Asian Journal of Education and Training
Vol. 12, No. 2, 28-35, 2026
ISSN(E) 2519-5387
DOI: 10.20448/edu.v12i2.8714

Task-differentiated Generative AI tool selection among pre-service teachers: An observational study

Leonardo D. Tejano

Mariano Marcos State University, Laoag City, Philippines.
Email: ldtejano@mmsu.edu.ph

Abstract

Generative AI tools now populate higher education at a scale that has shifted the central challenge from simple adoption to strategic, task-aligned selection. Pre-service teacher education in the Philippine context remains underrepresented in longitudinal, task-level research on this phenomenon. This descriptive-quantitative study tracked AI tool selection by 103 pre-service teachers enrolled in action research courses at a Philippine state university across 25 sequential course tasks distributed over three academic phases within one semester. A structured, course-embedded observational tracker captured tool selections, use intensity ratings, multi-tool strategies, and AI-assisted content estimates. ChatGPT dominated every phase with 1,341 total uses, yet selection was meaningfully differentiated by task type. Elicit and Consensus AI-led literature synthesis tasks; Canva AI and Gamma AI dominated presentation work; Grammarly AI, Quillbot, and Writefull clustered around manuscript editing. A chi-square test confirmed a statistically significant association between academic phase and tool selection (chi-square [18] = 360.40, $p < 0.001$, Cramer's V = 0.20). AI use intensity increased monotonically across phases (Early: M = 2.00; Late: M = 3.86), and multi-tool usage rose from 43.1% to 58.6%. These patterns indicate that pre-service teachers exercise purposive, task-calibrated AI selection consistent with emergent domain-specific AI literacy. The findings carry direct implications for teacher education curricula and institutional AI policy within the Philippine higher education context.

Keywords: AI literacy, AI tool selection, Generative AI, Pre-service teachers, Teacher education.

Citation | Tejano , L. D. (2026). Task-differentiated Generative AI tool selection among pre-service teachers: An observational study. *Asian Journal of Education and Training, 12*(2), 28–35. 10.20448/edu.v12i2.8714

History:
Received: 10 April 2026
Revised: 13 May 2026
Accepted: 18 May 2026
Published: 22 May 2026

Publisher: Asian Online Journal Publishing Group

Funding: This study received no specific financial support.

Institutional Review Board Statement: The study involved minimal risk and followed ethical guidelines for social science fieldwork. Formal approval from an Institutional Review Board was not required under the policies of Institute for Research Ethics Committee of the [Mariano Marcos State University, Philippines]. Informed verbal consent was obtained from all participants, and all data were anonymized to protect participant confidentiality.

Transparency: The author confirms that the manuscript is an honest, accurate, and transparent account of the study; that no vital features of the study have been omitted; and that any discrepancies from the study as planned have been explained. This study followed all ethical practices during writing.

Competing Interests: The author declares that there are no conflicts of interest regarding the publication of this paper.

Disclosure of AI Use: The author used OpenAI's ChatGPT (GPT-4o) and Grammarly AI to assist with language editing and proofreading of select sections of this manuscript. All AI-generated suggestions were reviewed, evaluated, and revised by the author, who takes full responsibility for the published content. The use of these tools was limited to linguistic refinement and did not extend to the generation of empirical data, statistical analyses, or substantive intellectual claims. AI-assisted content within this manuscript does not exceed 15% of the total text, in compliance with this journal's publication guidelines.

Contents

Contribution of this paper to the literature
This manuscript represents original scholarly work. It has not been previously published, is not under consideration for publication elsewhere, and does not reproduce, in whole or in substantial part, any previously published material without appropriate citation and permission. The empirical data reported were collected specifically for this study. All sources used have been properly cited in accordance with the journal's citation guidelines.

1. Introduction

Since the public release of ChatGPT in late 2022, generative artificial intelligence (GenAI) has entered higher education at a pace that challenges both policy and pedagogy. Aggregate adoption data, however, masks sharp variation in how students actually deploy these tools. Wang and Fan (2025) showed in a meta-analysis that ChatGPT's effects on learning are heterogeneous across task types. Bond et al. (2024) cautioned that summary adoption metrics conceal qualitatively distinct ways in which different tools mediate different forms of academic work. Liu and Zhong (2025) found that pedagogically effective AI use requires disciplinary and task-specific knowledge that most current curricula have not yet systematically addressed. The central challenge for higher education is therefore not whether GenAI will be adopted, but whether students navigate the tool landscape with strategic competency or indiscriminate convenience.

Task-differentiated AI use, the documented tendency of students to select different tools depending on the nature of the academic task, has begun to attract dedicated theoretical attention. Hershkovitz et al. (2025) proposed a Task-Centered GenAI Literacy Framework grounded in Bloom's cognitive taxonomy, arguing that genuine AI literacy requires mapping specific tool affordances to specific task demands. Adamakis and Rachiotis (2025) similarly argued that AI literacy must encompass critical evaluation of tool affordances rather than operational familiarity alone. Singer-Freeman et al. (2025) found that GenAI usage varied substantially by academic level and task type. Li et al. (2026), in a systematic review and meta-analysis, found that learning outcomes were significantly moderated by tool-task alignment, with misaligned deployments yielding null or negative effects.

The Philippine context occupies a specific and underexamined position within this discourse. Quimba et al. (2024) documented that while enthusiasm for AI innovation is high, infrastructural disparities and variable digital literacy pose substantial challenges. Malacapay (2025) found that AI use motives are high among Filipino pre-service teachers, but formal preparedness remains inconsistent. The bibliometric analysis of Kuzu (2025) identified the underrepresentation of Southeast Asian and Philippine contexts as a substantive methodological gap, a point reinforced by Belkina et al. (2025), whose review drew predominantly from high-income, English-dominant settings. Ziying et al. (2026) separately identified the absence of longitudinal, task-tracking designs as a priority in the pre-service teacher education literature.

Three complementary theoretical frameworks anchor this study. Cognitive Load Theory (Sweller, 1988; Sweller, van Merrienboer, & Paas, 1998) provides the lens for understanding how students strategically offload different types of cognitive burden (extraneous, intrinsic, and germane) to different AI tools (Gerlich, 2025; Yao & Fan, 2025). An extended Technology Acceptance Model (Davis, 1989), incorporating task-technology fit, conceptualizes tool selection as a function of task-specific perceived usefulness (Mustofa et al., 2025; Saflor, 2025). The Task-Centered GenAI Literacy Framework of Hershkovitz et al. (2025) maps GenAI competencies directly to Bloom's cognitive domains, positing that purposive tool selection across task types is the primary behavioral indicator of domain-specific AI literacy.

Very few studies examine within-semester, task-level AI tool selection with sufficient temporal granularity to reveal which tools are chosen for which tasks across a full academic cycle. Most studies use cross-sectional surveys that treat 'AI use' as synonymous with 'ChatGPT use,' overlooking the rich ecosystem of specialized tools increasingly available to students (Gabay et al., 2026). The present study addresses this gap through a semester-long observational tracker, generating a Task x Tool Matrix across 25 course tasks. This design provides a granular, empirical analysis of how pre-service teachers differentiate their AI tool selection in response to evolving academic demands at a Philippine state university.

This study addresses four research questions:

1. What AI tools are most frequently selected by pre-service teachers for each of the three academic task phases?
2. Is there a statistically significant association between task type and the AI tool(s) selected?
3. How does AI tool diversity and complexity evolve across the 25 course tasks from the Early to Late phase of the semester?
4. What cognitive functions do students appear to delegate to AI on task-tool pairings?

2. Methods

2.1. Study Design, Population, and Setting

This study employed a descriptive-quantitative research design with exploratory mixed-methods elements, utilizing a non-experimental observational approach. A semester-long naturalistic tracking method captured the dynamic nature of AI tool selection, thereby avoiding the retrospective biases inherent in cross-sectional survey designs.

The research was conducted at the College of Teacher Education, Mariano Marcos State University (MMSU), Philippines, during the second semester of Academic Year 2024-2025, within the courses SNED 131: Research in Special Needs and Inclusive Education and ECED 119: Research in Early Childhood Education. The participant pool consisted of an intact cohort of 103 pre-service teachers. This course was selected because its curriculum requires students to execute a comprehensive action research project encompassing a wide spectrum of academic tasks, from initial literature exploration to empirical data analysis and formal manuscript defense, providing a robust environment for observing task-differentiated AI use.

2.2. Study Tools, Variables, and Data Collection

The primary data source was a structured, course-embedded observational record in which each student documented the AI tools used, estimated the percentage of AI-assisted content, AI use intensity ratings, and navigation strategies for each of the 25 academic tasks completed throughout the 18-week semester. The 25 tasks were distributed across three phases corresponding to the sequential stages of an action research project.

The Early Phase (Tasks 1-9, Weeks 1-6) comprised: Reflective Journal Entry; Short Essay (Action Research); Diagram Creation (AR Cycle); Observation Protocol Draft; Ethics Checklist; Initial Research Topic and Problem; Annotated Bibliography; Revised Problem Statement; and Draft Intervention Plan. The Middle Phase (Tasks 10-16, Weeks 7-12) comprised: Research Instrument Draft; Instrument Alignment Matrix; Draft Methodology Section; Action Research Proposal (Chapters 1-3); Implementation Readiness Packet; Reflective Journal (Implementation); and Fidelity Checklist Self-Assessment. The Late Phase (Tasks 17-25, Weeks 13-18) comprised: Data Management Plan; Progress Report on Data Collection; Coded Data Sample and Descriptive Statistics; Initial Findings Draft; Revised Findings and Data Visuals; Discussion and Conclusion Draft; Full Manuscript Draft; Final Research Presentation; and Final Action Research Manuscript.

AI use intensity was rated on a five-point scale: 1 (Minimal: Occasional reference, no content adoption); 2 (Light: several queries for ideation, no direct content adoption); 3 (Moderate: AI scaffolding that shaped structure with substantial student reworking); 4 (Substantial: AI-generated content forming the majority of the draft before significant editing); and 5 (Extensive: AI output forming the primary structural basis of the submission). Documentation was completed immediately following or concurrent with task completion to reduce retrospective recall bias. Data were reviewed at three structured checkpoint meetings corresponding to the end of each academic phase.

2.3. Data Analysis

For RQ1, frequency counts and percentage distributions were computed for each AI tool across all 25 tasks, grouped into three phases, and top-tool rankings were established based on absolute use counts. For RQ2, a Pearson chi-square test of independence was conducted using a contingency table comprising the top 10 most frequently used AI tools across the three phases. Task type was operationalized as academic phase (Early, Middle, Late) rather than individual task, as phase-level aggregation produced stable and interpretable contingency cell frequencies. Effect size was estimated using Cramer's V, interpreted following Cohen's (1988) benchmarks.

For RQ3, two complementary indices were computed for each of the 25 tasks: multi-tool usage percentage (proportion of students using two or more tools) and Shannon's entropy index (H' = -sum [pi x ln(pi)]; (Shannon, 1948)), a distributional diversity measure from information theory. AI use intensity trajectories were aggregated by phase. For RQ4, a theoretically derived cognitive function taxonomy was developed a priori, drawing on the tool affordance literature and the Task-Centered GenAI Literacy Framework of Hershkovitz et al. (2025). Each observed task-tool pairing was coded according to the cognitive function the tool's affordances most plausibly served. All statistical analyses were conducted in Python 3.11 using the SciPy and pandas libraries.

2.4. Ethical Considerations

The study adhered to stringent ethical protocols. All student data within the AI Tool Usage Tracker was pseudonymized prior to analysis to ensure confidentiality. Students were fully informed that the tracking data would be utilized for both course pedagogical improvement and scholarly research, and informed consent was obtained. The course operated under a transparent institutional policy permitting up to 20% AI-generated content in academic submissions, provided it was properly cited and utilized for scaffolding rather than outright generation. No deception was employed, and the tracking mechanism was subject to institutional course oversight.

3. Results

3.1. RQ1: AI Tool Frequency by Academic Phase

Analysis of tool-use records revealed distinct and interpretively meaningful patterns of AI tool selection across the three academic phases. Table 1 presents aggregate frequency counts of the top AI tools used by 103 pre-service teachers during each phase.

Table 1. Frequency of AI tool selections by academic phase.

AI Tool	Early Phase (Tasks 1-9)	Middle Phase (Tasks 10-16)	Late Phase (Tasks 17-25)	Total Uses
ChatGPT	404	418	519	1,341
Google Gemini	201	231	202	634
Grammarly AI	188	172	249	609
Claude (Anthropic)	86	117	265	468
Microsoft Copilot	81	177	82	340
Quillbot	90	81	128	299
Perplexity AI	135	46	22	203
Canva AI	45	39	106	190
Writefull	57	27	96	180
NotebookLM	22	55	83	160
Consensus AI	66	--	--	66
Elicit	80	--	--	80
Gamma AI	--	19	59	78

ChatGPT maintained its position as the dominant foundational tool across all phases, rising from 404 instances in the Early Phase to 519 in the Late Phase (1,341 total uses). Aggregate frequency figures, however, understate the purposive differentiation observable in phase-specific trajectories. Table 2 highlights the percentage of the student cohort selecting specific tools for representative tasks, documenting clear task-differentiated behavior.

Table 2. Task-Specific AI tool selection highlights (% of N = 103).

Representative Task	Dominant AI Tools Selected (% of N = 103)
Reflective Journal Entry	Grammarly AI (36.9%), Google Gemini (29.1%), Quillbot (26.2%), ChatGPT (24.3%)
Short Essay (Action Research)	Grammarly AI (48.5%), ChatGPT (40.8%), Quillbot (29.1%)
Diagram Creation (AR Cycle)	Canva AI (43.7%), ChatGPT (32.0%), Google Gemini (28.2%)
Initial Research Topic & Problem	ChatGPT (58.3%), Google Gemini (34.0%), Perplexity AI (33.0%), Elicit (24.3%)
Annotated Bibliography	Elicit (53.4%), Perplexity AI (49.5%), Consensus AI (47.6%), NotebookLM (21.4%)
Final Research Presentation	Canva AI (59.2%), Gamma AI (57.3%), ChatGPT (47.6%)
Discussion & Conclusion Draft	ChatGPT (67.0%), Claude (48.5%), Grammarly AI (44.7%)
Full Manuscript Draft	ChatGPT (63.1%), Grammarly AI (54.4%), Claude (48.5%), Quillbot (35.9%)
Final Action Research Manuscript	ChatGPT (66.0%), Claude (53.4%), Grammarly AI (46.6%), Writefull (42.7%)

For literature synthesis tasks such as the Annotated Bibliography, specialized academic search tools dominated: Elicit (53.4%) and Consensus AI (47.6%) were far more heavily selected than any general-purpose chatbot. Visual and presentation tasks produced a markedly different profile, with Canva AI and Gamma AI covering 59.2% and 57.3% of the cohort, respectively, for the Final Research Presentation. Editing-intensive tasks such as the Short Essay and the Full Manuscript triggered consistent co-selection of Grammarly AI and Quillbot alongside ChatGPT, indicating a sequential workflow in which generative drafting and linguistic polishing functioned as distinct cognitive operations requiring different tools.

Cross-phase tool trajectories carry additional interpretive weight. Perplexity AI declined from 135 uses in the Early Phase to 22 in the Late Phase, consistent with its role being tied to exploratory topic-scoping rather than sustained research support. Claude (Anthropic) escalated from 86 Early Phase uses to 265 in the Late Phase, reflecting students' assignment of increasingly demanding reasoning and argumentation tasks to this platform as manuscript-stage writing intensified. The near-absence of Elicit and Consensus AI in the Middle and Late phases aligns with their affordances being narrowly suited to literature synthesis demands.

3.2. RQ2: Association Between Task Type and Tool Selection

To determine whether observed variations in tool selection across the three academic phases were statistically significant, a Pearson chi-square test of independence was conducted. The analysis used a contingency table comprising the top 10 most frequently used AI tools across the Early, Middle, and Late phases (Table 3).

Table 3. Top 10 AI tools by academic phase.

AI Tool	Early Phase	Middle Phase	Late Phase
ChatGPT	404	418	519
Google Gemini	201	231	202
Grammarly AI	188	172	249
Claude (Anthropic)	86	117	265
Microsoft Copilot	81	177	82
Quillbot	90	81	128
Perplexity AI	135	46	22
Canva AI	45	39	106
Writefull	57	27	96
NotebookLM	22	55	83

Note: Chi-square (18) = 360.40, $p < 0.001$, Cramer's V = 0.20, N = 4,661 total tool-use instances.

The chi-square test yielded a highly significant result, chi-square (18) = 360.40, $p < 0.001$, confirming that tool selection across the three phases was not random. The effect size estimated using Cramer's V was .20, a small-to-moderate effect under Cohen's (1988) benchmarks. In a 10 x 3 contingency table with N = 4,661 total tool-use instances, this value represents a substantively meaningful effect because phase-level aggregation necessarily compresses within-phase heterogeneity, rendering the obtained effect size a conservative estimate of actual task-type differentiation at the individual task level.

Raw frequency trajectories offer additional specificity. Perplexity AI showed a sharp phase-related decline (135, 46, 22), indicating concentration in exploratory topic-scoping behavior. Microsoft Copilot exhibited a distinctive Middle Phase peak (81, 177, 82), corresponding with tasks related to methodology drafting and proposal formatting. Claude (Anthropic) demonstrated a dramatic late-phase escalation (86, 117, 265), consistent with growing reliance on extended reasoning capabilities for higher-order writing. Visual design tools peaked significantly in the Late Phase, aligning with final presentation requirements. NotebookLM showed a steady rise (22, 55, 83), consistent with data analysis tasks requiring source-grounded comprehension.

3.3. RQ3: Evolution of AI Tool Diversity and Complexity Across 25 Tasks

The evolution of AI tool diversity and students' technological ecosystems was quantified using Shannon's entropy index (H'), the percentage of students employing multi-tool strategies, and self-reported AI use intensity ratings. Table 4 presents the phase summary metrics.

Table 4. AI tool diversity and complexity metrics by phase.

Phase	Avg Unique Tools per Task	Avg Multi-Tool %	Avg Shannon H'	Total AI Uses
Early Phase	5.00	43.1%	1.544	1,468
Middle Phase	4.78	36.8%	1.513	1,382
Late Phase	4.78	58.6%	1.511	1,811

Note: H' = Shannon's entropy index; higher values indicate more evenly distributed tool use across multiple tools.

The phase summary data reveal a trajectory that resists simple linear description. While the average number of unique tools per task remained relatively stable across phases (5.00 in Early; 4.78 in Late), the proportion of students employing multi-tool strategies rose from 43.1% in the Early Phase to 58.6% in the Late Phase. This divergence is consistent with the interpretation that the overall tool palette stabilized as students developed familiarity with the available ecosystem, while the depth of per-task multi-tool deployment increased in response to rising task complexity.

A detailed analysis of per-task data reveals patterns that phase-level averages do not fully capture. High multi-tool usage is concentrated in cognitively demanding research writing tasks: the Annotated Bibliography reached 68.9% multi-tool use, the Draft Intervention Plan reached 62.1%, the Action Research Proposal recorded 65.0%, and the Final Action Research Manuscript peaked at 78.6% multi-tool use with a Shannon entropy of H' = 1.590. Self-reported AI use intensity demonstrated a monotonic increase throughout the semester: mean intensity rose from 2.00 (SD = 0.78) in the Early Phase to 3.16 (SD = 1.05) in the Middle Phase and reached 3.86 (SD = 1.08) in the Late Phase. Trend analysis indicated that 92.2% of pre-service teachers exhibited an increasing trend in AI reliance as the semester progressed; 7.8% remained stable; none showed a decreasing trend.

3.4. RQ4: Cognitive Functions Delegated to AI on Task-Tool Pairings

By cross-referencing empirical tool selection frequencies with the cognitive demands of the 25 course tasks and applying the a priori cognitive function taxonomy, five dominant patterns of cognitive delegation were identified within pre-service teachers' AI workflows.

The first pattern, generative-synthetic delegation, was anchored by the co-deployment of ChatGPT and Claude for extended manuscript writing. In drafting the full manuscript, 63.1% of students utilized ChatGPT for structural generation, while 48.5% employed Claude to refine the depth of argumentation. Interpreted through Cognitive Load Theory (Sweller, 1988; Sweller et al., 1998), this pairing reflects a strategic decomposition of germane cognitive load: ChatGPT absorbed organizational and generative demands, while Claude's extended-context reasoning capacities managed higher-order inferential synthesis.

The second pattern, literature intelligence delegation, involved the tripartite cluster of Elicit, Consensus AI, and Perplexity AI deployed specifically for the Annotated Bibliography and related bibliographic tasks. These tools' affordances for structured academic search, citation verification, and claim-evidence mapping served to manage the high intrinsic cognitive load of systematic bibliographic work, a task type for which general-purpose chatbots have well-documented limitations, including confabulated citations (Abubakar et al., 2025). The sharp discontinuation of both tools after the Early Phase is consistent with the hypothesis that their utility is task-narrow rather than broadly applicable.

The third pattern, visual cognition delegation, involved pairing Canva AI and Gamma AI for presentation and diagrammatic tasks. For the Final Research Presentation, 59.2% of students selected Canva AI, and 57.3% selected Gamma AI. These tools offload the multimodal representation demands described by Gerlich (2025): the cognitive work of translating abstract research findings into visually coherent slides and infographics. Canva AI appeared as early as the Diagram Creation task in the Early Phase, suggesting that visual cognition delegation was established as a behavioral pattern from the outset of the semester.

The fourth pattern, language polishing delegation, was enacted through the co-deployment of Grammarly AI, Quillbot, and Writefull across all three phases, with particular concentration in late-phase manuscript tasks. For the Final Action Research Manuscript, 46.6% of students selected Grammarly AI, 42.7% selected Writefull, and a subset also engaged Quillbot. These tools operate as a meta-linguistic editing stack, absorbing the extraneous cognitive load of syntax correction, grammatical compliance, and academic register enforcement, consistent with Sweller's (1988) distinction between extraneous and germane cognitive load.

The fifth pattern, responsible AI adherence, was evident across all tasks. Despite increasing intensity and complexity of AI tool usage, all 103 students stayed within the course's 20% AI content policy, with a mean estimated AI content of 12.8% (SD = 4.4%). This indicates that pre-service teachers mainly used AI tools for scaffolding, ideation, structural organization, and linguistic refinement, activities that boost productivity while maintaining authorial voice and intellectual ownership (Adamakis & Rachiotis, 2025).

4. Discussion

4.1. ChatGPT as the Default Cognitive Partner

The empirical data clearly establish ChatGPT as the foundational node within the pre-service teachers' AI ecosystems, maintaining dominance across all three academic phases with 1,341 total uses. This sustained prevalence aligns with findings by Zhao (2025) and Lendvai (2025), who noted that ChatGPT's versatility and conversational fluidity position it as an optimal baseline tool for academic work. The current study extends this understanding by demonstrating that ChatGPT is rarely used in isolation for complex tasks. Instead, it functions as a default cognitive partner, serving as the primary interface for initial brainstorming, structural outlining, and general synthesis, before students pivot to more specialized tools. This behavior suggests that while pre-service teachers recognize the broad utility of ChatGPT, they also acknowledge its limitations in specialized domains such as rigorous literature mapping or deep academic argumentation.

4.2. Phase-Dependent Tool Specialization

A critical finding of this study is the highly significant association between academic phase and tool selection (chi-square [18] = 360.40, $p < 0.001$), which underscores the dynamic nature of students' epistemic needs. The sharp decline in Perplexity AI usage (135 to 46 to 22) illustrates this phenomenon: as students transitioned from exploratory topic scoping to rigorous empirical synthesis, the utility of a broad search-oriented AI diminished. The deployment of Elicit and Consensus AI during the Annotated Bibliography reflects a sophisticated literature intelligence delegation strategy, corroborating Mtotywa et al. (2026), who emphasized the necessity of specialized AI for methodological literature reviews. The mid-semester peak of Microsoft Copilot aligns with proposal formatting demands, while the late-phase escalation of Claude (Anthropic) indicates a strategic shift toward tools capable of handling the high cognitive load of data interpretation and conclusion drafting. This phase-dependent specialization demonstrates that pre-service teachers continuously recalibrate their tool selections to match the evolving cognitive architecture of their research projects.

4.3. Tool Complexity Escalation as Proxy for Academic Socialization

The longitudinal data reveal a clear escalation in both AI use intensity (from a mean of 2.00 to 3.86) and multi-tool strategy deployment (from 43.1% to 58.6%). This monotonic escalation can be interpreted as a proxy for academic socialization within the GenAI era. As pre-service teachers progressed through the rigorous curriculum, their technological behaviors matured from simple, single-prompt interactions to complex, multi-tool chaining. This progression mirrors the findings of Gabay et al. (2026) and Wu et al. (2026), who observed that advanced academic writing requires the orchestration of multiple AI applications. The high Shannon entropy values recorded during late-stage manuscript drafting (e.g., $H' = 1.590$ for the Final Action Research Manuscript) mathematically quantify this diversity, indicating that the cohort developed a rich and varied technological repertoire as intrinsic task difficulty increased.

4.4. Cognitive Delegation as an Emergent AI Literacy Construct

The identification of five dominant cognitive delegation patterns provides robust empirical support for the theoretical frameworks guiding this study. Viewed through Cognitive Load Theory (Sweller, 1988; Yao & Fan, 2025), the language polishing delegation pattern, utilizing Grammarly AI, Quillbot, and Writefull, represents a strategic offloading of extraneous cognitive load, preserving working memory capacity for the intrinsic load of empirical argumentation. These delegation patterns also operationalize the Task-Centered GenAI Literacy Framework (Hershkovitz et al., 2025): the pre-service teachers' capacity to accurately map specific tools to specific cognitive domains demonstrates functional, domain-specific AI literacy that transcends basic operational competence. The cohort's responsible AI adherence, with all students within the 20% content policy, and an average AI-assisted content of 12.8%, further suggests that high-frequency, multi-tool AI use does not inherently constitute academic dishonesty but rather reflects scaffolding and linguistic refinement aligned with the policy framework (Adamakis & Rachiotis, 2025; Jin et al., 2025).

4.5. Implications for Teacher Education and Institutional Policy

The findings necessitate a paradigm shift in how educational technology is integrated into teacher preparation programs. Generic prompting workshops are insufficient for developing the AI literacy required for modern academic work. Teacher education curricula must evolve to explicitly teach task-differentiated AI selection strategies, exposing pre-service teachers to a diverse ecosystem of specialized tools and training them to evaluate tool affordances and limitations in relation to specific cognitive tasks. As these pre-service teachers transition into professional educators, their capacity to model nuanced, task-aligned AI use will shape the digital literacy of their future students. AI literacy should therefore be embedded as a core component of pedagogical content knowledge, as advocated by Liu and Zhong (2025).

Within the Philippine higher education context, the study provides empirical grounding for the development of nuanced, task-sensitive AI policies. The successful adherence to the 20% AI content threshold demonstrates that students can engage deeply with AI technologies while maintaining academic integrity, provided that parameters are clear and pedagogically aligned. Institutional policies guided by Commission on Higher Education (CHED) directives should move beyond blanket prohibitions to acknowledge the multi-tool reality of student workflows and provide specific guidelines on acceptable cognitive delegation practices.

The generalizability of these findings must be interpreted with caution. The tool selection patterns observed at MMSU, a state university whose students operate under the infrastructural and economic constraints documented by Quimba et al. (2024), may differ substantially from those that would emerge in higher-resource institutional contexts. In settings characterized by reliable broadband connectivity, wider access to premium tool subscriptions, and more systematic pre-entry AI literacy preparation, the diversity and sophistication of tool deployment may be markedly elevated. Conversely, in more severely under-resourced Philippine schools where variable internet connectivity and subscription costs remain prohibitive (Quimba et al., 2024; Rodrigo & Talandron-Felipe, 2024), the tool ecosystem may be even more constrained than that observed here, with students relying disproportionately on free-tier general-purpose chatbots. These differential resource contexts imply that the Task x Tool Matrix should be understood as context-dependent rather than universally applicable. Future multi-institutional, comparative studies spanning diverse socioeconomic and infrastructural settings are necessary to determine which aspects of task-differentiated AI literacy are robust across resource gradients and which are shaped by access conditions specific to the Philippine context. Institutions must also consider promoting open-access alternatives to premium tools and advocating for institutional licensing that levels the technological playing field (Funa & Gabay, 2025).

5. Conclusion

This study provides a comprehensive empirical mapping of task-differentiated AI tool selection among 103 pre-service teachers across 25 course tasks over a full academic semester, demonstrating a decisive shift from monolithic AI reliance to sophisticated, multi-tool cognitive delegation. The Task x Tool Matrix analysis reveals that while ChatGPT remained the foundational interface, pre-service teachers strategically deployed specialized platforms (Elicit for literature synthesis, Claude for deep argumentation, Canva AI, and Gamma AI for visual presentation) in response to the evolving cognitive demands of their academic workflows. The statistically significant association between task phase and tool selection (chi-square [18] = 360.40, $p < .001$, Cramer's $V = .20$), coupled with the monotonic escalation in multi-tool usage (43.1% to 58.6%) and AI use intensity ($M = 2.00$ to $M = 3.86$), indicates the emergence of a highly functional, domain-specific AI literacy. This complex technological engagement was achieved while students maintained adherence to institutional academic integrity policies, a finding that carries particular weight given the frequent assumption that high AI use intensity correlates with policy non-compliance.

Philippine teacher education institutions, in alignment with CHED policy directives, are strongly recommended to develop and implement task-sensitive AI tool integration frameworks. Curricula must transcend basic generative prompting to explicitly train pre-service teachers in the critical evaluation and strategic selection of specialized AI ecosystems. National AI literacy standards for pre-service teachers are urgently needed to ensure that the next generation of educators is equipped not only to use AI tools for their own academic productivity but to ethically and effectively integrate them into the pedagogical fabric of the Philippine basic education system (Malacapay, 2025; Sinsay-Villanueva et al., 2025).

Several limitations must be acknowledged. The single-institution, single-course design limits immediate generalizability to other academic disciplines or institutional contexts. The specific infrastructural and economic realities of the Philippine context may produce selection patterns that differ substantially from those observed in more highly-resourced settings.

A second limitation concerns reliance on self-reported entries within the AI Tool Usage Tracker. While three structured checkpoint meetings were employed to verify submitted entries, social desirability bias cannot be fully discounted. Pressures may have operated bidirectionally: students may have under-reported tools or use intensity to appear less reliant on AI; conversely, in an environment that transparently permitted and monitored AI use, some students may have over-reported tool diversity to signal technological competence. The mean AI-assisted content figure of 12.8%, which falls within the institutional 20% threshold, should be interpreted with this caveat in mind. Future studies should consider supplementary verification strategies such as timestamped digital logs, browser extension usage trackers, or anonymous parallel reporting mechanisms.

A third limitation pertains to the metric used to estimate AI-generated content. AI-assisted content proportions were derived from instructor-approximated and student-reported estimations rather than from objective algorithmic detection tools. This introduces a degree of measurement imprecision. Future iterations of this research design would benefit from incorporating a parallel comparison between subjective estimates and the outputs of multiple AI detection tools (e.g., Turnitin AI, GPTZero, Copyleaks), with full acknowledgment of each tool's known error profile. The study also lacks a control or comparison group, and the rapid evolution of GenAI platforms necessitates continuous replication. Future research should employ quasi-experimental or longitudinal designs spanning multiple semesters and institutional contexts to validate these observational findings.

References

Abubakar, S., Jeilani, A., & Yusuf, M. (2025). The role of over-reliance on AI in the negative consequences of student learning: The moderating effects of ethical concerns and institutional policies. *Cogent Education, 12*(1), 2591503. https://doi.org/10.1080/2331186X.2025.2591503

Adamakis, M., & Rachiotis, T. (2025). Artificial intelligence in higher education: A state-of-the-art overview of pedagogical integrity, artificial intelligence literacy, and policy integration. *Encyclopedia, 5*(4), 180. https://doi.org/10.3390/encyclopedia5040180

Belkina, M., Daniel, S., Nikolic, S., Haque, R., Lyden, S., Neal, P., . . . Hassan, G. M. (2025). Implementing generative AI (GenAI) in higher education: A systematic review of case studies. *Computers and Education: Artificial Intelligence, 8*, 100407. https://doi.org/10.1016/j.caeai.2025.100407

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., . . . Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour. *International Journal of Educational Technology in Higher Education, 21*(1), 4. https://doi.org/10.1186/s41239-023-00436-z

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (*2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Funa, A. A., & Gabay, R. A. E. (2025). Policy guidelines and recommendations on AI use in teaching and learning: A meta-synthesis study. *Social Sciences & Humanities Open, 11*, 101221. https://doi.org/10.1016/j.ssaho.2024.101221

Gabay, R. A. E., Funa, A. A., & Ricafort, J. D. (2026). Generative artificial intelligence (GenAI) for academic writing in higher education: A scoping review of applications, challenges, and implications. *International Journal of Education in Mathematics, Science and Technology, 14*(1), 200-232.

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies, 15*(1), 6. https://doi.org/10.3390/soc15010006

Hershkovitz, A., Tabach, M., Reich, Y., Lurie, L., & Cholcman, T. (2025). Framing and evaluating task-centered generative artificial intelligence literacy for higher education students. *Systems, 13*(7), 518. https://doi.org/10.3390/systems13070518

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence, 8*, 100348. https://doi.org/10.1016/j.caeai.2024.100348

Kuzu, S. (2025). Artificial intelligence in pre-service teacher education: Bibliometrics analysis. *Pedagogical Research, 10*(4), em0249. https://doi.org/10.29333/pr/17402

Lendvai, G. F. (2025). ChatGPT in academic writing: A scientometric analysis of literature published between 2022 and 2023. *Journal of Empirical Research on Human Research Ethics, 20*(3), 131-148. https://doi.org/10.1177/15562646251350203

Li, Y., Shadiev, R., & Chiu, T. K. F. (2026). Applying generative artificial intelligence to task-based language teaching and learning: A systematic review and meta-analysis. *TechTrends, 70*, 150–163. https://doi.org/10.1007/s11528-025-01140-7

Liu, X., & Zhong, B. (2025). Integrating generative artificial intelligence into student learning: A systematic review from a TPACK perspective. *Educational Research Review, 49*, 100741. https://doi.org/10.1016/j.edurev.2025.100741

Malacapay, M. (2025). AI literacy, preparedness, and use motives of elementary education pre-service teachers of a state university in the Philippines. *Edelweiss Applied Science and Technology, 9*(9), 1120-1136. https://doi.org/10.55214/2576-8484.v9i9.10059

Mtotywa, M. M., Mowers, J.-L. J., Ndou, W., Moleko, T. V. Q., & Ledwaba, M. J. (2026). Artificial intelligence in literature review synthesis: A step-by-step methodological approach for researchers and academics. *Informatics, 13*(3), 43. https://doi.org/10.3390/informatics13030043

Mustofa, R. H., Kuncoro, T. G., Atmono, D., Hermawan, H. D., & Sukirman. (2025). Extending the technology acceptance model: The role of subjective norms, ethics, and trust in AI tool adoption among students. *Computers and Education: Artificial Intelligence, 8*, 100379. https://doi.org/10.1016/j.caeai.2025.100379

Quimba, F. M. A., Moreno, N. I. S., & Salazar, A. M. C. (2024). *Readiness for AI adoption of Philippine business and industry: The government's role in fostering innovation- and AI-driven industrial development.* PIDS Discussion Paper Series No. 2024-35 Philippine Institute for Development Studies.

Rodrigo, M. M. T., & Talandron-Felipe, M. M. P. (2024). *Using artificial intelligence to support basic education teachers in under-resourced contexts (UP CIDS Monograph Series 2024-10).* Philippines: University of the Philippines Center for Integrative and Development Studies.

Saflor, C. S. (2025). Insights from higher education AI integration initiatives: Looking beyond acceptance. *Social Sciences & Humanities Open, 12*, 101454.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal, 27*(3), 379-423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Singer-Freeman, K. E., Verbeke, K., & Barre, B. (2025). Generative AI usage among university students depends on academic level and task. *Higher Learning Research Communications, 15*(2), 1-25.

Sinsay-Villanueva, J. M., Fermin, E. M., Fernando, R. T., Marzon, L., Balasa, E. L., Dela Cruz, B. F. A., . . . Zosa, R. V. (2025). *Quality education starting with teacher education: The state of teacher education in the Philippines*. PIDS Discussion Paper Series No. 2025-04 Philippine Institute for Development Studies.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257-285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J., van Merrienboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. https://doi.org/10.1023/A:1022193728205

Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications, 12*(1), 621.

Wu, C., Moorhouse, B. L., Wan, Y., & Wu, M. (2026). Exploring PhD students' utilization of generative AI in academic writing for publication purposes: Insights for EAP. *Journal of English for Academic Purposes, 79*, 101612. https://doi.org/10.1016/j.jeap.2025.101612

Yao, G., & Fan, L. (2025). Cognitive load scale for AI-assisted L2 writing: Scale development and validation. *Frontiers in Psychology, 16*, 1666974. https://doi.org/10.3389/fpsyg.2025.1666974

Zhao, W. (2025). Enhancing academic writing efficiency with ChatGPT: A natural language processing framework for innovation, opportunities, and challenges. *Journal of ICT Standardization, 13*(4), 359-382. https://doi.org/10.13052/jicts2245-800X.1341

Ziying, L., Yongchun, H., & Qiaoping, Z. (2026). Harnessing artificial intelligence for preservice teachers' development: A scoping review of applications, benefits, and challenges. *Computers and Education Open, 10*, 100330. https://doi.org/10.1016/j.caeo.2026.100330