JOURNAL OF INTERDISCIPLINARY PERSPECTIVES

ISSN Print: 2984-8288, ISSN Online: 2984-8385
Vol. 4 No. 1, pp. 302-312, January 2026

Original Article

The Mediating Role of Artificial Intelligence (AI) Literacy in Graduate Seminar Coursework and Ethical Preparedness: Insights from the Graduate Students

Artemio L. Florendo, Jr.

Author Information:

Graduate School Department, Polytechnic College of La Union, La Union, Philippines

Correspondence:
artemioflorendo21@gmail.com

Article History:

Date received: October 6, 2025
Date revised: December 19, 2025
Date accepted: December 30, 2025

Recommended citation:

Florendo, A., Jr. (2026). The mediating role of Artificial Intelligence (AI) literacy in graduate seminar coursework and ethical preparedness: Insights from the graduate students. *Journal of Interdisciplinary Perspectives*, *4*(1), 302-312. https://doi.org/10.69569/jip.2025.698

Abstract. Given the lack of national and institutional Artificial Intelligence (AI) governance frameworks, the governance and education of AI in Philippine higher education often rest with faculty. This makes course-level integration a productive starting point. This pilot study examined how graduate students develop ethical preparedness for AI, arguing that the relationship between such preparedness and their exposure to AI ethics instruction is mediated through operational literacy on AI. A mixed-methods intervention was conducted with 39 graduate students enrolled in the Master of Arts Major in Educational Management, who participated in a 20-hour AI literacy and ethics course embedded in a Graduate Seminar. To capture the Philippine education realities often overlooked in global frameworks, a researcher-developed questionnaire was employed to reconceptualize and measure AI literacy and ethical preparedness in the local context. Quantitative results showed significant improvements, with Artificial Intelligence literacy mediating the relationship between ethical preparedness and artificial intelligence literacy ($\beta = .46$, $p < .001$). Qualitative insights through classroom observations illustrated how students translated abstract principles into situated judgment. This paper extends AI ethics discourse by reframing it not only as a matter of principles but also as a function of operational literacies that enable educators to exercise situated ethical judgment, especially within local contexts.

Keywords: *AI education; AI ethics; Artificial Intelligence; AI literacy; Education management.*

Artificial intelligence is already entrenched in the daily life of higher learning. Graduate students rely on AI to draft outlines, analyze data, and refine arguments, while professors use it to automate administrative tasks or create course materials. On the one hand, such tools hold out the promise of efficiency; on the other, they pose deep ethical challenges. In the absence of institutional policies and structured training programs governing the use of AI, faculty have taken the lead in incorporating AI ethics into their curricula. However, on closer examination of case studies of these integrated courses, I observe that most focus on articulating principles, with little attention to operational literacy to make those principles useful in practice (Sperling et al., 2024). Such an insufficiently designed AI literacy course is troubling, particularly for graduate students in education. Daher (2025) contends that teacher training must extend far beyond ethical principles to encompass the functions of AI systems, their design, failures, and the ways bias is coded. This paper argues that ethical preparedness, particularly among teachers, is mediated by AI literacy. Ethical preparedness cannot rest on principles alone; a technical foundation is needed to ensure that values are not merely stated in theory but responsibly implemented

in the day-to-day activities of teaching and research.

The varying extent to which AI literacy is integrated into courses may be attributable to limited AI governance models and the systematic inclusion of AI topics in higher education curricula. In the Philippines, for instance, there are no clear guidelines yet from the Commission on Higher Education (CHED) on the use of AI in the academe (Villanueva, 2025), nor is there any directive on the systematic integration of AI literacy. For now, the task of instructing and regulating its use rests with institutions or, at times, with individual faculty members' discretion. In Villanueva (2025), it was noted that at least eight universities in the country have issued formal guidelines on the use of AI. Among them is the University of the Philippines (UP) System, which issued principles on responsible use of AI. Broad principles, as they stand, are provided to UP campuses, departments, and individual faculty members for implementation. Similarly, the guidelines issued by the Ateneo de Manila University (ADMU) grant faculty the liberty to discuss the ethical implications of AI use in academic contexts, as well as regulatory powers, such as developing their own course policy on GenAI and including it in their syllabi. Wang et al. (2021) termed this "professors' academic governance," in which decision-making is generally decentralized, with professors granted authority to regulate the use of AI in their respective courses.

Aside from the lack of a standard AI governance model, equally concerning is the absence of systematic instruction on AI in the country's higher education. Outside of computer science and related disciplines, students rarely encounter courses that introduce AI concepts in a structured way. AI training is largely absent from most curricula and remains a decision made by professors. This is problematic, particularly in education courses, since AI is applied not just in study but also in their instruction and daily settings. According to some scholars, AI ethics is also incorporated only "if time allows," rather than as a core element of education (Garrett et al., 2020). This policy-pedagogical gap gives rise to pressing questions. How do teachers cope with the unprecedented rate of AI innovations while ethical discourse lags? How do we establish best practices when AI literacy and ethics research remains fragmented and unfinished? Frieary (2025) suggests a modest but promising start: intentional course-level integration. In the context of the decentralized role of faculty discretion in AI education, owing to the lack of standard guidelines and systematic curricula, course-level integration is a productive starting point. This is why, as a professor who teaches Graduate Seminar in the Master of Arts Major in Education Management at Polytechnic College of La Union, I sought to integrate AI ethics into the course syllabus. During the 1st semester of the 2025-2026 academic year, I dedicated a total of 20 hours to AI ethics, particularly to discussions on "AI Use and Ethics in Research" and other research-related topics.

Here, the idea of AI literacy and ethical preparedness becomes critical. AI literacy is more than knowing how to operate a platform or command a chatbot (Matthews & Bartley, 2025). It is a matter of developing the insight and judgment to understand what these systems accomplish, how they operate, and where their boundaries lie. It has the sense to know when AI can facilitate scholarship and when it undermines the very values research is designed to maintain. At the same time, AI ethical readiness, though not formally defined in current scholarship, should be understood here, ideally, not as a set of rules to memorize, but as a mindset that prepares teachers to address the challenges of AI application in instruction and research. At the graduate level, this is a state of mind and a set of skills. This worldview will foster responsibility and practices that provide confidence to move forward when emerging technologies obscure ethical boundaries.

Why then consider the mediating role of AI literacy? Because AI ethics without literacy is incomplete. Any effort to ethically equip educators cannot be achieved by discussing ethics alone; it requires AI literacy. For example, governance models emphasize ethical guardrails (i.e., what not to do and what to ban) but overlook the operational know-how regarding what AI is, how it operates, and where it could go wrong. This gap matters. Consider the dependence on plagiarism-detection software. Without a practical understanding of how AI-generated content can be identified or evaded, students can be unfairly targeted because "writing too well raises a red flag" (Showewimo, 2025), or they can cleverly avoid detection by passing their work through humanizing software. On the other hand, technical literacy without ethics threatens to produce well-skilled but unreflective consumers who are efficient but oblivious to its ethical implications. The actual challenge lies at the intersection, where AI literacy enables learners to apply ethical principles in practice.

This research fills that gap. It examines the role of AI literacy in mediating the relationship between exposure to

integrated AI ethics coursework and AI ethical preparedness among students of Master of Arts Major in Education Management. The graduate seminar, as a course that emphasizes research training and professional reflection, provides a conducive environment for integrating AI ethics. In doing so, the research reframes the discourse on general calls to "teach AI ethics" toward identifying the essential competencies. It argues that ethical preparedness is not an automatic outcome of AI ethics training but is shaped by the extent to which students are AI-literate. This mediating function is the one that must be apprehended if graduate programs are to be compliant with the ethical and intellectual challenges of an age of AI.

Methodology

Research Design

This study employed a quasi-experimental pretest–posttest design with mixed methods to examine how graduate students' AI literacy and AI ethical preparedness evolved after completing a structured Graduate Seminar coursework. The design was quasi-experimental because the seminar class was an intact group rather than a randomized pool. The pretest–posttest design provided a credible means of monitoring within-subject change. Given the relatively small sample size (n=39), I conducted this as both a pilot study and a preliminary investigation. No attempt was made toward broad generalizability; instead, an initial probe of the mediating role of AI literacy in the relationship between the integration of graduate coursework and ethical readiness was undertaken. As a pilot study, the research also helped assess the practicality of the research design and the validity of the instruments, particularly the AI Ethical Preparedness Scale, which was constructed in this study based on contextual considerations. Quantitative analysis was supplemented by qualitative data derived from my observations of the course while the professor facilitated the course.

Participants and Sampling Technique

The participants in the study were Master of Arts students majoring in Education Management enrolled in the Graduate Seminar course at the Polytechnic College of La Union in the Philippines during the 2024–2025 academic year. A total of 39 graduate students participated, representing a range of fields of specialization, including educational management, curriculum studies, and instructional leadership. Since the study was conducted within an intact class, no sampling was necessary; the entire cohort was included as the accessible population for analysis. This aligned with the study's character as a preliminary pilot investigation, in which the aim was to examine feasibility, explore emerging patterns, and test instrument reliability rather than to generalize findings to a broader population.

Research Instrument

Two instruments were developed for this study: an AI Literacy Scale and an AI Ethical Preparedness Scale, both contextualized to the Philippine graduate education setting.

AI Literacy Scale. This study modified the Artificial Intelligence Literacy Scale for Teachers (AILST) initially developed by Ning et al. (2025). The scale was structured around four core dimensions: AI perception, AI knowledge and skills, AI applications and innovation, and AI ethics. Each dimension included a range of items designed to capture teachers' understanding, abilities, and critical awareness of AI use in educational contexts. Responses were recorded on a five-point Likert-type scale with options labeled A to E, ranging from *Fully Applicable* to *Completely Not Applicable*. The instrument was expanded and localized by incorporating additional items. These adjustments addressed:

(a) The incipient stage of AI adoption in the Philippines.
(b) Resource constraints and digital divide issues, such as limited connectivity and aging hardware.
(c) Techno-linguistic bias in multilingual classrooms.
(d) The dual role of graduate students as both learners and teachers (feedback loop).
(e) Broader structural issues, including inequality and the marginalization of rural schools.

These inclusions ensured that the instrument not only measured general AI literacy but also captured its situated and context-dependent nature in Philippine higher education.

AI Ethical Preparedness Scale. To measure teachers' ethical preparedness regarding AI, the researcher developed

the AI Ethical Preparedness Scale (AIEPS). The instrument was primarily anchored in the University of the Philippines' Principles for Responsible and Trustworthy Artificial Intelligence (UP, 2023), which were adapted to address the lack of a locally grounded framework for AI ethics in Philippine education. The development of the scale followed a twofold process.

First, the fifteen UP principles were reviewed and conceptually consolidated into three higher-order constructs: Societal–Ethical Orientation (common good, empowerment, cultural sensitivity, privacy, and accountability), Research and Development Ethics (meaningful human control, transparency, fairness, safety, and environmental sustainability), and Pedagogical–Professional Preparedness (primacy of learning goals, human capital development, capacity building, education management, and collaboration). Each construct was then broken down into observable indicators that could be translated into measurable survey items.

Second, the researcher integrated contextual considerations identified as gaps in the existing AI ethics and literacy literature. These include:

(1) The realities of resource constraints and the digital divide (Espinosa et al., 2025).
(2) The incipient stage of AI adoption in the Philippines, where teachers must exercise caution in the absence of comprehensive policies.
(3) Issues of language and techno-linguistic bias (Noor & Kanitroj, 2025).
(4) The feedback loop is inherent in the dual role of MA Education Management students as both teachers and learners.

The resulting scale comprised 22 items distributed across the three constructs. Items were rated on a five-point Likert scale ranging from 1 ("Strongly Disagree") to 5 ("Strongly Agree").

Validation Process. The instrument was pilot-tested with the study participants (n = 39) to examine its initial psychometric properties. Reliability testing using Cronbach's alpha was conducted to assess internal consistency, with 0.70 used as the threshold for acceptability (Taber, 2018).

Table 1. *Reliability of the AI Literacy Scale (AILS) and AI Ethical Preparedness Scale (AIEPS)*

Instrument / Dimension	No. of Items	Sample (n)	Cronbach's α
AI Literacy Scale (AILS)		39	
AI Perception	6		.81
AI Knowledge and Skills	7		.84
AI Applications and Innovation	6		.79
AI Ethics	5		.83
Total AILS	**24**		**.88**
AI Ethical Preparedness Scale (AIEPS)		39	
Societal–Ethical Orientation	8		.85
Research and Development Ethics	7		.82
Pedagogical–Professional Preparedness	7		.87
Total AIEPS	**22**		**.91**

Intervention

As noted in the preceding discussion, the integration of artificial intelligence in higher education is often treated as an optional add-on, addressed only when extra time permits or when individual instructors take the initiative. International organizations such as the OECD have noted that AI in education is often adopted in a fragmented, ad hoc manner, leaving many students with uneven exposure to both the opportunities and the risks of AI use (OECD, 2021). The present study addresses this gap by situating AI literacy and AI ethical preparedness within the Graduate Seminar course for students of Master of Arts Major in Education Management. In the university's curriculum, the Graduate Seminar is a capstone subject that orients students toward the rigor of educational research. It covers essential areas, including formulating research problems, conducting literature reviews, communicating research findings, and preparing for a thesis or dissertation. Accordingly, a structured 20-hour intervention was developed, which combined (a) one dedicated lesson on AI ethics with (b) the deliberate infusion of AI-related literacy and ethical considerations across multiple seminar topics as shown below:

Selecting a Research Topic (2 Hours). The students were exposed to AI technologies for generating ideas on

research trends and areas of need. This was matched with societal–ethical direction, noting that research topics had to address real educational needs without resorting to superficial or AI-led framing.

Developing Research Problems (2 Hours). The learners employed AI to define variables and scope. The session emphasized research and development ethics in ensuring transparency and human control in defining research problems.

Review of Related Literature (3 Hours). AI-supported tools were used for citation assistance, article summarization, and keyword search. Ethical emphasis continued to be placed on a societal–ethical orientation, particularly honesty and accountability in the management of scholarly materials.

Action Research and Design (3 Hours). AI was used to offer methodological recommendations and design alternatives. This task upheld research and development ethics by ensuring the fairness and appropriateness of methodological choices.

Statistical Techniques and Data Analysis (2 Hours). Students investigated AI-augmented statistical methods. The focus was on research and development ethics, particularly the demand for accuracy and effective human oversight of data interpretation.

AI Research Ethics (4 Hours, Dedicated Session). Students discussed AI ethics, governance, and disclosure standards. This session integrated society–ethical orientation and research ethics, promoting responsible use of AI, fairness, and transparency.

Communicating Research (2 Hours). Structuring arguments, writing outlines, and creating presentations were done with AI tools. The ethical alignment was pedagogical–professional readiness, emphasizing the principle that AI should assist—but not substitute for—the genuine communication of knowledge.

Research Practice and Academic Integrity (2 Hours). Students explored AI-detection tools such as Turnitin. This was connected to pedagogical–professional preparedness, with a focus on the educator's role in encouraging learners to use AI ethically and responsibly.

Data Collection Procedure

At the start of the Graduate Seminar course, participants completed two instruments:
(1) The Modified Artificial Intelligence Literacy Scale for Teachers
(2) The AI Ethical Preparedness Scale

The pretest established baseline measures of students' AI literacy and ethical preparedness before their exposure to the integrated AI ethics instruction. Over six weeks, AI ethics instruction was integrated into the Graduate Seminar course through lectures and guided discussions. As part of the data collection process, class discussions were also documented to capture qualitative insights into how participants interpreted and applied the ethical frameworks. The instruments were administered to participants at the end of the intervention. This gave post-intervention scores for comparison with the pretest. Students' attendance was also monitored as a measure of exposure to the course's AI ethics components. Responses were collected, and reliability was assessed using Cronbach's alpha to evaluate the internal consistency of the instruments. Last, the discussion transcripts were examined, coded, and analyzed thematically.

Data Analysis Procedure

The data analysis in this study combined statistical methods with qualitative interpretation. On the quantitative side, descriptive statistics, such as means and standard deviations, were used to provide baseline data on students' AI literacy and ethical preparedness, both before and after their participation in the Graduate Seminar. Exposure to the integrated coursework was also quantified by attendance at sessions that explicitly addressed the ethics of AI use. To determine whether participation made a measurable difference, the study employed the Wilcoxon signed-rank test, a nonparametric alternative suitable for small samples that may not meet the assumption of normality. This test assessed changes in pretest and posttest scores across the constructs of literacy and ethical

preparedness. The magnitude of these changes was evaluated using rank-biserial correlations, which provided insight into the strength of the observed effects.

Relationships among coursework exposure, AI literacy, and ethical preparedness were further examined using nonparametric correlation analyses. Beyond these correlations, the research tested a key hypothesis: that literacy mediated the relationship between coursework integration and ethical readiness. To measure this, a bootstrapped mediation analysis was used, creating bias-adjusted confidence intervals for the indirect effect. Bootstrapping was used because it is robust to non-normal sampling distributions and is suitable for modest sample sizes. Quantitative results were supplemented with qualitative analysis. Classroom exchanges and reflective discussions were subjected to thematic analysis, which allowed participants' voices to surface and enrich the interpretation of statistical results. These narratives illustrated, for instance, how students navigated the tension between AI's affordances and limitations, or how pedagogical values influenced their engagement with ethical issues.

Ethical Consideration

Each participant was provided with an informed consent form that explained the study's objectives, the nature of the intervention, and the types of data to be collected. Students were reminded throughout the semester of their right to withdraw without penalty and of the voluntary nature of their involvement. Confidentiality was also handled with great care. Identifiers were excluded, and only aggregate results were reported. In the classroom observation component, reflections and stories were also coded to conceal individual identities while retaining the richness of their insights. Lastly, the researcher recognized the ethical reflexivity required by the subject matter itself, particularly given that the researcher also served as the course instructor. The students were explicitly informed that participation in the surveys was voluntary and unrelated to their course grades. Anonymity was maintained in reporting, and students were assured that no individual responses would influence academic standing. By consciously separating the instructional and research roles, the researcher sought to reduce power imbalances and protect the integrity of student voices.

Results and Discussion

Quantitative Results

Preliminary checks of the data indicated that pretest and posttest scores for AI literacy and ethical preparedness were not normally distributed, as assessed through Shapiro–Wilk tests ($p < .05$). Skewness values at pretest ranged from −0.82 to −0.21 (SE = 0.38), and kurtosis values ranged from −0.59 to 0.72 (SE = 0.75). In contrast, posttest skewness ranged from −1.03 to −0.11 (SE = 0.38) and kurtosis ranged from −0.41 to 1.11 (SE = 0.75). Given these data characteristics, nonparametric analyses were employed to assess temporal changes. Specifically, Wilcoxon signed-rank tests were used to compare pretest and posttest scores across all constructs of AI literacy and ethical preparedness. This test enabled evaluation of median differences while accounting for the ordinal nature and non-normal distribution of the data (Okoye & Hosseini, 2024). To complement these analyses, a bootstrapped mediation test with 5,000 resamples was conducted to examine whether AI literacy mediated the relationship between instructional exposure and ethical preparedness. The following subsections report the descriptive and inferential findings in detail.

Exposure to AI Ethics Integrated Coursework

As shown in Table 2, most graduate students (76.9%) received the full 16–20 hours of AI ethics–integrated instruction. This indicates that the majority were consistently present and engaged in the intervention sessions. A smaller group, however, had lower exposure (0–15 hours), mainly due to missed classes. Although these cases were few, they indicate variation in the extent to which students could fully experience the intervention.

Table 2. *Distribution of Student Exposure to AI Ethics–Integrated Lessons*

Hours of Exposure	Frequency	Percentage (%)
0–5 Hours	1	2.6
6–10 Hours	3	7.7
11–15 Hours	5	12.8
16–20 Hours	30	76.9
Total	**39**	**100.0**

Pretest and Posttest Scores

Table 3 shows that the 20-hour Graduate Seminar intervention significantly enhanced AI literacy and AI ethical readiness across all measured domains, with large effect sizes in all cases. Using the Wilcoxon signed-rank test, the mean score of AI literacy climbed from 3.12 (SD = 0.47) to 4.01 (SD = 0.39), while ethical preparedness on average improved from 3.28 (SD = 0.51) to 3.92 (SD = 0.44). At the literacy level, students exhibited significant improvement in perception (Mdiff = 0.80, r = 0.78), knowledge and skills (Mdiff = 0.85, r = 0.84), applications and innovation (Mdiff = 0.82, r = 0.80), and ethics (Mdiff = 0.87, r = 0.82). These improvements imply that students not only became more technically proficient but also more critically engaged with the use of AI in education, research, and professional activities.

Table 3. *Pretest & Posttest AI Literacy and AI Ethical Preparedness Scores (n = 39)*

Construct	Pretest Mean (SD)	Posttest Mean (SD)	Median Diff.	Z	p-value	r (Effect Size)
AI Literacy						
AI Perception	3.05 (0.51)	3.92 (0.43)	0.80	-5.02	.000004	0.78
AI Knowledge & Skills	3.10 (0.48)	4.05 (0.41)	0.85	-5.33	.000001	0.84
AI Applications & Innovation	3.18 (0.44)	4.07 (0.40)	0.82	-5.11	.000009	0.80
AI Ethics	3.15 (0.49)	4.02 (0.38)	0.87	-5.27	.000003	0.82
Overall AI Literacy	**3.12 (0.47)**	**4.01 (0.39)**	**0.85**	**-5.41**	**.000001**	**0.86**
AI Ethical Preparedness						
Societal–Ethical Orientation	3.22 (0.53)	3.91 (0.46)	0.75	-5.08	.000006	0.79
Research & Development Ethics	3.29 (0.50)	3.94 (0.42)	0.72	-5.21	.000002	0.83
Pedagogical–Professional Preparedness	3.33 (0.52)	3.91 (0.44)	0.68	-4.95	.000012	0.76
Overall Ethical Preparedness	**3.28 (0.51)**	**3.92 (0.44)**	**0.70**	**-5.37**	**.000002**	**0.84**

Concurrent developments in ethical preparedness were seen. Students indicated higher societal–ethical orientation (Mdiff = 0.75, r = 0.79), enhanced responsiveness to research and development ethics (Mdiff = 0.72, r = 0.83), and higher pedagogical–professional readiness (Mdiff = 0.68, r = 0.76). These findings suggest that students became more aware of the long-term implications of AI use, ranging from maintaining fairness and transparency to exemplifying responsible practice as future teachers. Taken together, these results confirm that technical proficiency and ethical readiness support one another: technical competence without an ethical foundation risks being shallow, while ethical norms without technical appreciation remain abstract. The integrated structure of the Graduate Seminar, hence, offered a unified path for graduate students to transfer AI knowledge into responsible, situated pedagogical practice.

Mediation Analysis

A bootstrapped mediation analysis was conducted to investigate whether improvements in AI literacy mediated the effect of instruction on AI ethical preparedness. The study showed a significant indirect effect ($\beta = .46$, 95% CI [.30, .65], $p < .001$). This suggests that gains in ethical preparedness were not solely the immediate outcome of instruction ($\beta = .22$, $p = .024$) but were significantly enhanced by students' increasing literacy. At the practical level, this implies that learning about AI ethics in the abstract was not sufficient; those students who also gained some insight into AI systems—what they can do, what they cannot, and how they can be biased—were better positioned to address tricky questions around societal impact, research integrity, and classroom practice. Ethics education made a modest difference, but its radicalizing potential unfolded only when literacy served as a mediator, enabling students to apply principles as situated, actionable judgment.

Table 4. *Bootstrapped Mediation Analysis of AI Literacy on the Relationship Between Instruction and Ethical Preparedness (n = 39)*

Pathway	β	SE	95%CI (Bootstrapped)	p-value
Instruction → AI Literacy (a)	.78	.09	[.60, .94]	<.001
AI Literacy → AI Ethics (b)	.59	.11	[.35, .81]	<.001
Instruction → AI Ethics (c′, Direct)	.22	.10	[.03, .41]	.024
Indirect Effect (a × b)	.46	.08	[.30, .65]	<.001

Note. 5,000 bootstrap samples, bias-corrected confidence intervals.

Qualitative Results

The quantitative findings indicated significant improvements in AI literacy and ethical preparedness, and that the former mediates between exposure to AI ethics integrated studies and AI ethical readiness. In unpacking this, I have taken notes on points of discussion raised during the Graduate Seminar, including questions, tensions, and

the gradual development of reflexive stances towards AI. Across twenty hours of combined lessons, I observed that students' experiences with AI-related challenges spanned three areas of ethical readiness: societal–ethical orientation, research ethics, and pedagogical–professional readiness.

Societal–Ethical Orientation

AI Detection and Fairness. In one of the early sessions, our discussion of Turnitin's AI-detection feature prompted widespread concern. A student raised: "Sir, what if someone gets a 0% plagiarism score but the write-up still sounds like it was AI output? Would that be okay?" This question from one of my students, who herself teaches high school students, generated interest in my class. Some of my students argued that a 0% score was proof of honesty, while others worried that it encouraged gaming the system. I reminded the class that reducing ethics to technical detection was dangerous, echoing OECD's (2021) critique that education systems risk turning integrity into compliance. We concluded that actual preparedness required discernment and honesty, not just passing a system check. This is where I discussed the technicalities and limitations of AI detection tools such as Turnitin and encouraged them to read the paper rather than fixating on the similarity score. First, I discussed the distinction between Similarity and Plagiarism, which Turnitin reports as Similarity rather than Plagiarism. However, high similarity can indicate potential plagiarism. However, I also stressed that 0% is still as dangerous as 100% because Turnitin's similarity tool is a text-matching system; therefore, at least the referenced material should be highlighted. Secondly, I discussed how these weaknesses of Turnitin are being manipulated by some students for academic dishonesty, including the use of white quotation marks, uploading an image to display text, replacing characters, using Macros, and using AI to paraphrase. In this instance, the discussion is not only about the concept of academic honesty but also about their literacy with AI, which prompted further discussion of what is ethically permissible, particularly regarding plagiarism.

Accountability in Shared Work. Another time, during our session on Literature Review, a participant asked: "If AI helped us summarize articles for this project, who takes the blame if the summary is wrong—us or the AI?" Some laughed nervously, admitting they had used AI to speed up reading tasks. I guided them through a discussion about accountability, stressing that while AI may assist, authorship cannot be outsourced. We unpacked Floridi and Cowls' (2019) principle of accountability and linked it to their dual role as graduate students and future teachers. The consensus reached was clear: responsibility rests with humans, even when tools are involved.

Cultural Sensitivity. Later, when students tested AI translation tools for local dialects, one participant exclaimed in frustration: "It is like the AI does not even know regions exist!" The class laughed, but the comment revealed that, more often than not, AI is tone-deaf to the cultural specificities of local languages. I used this moment to highlight techno-linguistic bias and the need for culturally sensitive AI practices. This was another opportunity to discuss how AI was built and its current limitations. Espinosa et al. (2025) argued that localized, frugal AI practices are essential in resource-constrained contexts such as the Philippines. Students reflected that ethical preparedness required not only personal honesty but also cultural vigilance to ensure inclusivity in the use of AI.

Research Ethics

Fictitious References. During our literature review session, a student admitted, "Last semester, I used ChatGPT for citations, but many of them turned out to be fake. I did not know what to do." The class reacted with concern, and some reported similar experiences. I reminded them that citation is not a mechanical task but a matter of integrity. Ng et al. (2022) argued that AI literacy requires critical awareness rather than blind reliance. I discussed that ChatGPT is not a reliable tool for Literature Reviews because it tends to fabricate data and generate fictitious references. However, I noted that formatting credible references in a particular style using AI may be less of an ethical concern. Moreover, if AI is necessary, there is a way to not only declare but also cite it properly. This practice, however, permissible for transparency, could make the review less credible. We practiced checking sources manually and discussed disclosure as an act of academic honesty. This reinforced the view that AI literacy (i.e., understanding what AI can and cannot do) and ethics intersect most sharply in research.

AI in Statistics. When I discussed Research Methods and touched on data analysis, another student asked directly: "Can ChatGPT just run the test for us? It is faster." This sparked a heated exchange—some argued for efficiency, while others worried about accuracy. I intervened by distinguishing between legitimate uses of AI for decision

support (e.g., clarifying which test to use) and inappropriate substitution (e.g., allowing AI to generate fabricated outputs). I have disclosed that, in some instances, I must make research decisions; it has become second nature to me to consult ChatGPT. The problem arises when ChatGPT is used to calculate statistics, potentially undermining the accuracy of the quantitative results. This distinction is essential in research ethics: automation should support, not replace, professional judgment. Students left recognizing that shortcuts in research carry ethical costs.

Declaration of AI Use. When I introduced the idea of declaring the use of AI in research papers, several students appeared uneasy. One asked: "But sir, if we disclose it, won't professors penalize us?" I reassured them that disclosure is not about self-incrimination but about transparency, a key principle of most AI governance models. I discussed Villanueva's (2025) research on the regulatory frameworks of eight leading universities in the Philippines, noting a commonality among them: the declaration of AI use. I mentioned that even journals for publication are encouraging authors to declare the use of AI and provide templates for doing so. We explored how disclosure builds credibility, and one student reflected: "It is like admitting mistakes in class—it feels risky, but it shows honesty." This conversation redefined ethics not as hiding AI use but as practicing openness in scholarship.

Pedagogical–Professional Preparedness

Teaching Students About AI. In one session, a student (who also taught at a public high school) asked candidly: "If my high school students use ChatGPT for essays, do I punish them or guide them?" Some admitted they would feel betrayed, whereas others considered guidance more constructive. I framed the issue around the critique of punitive approaches, emphasizing that teachers should guide rather than police. This is where I revisited regulatory frameworks that range from "soft" to "hard" approaches, from lenient, principle-based to punitive, depending on the situation (Villanueva, 2025). We discussed practical strategies, such as requiring process work alongside final drafts, to balance accountability with learning.

Balancing Efficiency and Learning. Another conversation unfolded around efficiency. A student shared: "I use AI to draft my outlines—it saves me hours." Immediately, another countered: "But if our students always do that, will they still learn how to think critically?" This dilemma illustrated pedagogical ethics: balancing productivity with deep learning. I reminded them that their responsibility as teachers was not to ban technology but to design tasks that preserved the primacy of learning goals. One practical strategy we discussed was having students first draft ideas by hand in notebooks before consulting AI tools, an approach that preserves human thought, slows the reasoning process, and strengthens critical engagement. This mirrored the call of Garrett et al. (2020) to foreground ethics in pedagogy rather than treat it as an afterthought.

The Feedback Loop of Dual Roles. Perhaps the most striking moment came when one participant reflected aloud: "What we struggle with here as graduate students—our own students will struggle with too. How we respond now is how we will teach later." This realization captured the essence of the feedback loop: being both learners and teachers, they were practicing the same reflexivity they would later cultivate in their students. To make this concrete, I encouraged them to integrate AI literacy into the subjects they teach explicitly. I also invited them to design small-scale action research projects related to AI. Through these activities, students experienced firsthand how ethical reflection and practical literacy intersect, thereby embodying the dual roles of informed educator and reflective practitioner, and recognizing how the skills they cultivated now would directly shape their future classrooms.

Quantitative-Qualitative Integration

Taken together, the quantitative and qualitative results indicate a statistical relationship in which AI literacy mediates the path from instruction to ethical preparedness ($\beta = .46$, $p < .001$), as further supported by classroom observations. As I have articulated in the qualitative phase, literacy is not a static set of competencies but a dynamic process of sense-making. We observe students repeatedly encountering ethical dilemmas, drawing on their growing technical and critical understanding of AI, and, in the process, forging a more personalized form of ethical readiness.

Consider the pervasive anxiety around AI detection. The quantitative phase showed a significant jump in "Societal–Ethical Orientation." Qualitatively, we saw this not as a simple memorization of rules, but as a literacy-

fueled debate. When a student questioned the fairness of a "0% plagiarism" score, the issue could not be resolved by ethical considerations alone. It was their collective literacy, their understanding of how detection tools work, their weaknesses, and the ways they can be manipulated, that allowed them to move beyond a binary of "cheating" or "not cheating." They were constructing an ethical stance that integrity is not about outsmarting a system but about fostering a culture of honesty, a stance made possible precisely because their literacy allowed them to deconstruct the technology's authority.

Similarly, the significant improvement in "Research and Development Ethics" observed in the quantitative phase was reflected in students' reports of having encountered "fictitious references." This was critical. The ethical principle of "transparency" is useless if a student lacks the literacy to recognize a fabricated citation. The student's dismay, "I did not know what to do," marks the very moment where ethics and literacy collide. The pedagogical response was not to reiterate the principle of honesty, but to build the operational literacy: teaching them how to cross-check sources, why LLMs hallucinate, and when it is permissible to use AI for formatting but not for sourcing. Here, literacy mediated the relationship by providing the tools to enact the ethical principle, transforming a moment of failure into one of empowered practice.

Finally, the "Pedagogical–Professional Preparedness" gains were crystallized in the reflective remark of a teacher-student: "What we struggle with here as graduate students—our own students will struggle with too." This is the feedback loop in action, driven by literacy. The quantitative model shows preparedness as an outcome; the qualitative data show it as a dialogic process. Their ethical reasoning was not developed in a vacuum but through applying their literacy to their dual roles. When they debated whether to "punish or guide" a high school student using ChatGPT, they were simultaneously acting as learners grappling with AI and as teachers formulating a pedagogical philosophy. Their literacy allowed them to see the connection between their own practices and their future responsibilities, making their preparedness immediately relevant and deeply personal.

In sum, the integration of findings reveals a crucial narrative. The 20-hour intervention did not simply add two separate sets of skills: technical and ethical. Instead, it created a generative space where each ethical dilemma demanded literacy for its resolution, and each act of literacy, in turn, deepened their ethical understanding. The mediation effect is thus not only a statistical figure but a representation of this lived, iterative process. AI literacy is the missing link that transforms ethical instruction from a syllabus mandate into a capacity for situated judgment, equipping these future educational leaders to navigate the ambiguous, and often uncharted, ethical terrain of AI in their own classrooms and institutions.

Conclusion

This study demonstrated that modest yet deliberate efforts to embed AI instruction in graduate education yielded substantial gains in both literacy and ethical preparedness. The quantitative results confirmed significant improvements across constructs, while classroom observations revealed how these changes unfolded in real time. When students were given structured opportunities to engage critically with AI, not merely as a technical tool but as a social and ethical phenomenon, they learned to question, apply, and act responsibly. When taught together, literacy and ethics provided the scaffolding for the urgent preparedness needed in a rapidly shifting educational landscape. As a pilot study, the research necessarily had limitations: it was confined to a single institution, relied on a modest cohort of 39 students, and did not include a control group. However, these constraints do not diminish its contribution. On the contrary, the intervention established proof of concept, demonstrated the feasibility of embedding AI literacy and ethics in existing coursework, and generated validated instruments that future research can build upon. More importantly, it provided fine-grained qualitative observations that larger studies might miss.

However, the findings do more than affirm the value of course-level interventions; they invite more profound reflection on what higher education in the Philippines should look like in an AI-driven era. At the policy level, the absence of CHED guidelines leaves institutions and faculty members to navigate AI integration independently. This autonomy enables innovation but also entails risks of uneven implementation. In this study, students' repeated questions—such as whether they were "allowed" to disclose AI use in research papers—exposed the uncertainty created by this policy vacuum. Without guidance, teachers are left to negotiate between fear of sanctions and the desire for transparency. The evidence here suggests that CHED and related agencies should not

delay in issuing frameworks for AI education—frameworks that move beyond a compliance mindset ("what to prohibit") toward enabling practices ("how to cultivate responsible engagement").

At the pedagogical level, the findings show that AI instruction does not require new courses or sweeping revisions. Integration within existing subjects proved sufficient to spark meaningful learning. What mattered was deliberate framing. Case-based discussions, peer reflections, and tool demonstrations can bridge the gap between abstract principles and concrete practice. Graduate education, particularly in teacher education, needs to prioritize this type of integration, recognizing the multiplier effect that teachers have on their students.

Overall, this study makes three contributions. Locally situated, it presents a Philippine-centered approach to AI literacy and ethical readiness among teachers, while remaining sensitive to digital divides, multilingualism, and uneven infrastructure. Conceptually, it reconceptualizes AI literacy not as an adjunct skill but as a mediating factor that facilitates ethical practice. As a pilot study, the findings are necessarily tentative. However, these are initial steps toward scaling. Extensive comparative studies will benefit from the proof of concept developed here, the tools piloted here, and the contextual knowledge gained in this research.

Contributions of Authors

The author solely contributed to the conception, design, data collection, analysis, and writing of this manuscript.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Conflict of Interests

The author declares no conflict of interest. The respondents in this study were the author's students at the Polytechnic College of La Union; however, all ethical protocols, including voluntary participation and confidentiality, were strictly observed to minimize potential bias.

Acknowledgment

The author extends sincere appreciation to the anonymous reviewers for their valuable feedback and to the students of the Polytechnic College of La Union, Master of Arts Major in Educational Management program, for their participation and support in this study.

References

Daher, R. (2025). Integrating AI literacy into teacher education: A critical perspective paper. Discover Artificial Intelligence, 5, 217. https://doi.org/10.1007/s44163-025-00475-7

Espinosa, A., Gomez, M.A., Miranda, P., David, A., Abulon, E.L., Hermosisima, M.V., Quinosa, E., Jr., Soliman, A., de Vera, J., Claros, I.H., Cruz, H.G., & Gonzales, N. (2025). Bridging a digital divide: A critical analysis of contextual factors affecting ICT integration in Philippine schools. Issues in Educational Research, 35(2): 526–549. http://www.iier.org.au/iier35/espinosa-2.pdf

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. Harvard Data Science Review, 1(1). https://doi.org/10.1162/99608f92.8cd550d1

Frieary, A. (2025). The integration of Artificial Intelligence into curriculum design in higher education (Applied Research Project). Purdue Global. https://purdueglobal.dspacedirect.org

Garrett, N., Beard, N., & Fiesler, C. (2020, February). More than "If Time Allows": The role of ethics in AI education. AAAI/ACM Conference on AI, Ethics, and Society. https://cmci.colorado.edu/~cafi5706/AIES_EthicsEducation.pdf

Matthews, A.S., & Bartley, B. (2025). Pay attention to the chatbot behind the curtain when AI 'is no place like home': A framework and toolkit for integrating critical thinking and information literacy in educational and professional settings. Advances in Online Education, 3(3):247–273. https://doi.org/10.69554/FMAI7138

Ng, D.T.K., Leung, J.K.L., Chu, K.W.S., & Qiao, M.S. (2022). AI literacy: Definition, teaching, evaluation and ethical issues. Proceedings of the Association for Information Science and Technology, 58(1): 504–509. https://doi.org/10.1002/pra2.487

Ning, Y., Zhang, W., Yao, D., Fang, B., Xu, B., & Wijaya, T.T. (2025). Development and validation of the Artificial Intelligence Literacy Scale for Teachers (AILST). Education and Information Technologies. https://doi.org/10.1007/s10639-025-13347-5

Noor, E., & Kanitroj, B. (2025). Speaking in code: Contextualizing large language models in Southeast Asia. Carnegie Endowment for International Peace. https://shorturl.at/cDM4d

Okoye, K., & Hosseini, S. (2024). Wilcoxon statistics in R: Signed-Rank Test and Rank-Sum Test. In: R Programming. Springer, Singapore. https://doi.org/10.1007/978-981-97-3385-9_13

Organisation for Economic Co-operation and Development. (2021). AI and the future of skills, volume 1: Capabilities and assessments. OECD Publishing. https://doi.org/10.1787/5ee71f34-en

Showemimo, P. (2025). The AI paradox in academia: When writing too well raises red flags. SSRN. https://doi.org/10.2139/ssrn.5255976

Sperling, K., Stenliden, L., Mannila, L., Hallström, J., Nordlöf, C., & Heintz, F. (2025). Perspectives on AI literacy in middle school classrooms: An integrative review. Postdigital Science and Education, 7: 719–749. https://doi.org/10.1007/s42438-025-00560-1

Taber, K. (2018). The use of Cronbach's Alpha when developing and reporting research instruments in Science Education. Research in Science Education, 48: 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

University of the Philippines. (2023). University of the Philippines principles for responsible and trustworthy Artificial Intelligence. https://tinyurl.com/mpz4sdjd

Villanueva, M. (2025). Towards national standards: Comparative analysis of AI governance models in Philippine higher education. Ignatian Journal for Multidisciplinary Research. https://doi.org/10.5281/zenodo.15212683

Wang, R., Li, J., Shi, W., & Li, X. (2021). Application of artificial intelligence techniques in operating mode of professors' academic governance in American research universities. Wireless Communications and Mobile Computing. https://doi.org/10.1155/2021/3415125