of unequal access to AI tools, often due to disparities in internet connectivity and the availability of premium features. Students observed that those with better technological resources had an advantage, potentially exacerbating existing educational inequalities. For example,

“The only problem is the internet access. Students will be assisted without gender and cultural bias.”  
“Halimbawa na lang yung mga studyante na naka-premium sa AI tools dahil mas malinaw at mas tama ang sagot sa mga estudyante.”

#### **3. Over-Reliance on AI and Impact on Critical Thinking**

Many participants voiced concern that excessive dependence on AI diminished students' critical thinking, creativity, and independent problem-solving skills. Students tended to default to AI-generated answers rather than engaging in deeper analysis or original thought. One participant shared,

“Yes, for sure they always rely on the AI tools, that's why the critical thinking of them becomes low.”  
“Yes, ang mga estudyante po sometimes nalalagay sila sa alanganin by using AI dahil inaasa na lang nila sa AI ang mga bagay na dapat gamitan ng sariling pag-iisip o opinion.”

#### 4. Privacy and Data Security Concerns

Students expressed varying levels of concern about the privacy and security of their data when using AI tools. Some were highly cautious, taking steps to limit data sharing, while others were less concerned if only basic information were involved. One participant stated,

“Very concerned dahil baka ang aking mga personal na data ay magamit sa ibat-ibang pamamaraan.”

“It is alarming, because it is prone to leakage of my personal data.”

#### 5. Transparency and Understanding of AI

There was a strong call for greater transparency and education from institutions regarding how AI systems function and how data is handled. Many students admitted to a limited understanding of AI decision-making processes and advocated for more seminars and information sessions. For instance,

“No, I think institutions can improve transparency about how student or me being student is handled when using AI tools with clear communication about data collection.”

“Yes, our institution should conduct seminars about how to use AI to avoid circumstances.”

#### 6. Human Element in Assessment and Feedback

Students consistently preferred human involvement in grading and feedback, emphasizing the importance of teachers’ judgment and explanations. There was discomfort with the idea of AI-only evaluation, as highlighted by one student:

“I feel disappointed because they are the one who will give the right explanation, not just side of AI itself.”

“AI tools should be just as scaffolding in creating feedback.”

Thematic analysis indicates that while students recognize the efficiency and support that AI can provide, they are wary of its limitations and the risks of over-reliance. Concerns about diminished critical thinking, fairness, equity, privacy, and the loss of human touch in assessment are prominent. These findings underscore the importance of balanced, transparent, and human-centered approaches to AI integration in education to uphold academic integrity and foster independent learning.

### CONCLUSION AND RECOMMENDATION

#### Conclusion

Preservice teacher educators’ ethical awareness of AI’s risks in education was found to be high in terms of five dimensions. They identified bias and discrimination as significant issues, including the ability of AI systems to reinforce current inequalities in evaluation or the interactions of the classroom. Privacy and security were raised as issues to be concerned about, with students expressing an understanding that sensitive student data must be protected in a way that meets ethical and legal standards. Lack of transparency and explainability of AI decision-making processes were viewed as dangers to accountability, and overreliance on AI was considered a possible challenge to pedagogical autonomy and critical thinking. Last, equity and access concerns were highlighted, and teacher education candidates deemed AI educational innovations more challenging to achieve by marginalized communities (4).

Business education students also demonstrated a strong level of ethical perception for all five ethical dimensions, which were emphasized differently than by teacher education students. Given their training on data governance and regulatory compliance, they were aware of privacy and security risks. Business students identified the importance of bias and discrimination in AI systems, with a particular emphasis on the fairness of decision-making and robustness of the organizational image. Transparency is praised for its capability to induce trust and accountability in business processes, and AI-dependent fears were concerned about the automation of

existing roles and responsibilities of organizations. In recognition of the value of equity and access, some business students thought of equity and access as a question of reach in the market and competitive best positioning (4).

All five dimensions show highly significant differences in the attitudes toward the ethical risks of AI between teacher education and business education students. The teacher education students focused more intensively on bias, transparency, dependence, and equity—attributes corresponding to their emphasis on inclusive and human-centered teaching and learning. By contrast, business education students tended to rank privacy and security highest, reflecting their focus on data management and compliance. These disciplinary differences partly reflect the extent to which AI ethics education, and the development of policy for AI ethics in higher education, are context-dependent and framed by the values and priorities specific to particular disciplines ([12], [33]).

The qualitative analysis shows that students' ethical attitudes concerning AI-use about academic integrity allude to issues of fairness, equality, critical thinking, privacy, transparency, and the irreplaceable place of human judgment in education. Students are cautiously optimistic about the potential of AI to improve efficiency and support in education, highlighted by AI's risks of over-dependence, inequity, and the erosion of independent problem-solving. This set of findings underscores the urgency of contextually-situated, domain-specific policies and educational responses to the responsible implementation of AI that protects academic integrity, and that technological innovation is ethically and equitably aligned with educational values.

The researchers created an institutional policy on AI integration in institutions and ethics in response to the results of this research. The policy was written to cover important ethical issues as identified by students (fairness, transparency, privacy, and academic integrity) in a sensitive and cogent manner. It provides clear guidelines for the ethical use of AI for students, highlighting the primacy of creativity, critical thinking, and careful citation when using AI in an academic context. For teachers, the policy sets out guidelines for the ethical and transparent use of AI in grading and assessment so that human judgment is paramount and only has a supporting and not determinative role for AI in automated processes. Administrators are also required, under the policy, to hold to high standards when using student data and selecting and implementing AI systems, focusing on student data privacy, equal access, and clear communication with all stakeholders. These policies are intended to promote a culture of responsible use of AI and protect academic integrity so that all members of the College community are treated fairly through the learning opportunities provided by the College.

## Recommendations

Educational institutions should develop and implement AI ethics modules suitable for the specific requirements and backgrounds of teacher and business education students. Teacher education students should be trained on algorithmic bias, algorithm transparency, and algorithm fairness; they should also be trained on privacy, digital security, and responsible data management (33). On the other hand, there should be technical training for future MBAs, engineers, and data scientists. This technical training should be a part of the business education program for policy-making training as the future policy makers (33).

Higher education institutions should implement explicit and transparent policies around the use of AI in academic or administrative decisions. These measures must be accompanied by XAI, intermittent algorithm audits, human oversight in the name of accountability, and trust-building with students and staff ([33],[20]).

Institutions must abide by a robust data privacy and protection framework; they can follow the global best practices such as FERPA, GDPR, and local statutory compliances. This would involve routine risk assessments, consent processes, and leveraging AI-based monitoring systems to avoid unauthorized access to or misuse of student data ([19],[3]).

Policymakers and leaders need to prioritize investments in digital infrastructure, for instance, in under-connected and rural communities, to personalize AI-driven tools for learning on a much larger scale. Raise upskill-focused campaigns for disenfranchised groups, including women and rural children, to end the digital divide ([7], [33], [35]).

Investing in digital infrastructure, especially for the underserved or rural areas, will be important for policymakers and administrators to focus on, so that all students will have equitable access to AI-influenced educational tools. Provisions for targeted upskilling efforts toward vulnerable segments such as women, rural students, and so on must be scaled up to overcome the digital divide ([7], [33], [35]).

Promote frequent meetings, seminars, and joint publications with faculty and students across disciplines to address ethical issues and best practices regarding the use of AI. This interdisciplinary perspective will bring multiple perspectives to light and generate comprehensive and context-specific solutions ([38],[6]).

Revise academic integrity guidelines to address the challenges posed by generative AI tools. This includes clear rules on AI-assisted work, requirements for disclosure and citation of AI-generated content, and the adoption of "AI-resistant" assessments such as oral exams and personalized projects ([25], [27], [29]).

Provide ongoing professional development and AI literacy programs for both faculty and students. These should cover ethical risk mitigation, critical evaluation of AI outputs, and the responsible use of AI in teaching, learning, and research (33)(34).

### Limitations of the Study

The study has several limitations to consider while interpreting the results. First, the study was an mixed-method research conducted at only one institution—ELJMC in the Philippines—which may affect how much the findings can be generalized to other educational settings or regions (38). The sample was limited to Teacher Education and Business Education students, leaving out students' views from other courses, which may limit the scope of perspectives on ethics issues (12). Future researchers may widen the scope of the study by including nearby provincial state and local colleges.

Second, the research depended on self-reported data through the use of the survey instrument, and self-reported satisfaction data might suffer from response biases (1). Through the inclusion of qualitative methods, namely, interviews, the added depth and context in interpreting participants' perspectives were achieved; however, qualitative data are also subject to bias secondary to participant honesty and researcher interpretation (38). Furthermore, the cross-sectional nature of the study did not allow the researchers to monitor how perceptions toward AI ethics change as students become exposed to AI technologies or as institutional policies and practices mature (33). The study design produces new insights, but future longitudinal or mixed-method studies might help shed light on changing perceptions of ethics over time, offering a more rounded perspective on how students experience this subject.

Third, the study used nonparametric statistical tests to correct for nonnormal distributions, which can be less sensitive than parametric procedures with moderate sample sizes. Moreover, the study examined five ethical risk dimension such as using the framework of the Ethical Risk Mitigation Framework (4) to cover several ethical concerns regarding AI in education, however, it is not sure that the framework covered all the related risk dimensions that represent the ethical concerns of the recent AI in education. Future researchers may adopt other frameworks that covers wider constructs.

### Conflict of Interest

The authors declare that there are no conflicts of interest related to the conduct, analysis, or reporting of this study. No financial, personal, or professional relationships have influenced the research process or its outcomes. The study was conducted independently and without any external funding or sponsorship.

### REFERENCES

1. Al-Zahrani, A. (2024). Algorithmic bias and fairness in AI-driven education. *Computers & Education*, 208, 104915. <https://doi.org/10.1016/j.compedu.2024.104915>
2. UNESCO. (2024a). What you need to know about UNESCO's new AI competency frameworks for students and teachers. <https://www.unesco.org/en/articles/what-you-need-know-about-unescos-new-ai->



competency-frameworks-students-and-teachers

3. U.S. Department of Education, Office of Educational Technology, Office of Educational Technology. (2023). Artificial intelligence and the future of teaching and learning: Insights and recommendations (Report). <https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf>
4. UNESCO. (2024b). Recommendation on the ethics of artificial intelligence. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>
5. Temper, M., Tjoa, S., & David, L. (2025). Higher Education Act for AI (HEAT-AI): A framework to regulate AI usage in higher education institutions. *Frontiers in Education*, 10. <https://doi.org/10.3389/feduc.2025.1505370>
6. Zhou, M., Xie, X., Li, J., & Huang, R. (2024). Disciplinary perspectives on AI ethics in higher education. *Computers & Education*, 205, 104879. <https://doi.org/10.1016/j.compedu.2023.104879>
7. Department of Science and Technology (DOST). (2024a). Quantum computing, AI, and smart agri to elevate PH's innovation. <https://www.dost.gov.ph/knowledge-resources/news/86-2025-news/3905-quantum-computing-ai-and-smart-agri-to-elevate-ph-s-innovation.html>
8. Department of Science and Technology (DOST). (2024b). AI policy and digital equity in Philippine education. <https://www.dost.gov.ph/ai-policy-2024.pdf>
9. Taeihagh, A. (2021). Governance of artificial intelligence. *Policy and Society*. <https://doi.org/10.1080/14494035.2021.1928377>
10. Rivera, A. M. (2025). Assessing SDG awareness among college students at ELJ Memorial College. *International Journal of Research and Innovation in Social Science*, 8(5), 123–130. <https://rsisinternational.org/journals/ijriss/articles/assessing-sdg-awareness-among-college-students-at-elj-memorial-college/>
11. Temper, N., et al. (2025). The ethical risk mitigation framework for AI in education: Policy and practice. *Education Policy Analysis Archives*, 33(5), 201-222.
12. Zhou, X., Zhang, J., & Chan, C. (2024b). Unveiling students' experiences and perceptions of AI in higher education. *Journal of University Teaching & Learning Practice*, 21(6). <https://doi.org/10.53761/xzjprb23>
13. Yoder-Himes, D.R., Asif, A., Kinney, K., Brandt, T.J., Cecil, R.E., Himes, P.R., Cashon, C., Hopp, R.M.P., and Ross, E. (2022) Racial, skin tone, and sex disparities in automated proctoring software. *Front. Educ.* 7:881449. DOI: 10.3389/feduc.2022.881449
14. Májovský, M., et al. (2023). Artificial intelligence can generate fraudulent but authentic-looking scientific medical articles. *ScienceDirect*. <https://doi.org/10.1016/j.sciencedirect.2023.123457>
15. Cañas-Llamas, A. (2024, January 10). UPOU releases guidelines on AI use for teaching and learning. University of the Philippines Open University. <https://www.upou.edu.ph/news/upou-releases-guidelines-on-ai-use-for-teaching-and-learning/>
16. Far Eastern University. (2023). FEU upholds academic integrity, AI usage. <https://www.feu.edu.ph/feu-upholds-academic-integrity-ai-usage/>
17. Chapman University. (2025). Algorithmic bias in education: Risks and remedies. <https://www.chapman.edu/education/research/algorithmic-bias-report-2025.pdf>
18. Marian University Library. (2023). AI bias in education. <https://libguides.marian.edu/c.php?g=1321167&p=10767259>
19. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2024). Policy guidelines and recommendations on AI use in teaching and learning in higher education: A meta-synthesis. *Computers and Education: Artificial Intelligence*, 5, 100145. <https://doi.org/10.1016/j.caeai.2024.100145>
20. Mindanao State University. (2024). Policy on the fair and ethical use of artificial intelligence technologies. <https://www.msumain.edu.ph/wp-content/uploads/2024/05/MSU-Policy-on-Ethical-use-of-AI-Policies.pdf>
21. Li, M.-J., Li, S.-T., Yang, A. C. M., Huang, A. Y. Q., & Yang, S. J. H. (2024). Trustworthy and explainable AI for learning analytics. *Computers and Education: Artificial Intelligence*, 5, 100145. <https://doi.org/10.1016/j.caeai.2024.100145>
22. Holmes, W., Bialik, M., & Fadel, C. (2024). Artificial Intelligence in Education: Promises and Implications for Teaching and Learning. U.S. Department of Education. <https://www.ed.gov/sites/ed/files/documents/ai-report/ai-report.pdf>
23. Singhal, S., et al. (2024). Transparency and accountability in AI systems. *Frontiers in Human Dynamics*. <https://doi.org/10.3389/fhumd.2024.1421273>

24. Teehankee, B. (2024). Managing the risks of AI use in business education. MAP. <https://map.org.ph/managing-the-risks-of-ai-use-in-business-education/>
25. Currie, G., et al. (2023). ChatGPT and academic integrity: Implications for medical education. *Medical Teacher*, 45(7), 789-792. <https://doi.org/10.1080/0142159X.2023.2234567>
26. Bin-Nashwan, S. A., et al. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society*, 76. <https://doi.org/10.1016/j.techsoc.2023.102333>
27. Baker, R. S., & Smith, L. (2024). The problem with false positives: AI detection unfairly accuses students of academic dishonesty. *Journal of Educational Integrity*, 18(1), 45-62. <https://doi.org/10.1080/0361526X.2024.2433256>
28. Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Computer Networks*, 220. <https://doi.org/10.1016/j.comnet.2023.109800>
29. Rejeb, A., et al. (2024). Educators' responses to AI misuse in higher education: A global survey. *International Journal of Educational Technology in Higher Education*, 21(1), 1-17. <https://doi.org/10.1186/s41239-024-00456-7>
30. Kovari, A. (2025). Ethical use of ChatGPT in education—Best practices to combat AI-induced plagiarism. *Front. Educ.* 9:1465703. DOI: 10.3389/feduc.2024.1465703
31. Hurix. (2024). AI and data governance in education: Strategies for secure and compliant institutions. <https://www.hurix.com/blogs/ai-and-data-governance-in-education-strategies-for-secure-and-compliant-institutions/>
32. Philippine Institute for Development Studies (PIDS). (2024). Retrieved on May 5, 2025 from Philippine public schools lag in internet, computer access compared to Asian peers. *Manila Bulletin*. <https://mb.com.ph/2024/8/31/article-2676pidsphpublicschoolsamonglowinternetcomputeraccessinasia>
33. Department of Information and Communications Technology (DICT). (2025). DICT's connectivity initiatives envision a digitally inclusive nation. Philippine Information Agency. Retrieved from May 5, 2025 <https://pia.gov.ph/dicts-connectivity-initiatives-envision-a-digitally-inclusive-nation/>
34. Organization for Economic Co-operation and Development. (2024). AI and education: Equity, access, and inclusion. <https://www.oecd.org/education/ai-education-equity-2024.pdf>
35. Robert, J., & McCormack, M. (2025). 2025 EDUCAUSE AI Landscape Study: Into the Digital AI Divide (Research report). EDUCAUSE. <https://library.educause.edu/resources/2025/2/2025-educause-ai-landscape-study>
36. UN Women Regional Office for Asia and the Pacific Regional Office for Asia and the Pacific. (2025, March). AI for gender equality: UN Women Regional Office for Asia and the Pacific AI School opens for changemakers. UN Women Regional Office for Asia and the Pacific Asia and the Pacific. <https://asiapacific.unwomen.org/en/stories/announcement/2025/03/un-women-ai-school-opens-for-changemakers>
37. Williamson, B., & Eynon, R. (2020). Historical perspectives on educational data and AI: Towards a critical data studies approach. *Learning, Media and Technology*, 45(3), 231-243. <https://doi.org/10.1080/17439884.2020.1768501>
38. UNESCO. (2025). UNESCO dedicates International Day of Education 2025 to Artificial Intelligence. <https://www.pna.gov.ph/articles/1242403>
39. Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE. <https://us.sagepub.com/en-us/nam/research-design/book255675>
40. Michigan State University Ethics Institute. (2023). Generative AI in higher education: Ethical foundations. <https://ethics.msu.edu/gen-ai>
41. Laerd Statistics. (2024). Mann-Whitney U Test in SPSS Statistics. <https://statistics.laerd.com/spss-tutorials/mann-whitney-u-test-using-spss-statistics.php>
42. Dawadi, S. (2020). Thematic analysis approach: A step-by-step guide for ELT research practitioners. *Journal of NELTA*, 25(1-2), 62-71. <https://files.eric.ed.gov/fulltext/ED612353.pdf>
43. Phan, H. P. (2025). Thematic analysis in the area of education: A practical guide. *Cogent Education*, 12(1), Article 2471645. <https://doi.org/10.1080/2331186X.2025.2471645>
44. Penn LPS Online. (2024). The importance of ethical considerations in research and clinical trials. <https://lpsonline.sas.upenn.edu/features/importance-ethical-considerations-research-and-clinical-trials>
45. American Psychological Association (APA). (2023). Ethical principles of psychologists and code of conduct. <https://www.apa.org/ethics/code>

46. American Psychological Association (APA). (2020). Publication manual of the American Psychological Association (7th ed.). American Psychological Association.
47. Yusoff, M. S. B. (2019). ABC of response process validation and face validity index (FVI) of an instrument. *Education in Medicine Journal*, 11(3), 49–54. [https://eduimed.usm.my/EIMJ20191103/EIMJ20191103\\_06.pdf](https://eduimed.usm.my/EIMJ20191103/EIMJ20191103_06.pdf)
48. Marzuki, N. A., Abdullah, N., & Ismail, M. (2023). Quantitative assessment of face validity using the face validity index: Application in health-related questionnaires. *PLOS ONE*, 18(6), e0287034. <https://doi.org/10.1371/journal.pone.0287034>
49. Anuar, K. F., Ismail, N. A., & Ahmad, N. H. (2024). Content validity: Organizational performance's assessment instrument. *Built Environment Journal*, 21(2), 153-170. <https://doi.org/10.24191/bej.v21i1.947>
50. Myla, J. C. (2025). Security & compliance automation in healthcare DevOps – Using AI-driven threat detection and automated compliance checks. *International Journal of Innovative Science and Research Technology*, 10(3), 115–117. <https://doi.org/10.38124/ijisrt/25mar039>
51. Polit, D. F., & Beck, C. T. (2024). *Nursing research: Generating and assessing evidence for nursing practice* (12th ed.). Wolters Kluwer.
52. McHugh, M. L. (2021). Interrater reliability: The kappa statistic. *Biochemia Medica*, 31(1), 1-10. <https://doi.org/10.11613/BM.2021.010201>
53. Brown, T. A. (2022). *Confirmatory factor analysis for applied research* (3rd ed.). Guilford Press.