Contextualizing a Proposed AI Governance Policy Framework within the Communication Program of St. Paul University Manila

Brian Bantugan, PhD

St. Paul University Manila

DOI: https://doi.org/10.51584/IJRIAS.2026.11050061

Received: 04 May 2026; Accepted: 10 May 2026; Published: 29 May 2026

ABSTRACT

This study investigated the dualistic role of artificial intelligence (AI) in shaping critical thinking within educational contexts, specifically focusing on the Communication Program at St. Paul University Manila. Employing a qualitative, interpretive research design, the study utilizes comparative discourse analysis of contradicting YouTube transcripts alongside a synthetic analysis of scholarly article by Brian Bantugan that proposed a policy framework on AI governance in St. Paul University Manila. The analytical framework integrates Cognitive Hierarchy Theory, Sociocultural Theory, and Cognitive Offloading Theory to move beyond binary perspectives of advocacy or resistance. Findings reveal two primary paradigms: AI as "cognitive augmentation," where the technology reconfigures intellect toward high-level task stewardship and information verification, and AI as "cognitive substitution," where overreliance leads to "passive consumption" and cognitive atrophy. Results indicate that while AI can enhance higher-order analysis, it risks "short-circuiting" the foundational mental effort required for deep learning. Consequently, the study proposes a progressive "Learn First, Augment Later" pedagogical model. In Basic Education, AI should serve as a limited "cognitive scaffold" to protect foundational skill formation. In Higher Education, AI transitions into a "cognitive partner" for advanced synthesis, provided learners maintain active agency. These insights culminate in a modified AI governance framework for St. Paul University Manila, emphasizing stage-based integration, process-based assessment, and the transition from output-oriented to cognition-centered evaluation. Ultimately, the study concludes that AI's impact is determined not by the technology itself, but by the developmental timing of its integration, ensuring students remain active seekers of knowledge rather than passive recipients of machine-generated content.

Keywords: Artificial Intelligence (AI), AI Governance Policy, Communication Program, Program Outcomes, St. Paul University Manila

INTRODUCTION

The rapid integration of artificial intelligence (AI) into educational contexts has generated a polarized discourse on its implications for learners' critical thinking. On one hand, emerging studies and practitioner perspectives suggest that AI tools can **reconfigure cognitive processes**, shifting learners' efforts from information retrieval toward higher-order functions such as verification, integration, and evaluation. On the other hand, a growing body of evidence warns that excessive reliance on AI may lead to **cognitive offloading**, reduced mental effort, and the gradual erosion of learners' capacity to independently analyze, synthesize, and generate knowledge. This tension underscores a fundamental pedagogical dilemma: whether AI serves as a **cognitive amplifier** that enhances critical thinking or a **cognitive substitute** that diminishes it.

This debate is particularly significant in the context of contemporary education systems that aim to develop learners' competencies across a broad spectrum of outcomes, including information literacy, communication, research, creativity, leadership, and ethical responsibility. These competencies—articulated in the communication program outcomes such as defining and accessing information (PO1), producing and communicating knowledge (PO2–PO3), designing communication strategies (PO4), conducting research (PO5), creating media outputs (PO6), exercising leadership (PO7), fostering entrepreneurial thinking (PO8), upholding ethical standards (PO9), and applying communication for development (PO10)—all presuppose a strong

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

foundation in critical thinking. The integration of AI into learning environments, therefore, raises urgent questions about how these outcomes can be achieved without compromising the very cognitive capacities they seek to cultivate.

At the core of this issue is not merely the presence of AI, but **how, when, and for what purposes it is used** in the learning process. If introduced prematurely or used as a replacement for foundational cognitive tasks, AI may encourage passivity and dependency. Conversely, when employed after learners have developed essential competencies, AI has the potential to extend intellectual work, enabling deeper analysis, more efficient problem-solving, and innovative knowledge production. This suggests the need for a **developmentally aligned and pedagogically grounded framework** for AI integration across basic and higher education.

Critical Thinking in the Context of Artificial Intelligence

Critical thinking has long been conceptualized as a hierarchy of higher-order cognitive processes—analysis, synthesis, and evaluation—anchored in frameworks such as Bloom's taxonomy (Bloom et al., 1956; Anderson & Krathwohl, 2001). These models frame critical thinking as both a structured competence and a measurable educational outcome. Contemporary perspectives, however, extend this view by emphasizing critical thinking as an **active and effortful engagement with information**, requiring learners to question assumptions, evaluate sources, and construct meaning independently (Facione, 1990; Paul & Elder, 2014).

The rise of artificial intelligence (AI) complicates and enriches this dual conception. By enabling rapid information generation, AI shifts cognitive demands from knowledge production to **evaluation and verification**, redefining critical thinking as the ability to interrogate machine-generated outputs (Luckin et al., 2016). This shift foregrounds AI's potential not only to transform how learners think, but also to reshape where cognitive effort is directed—thus raising questions about whether AI primarily **enhances or diminishes** critical thinking.

AI as Cognitive Augmentation and Constraint

On one hand, AI is widely regarded as a tool for **cognitive augmentation**, enabling learners to process complex information and engage in higher-order reasoning. Applications such as data analysis, pattern recognition, and simulation allow students to focus on interpretation and decision-making rather than routine tasks (Holmes et al., 2019; Zawacki-Richter et al., 2019). In this capacity, AI functions as a "thinking partner," supporting metacognitive awareness and self-regulation (Roll & Wylie, 2016), and aligning with the evolving demands of communication, research, and strategic planning.

On the other hand, this same capacity introduces the risk of **cognitive offloading**, wherein learners delegate mental processes to external systems (Risko & Gilbert, 2016). Excessive reliance on AI may reduce cognitive engagement, leading to diminished memory retention, weaker problem-solving abilities, and reduced analytical depth (Carr, 2010; Sparrow et al., 2011). In generative AI contexts, learners may become passive recipients of pre-constructed knowledge, undermining intellectual autonomy and critical thinking—particularly when AI is introduced before foundational skills are fully developed (Bender et al., 2021). This tension between augmentation and dependency underscores the need to examine **how AI use is developmentally structured**.

Developmental Integration of AI in Education

The impact of AI on critical thinking is therefore contingent on the **timing and context of its use**. Educational theories on scaffolding emphasize that learners must first develop foundational competencies before engaging with advanced tools (Vygotsky, 1978; Bruner, 1960). When introduced prematurely, AI may disrupt the formation of essential cognitive skills; when introduced appropriately, it can extend and deepen higher-order thinking.

This insight has led to **stage-based models of AI integration**, where AI use is limited in early educational stages and progressively expanded in higher education as learners develop greater cognitive and metacognitive capacity (Luckin et al., 2016). Such models align with outcome-based education frameworks, particularly in

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

communication programs, where competencies evolve from foundational literacy to strategic and professional practice. This progression highlights the importance of aligning AI use with specific learning outcomes.

Alignment with Communication Competencies or Program Outcomes (PO1–PO10 IN St. Paul University Manila)

The integration of AI in education must be contextualized within communication competencies such as PO1–PO10. Foundational outcomes—defining information needs (PO1), producing and sharing knowledge (PO2), and communicating across platforms (PO3)—require strong critical thinking to ensure accuracy, coherence, and relevance (Hobbs, 2010). At this stage, AI should function as a **supportive tool**, reinforcing rather than replacing cognitive processes.

As learners advance, competencies in communication planning (PO4), research (PO5), and media production (PO6) can benefit from AI's capacity to assist in complex tasks, provided that learners maintain critical oversight. Higher-level outcomes in leadership (PO7), entrepreneurship (PO8), and development communication (PO10) further require the ability to contextualize AI outputs within broader social, cultural, and institutional frameworks (Couldry & Hepp, 2017). Central to all these competencies is ethical practice (PO9), particularly in addressing issues of bias, transparency, and accountability in AI-generated content (Floridi et al., 2018). These considerations collectively point toward the necessity of a **balanced and ethically grounded approach** to AI integration.

Contrasting Perspectives on AI Use

The theoretical tension between AI as augmentation and as constraint is reflected in public discourse. A video by Al Jazeera English adopts an investigative-cautionary stance, highlighting neuroscientific evidence that early reliance on AI may reduce neural engagement associated with attention and memory, thereby reinforcing concerns about cognitive offloading. In contrast, a video by Sovorel advances a defensive-pedagogical perspective, arguing that AI shifts human effort from "task execution" to "task stewardship," enabling higher-order thinking when foundational skills are already established.

Together, these perspectives converge on a critical pedagogical insight: effective AI integration depends on a **"Learn First, Augment Later"** approach, where learners first develop the cognitive capacity for independent inquiry before using AI to extend their capabilities.

Toward Governance and Institutional Alignment

This pedagogical imperative necessitates institutional frameworks that regulate AI use in education. The work of Brian Bantugan (2026) addresses this need by proposing a governance model for AI integration in higher education, particularly within St. Paul University Manila. The study highlights ethical challenges such as corporate influence and techno-authoritarian risks, advocating for a **multisectoral, participatory approach** to AI governance. By emphasizing transparency, accountability, and stakeholder engagement, the framework ensures that AI remains aligned with human and institutional values.

Implications for Communication Education

Within this governance context, the program learning outcomes (PO1–PO10) of St. Paul University Manila provide a structured pathway for implementing AI in education. Foundational competencies (PO1–PO3) establish critical thinking and communication skills, while intermediate outcomes (PO4–PO6) develop strategic and technical expertise. Advanced outcomes (PO7–PO10) emphasize leadership, innovation, ethics, and social responsibility.

This progression reflects a **developmentally aligned model of AI integration**, where AI supports but does not replace cognitive engagement. Ultimately, the literature suggests that AI's educational value lies not in its capacity to generate knowledge, but in its ability to **enhance human thinking when used critically, ethically, and at the appropriate stage of learning**.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Synthesis

The current discourse on AI in education is defined by a tension between **cognitive augmentation** and **cognitive offloading**. Traditional models, such as Bloom's Taxonomy, are being re-evaluated as AI shifts the locus of human effort from knowledge production to **"task stewardship"** and information verification. Research indicates that while AI can serve as a powerful "thinking partner" for those with established expertise, it poses significant neurocognitive risks to developing minds. Specifically, "The Take" highlights that premature reliance on generative tools can lead to reduced brain connectivity in areas associated with memory and attention, effectively "atrophying" the mental muscles required for independent inquiry. Conversely, a "defensive-pedagogical" view suggests that AI is a neutral tool whose impact is determined by user agency and self-confidence.

To navigate this, Bantugan (2026) proposed a policy framework on AI use governance. A democratized governance policy—exemplified by the proposed framework for St. Paul University Manila— balances centralized risk management with participatory, multisectoral oversight. By grounding AI use in ethical accountability and Catholic social teachings, institutions can ensure that technology enhances, rather than substitutes, the "central humanness" of critical thinking.

Despite the wealth of literature on AI's theoretical impact, a significant **research gap** exists in understanding the **rhetorical and discursive divergence** between popular media and scholarly policy. While academic papers (like Bantugan, 2026) focus on structural governance, institutional accountability, and multisectoral policy, popular public discourse, especially on YouTube, tends to frame AI through highly emotional or deterministic lenses—either as a "brain-breaking" threat to human nature or a liberating tool for efficiency. There is a lack of empirical study on how these **contradictory public narratives** influence student and faculty perceptions of "legitimate" AI use.

Specifically, existing research has not sufficiently explored how a **comparative analysis of popular media transcripts** can be synthesized with **formal institutional outcomes** to create a "middle-ground" literacy. This study addresses this gap by triangulating neurocognitive warnings from popular media with academic governance frameworks, thereby proposing a localized, stage-based integration model that translates high-level theory into practical, everyday classroom policy.

THEORETICAL FRAMEWORK

This study is anchored on an integrative theoretical framework that views artificial intelligence not as an autonomous determinant of learning, but as a mediating tool whose impact is defined by cognitive development, technological mediation, and learner agency. By synthesizing **Bloom's Taxonomy** with **Vygotsky's Sociocultural Theory**, the framework posits that critical thinking emerges through a progression toward higher-order cognitive tasks within a learner's **Zone of Proximal Development (ZPD)**. In this context, AI serves as a "double-edged sword": it can act as a constructive scaffold that enhances analysis and synthesis when used to support the learner, or it can become a detrimental substitute that triggers **Cognitive Offloading**, leading to reduced memory retention and diminished problem-solving skills.

The framework further incorporates **Media and Technological Mediation Theory** alongside **Constructivist Learning** principles to highlight how AI transforms learners from active information seekers into potential receivers of pre-processed knowledge. This shift risks "short-circuiting" the active construction of knowledge necessary for independent inquiry. Ultimately, the model suggests that the transition from enhanced to diminished critical thinking outcomes depends on the **Mode of AI Use** (scaffold vs. substitute) and the **Level of Cognitive Engagement** (active vs. passive). Consequently, the study proposes that while AI can bolster higher-order thinking in advanced educational stages, its unregulated use in foundational stages may weaken the very cognitive processes it aims to assist.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Statement of the Problem

In light of these considerations, this study sought to examine the role of AI in shaping critical thinking within educational contexts, guided by the following research questions:

1. How do the two videos (for and against AI use) define critical thinking?

2. How does AI affect critical thinking based on the assertions of the two videos?

3. What are the agendas supporting the arguments in each video?

4. Given the arguments for and against the use of AI and the learning outcomes of the Communication Program of St. Paul University Manila, how should it be used in basic and higher education?

5. What revisions should be done on the preliminary framework on generative AI governance initially proposed to St. Paul University Manila?

By addressing these questions, the study aimed to contribute to the ongoing discourse on AI in education by moving beyond binary positions of advocacy or resistance, and instead proposing a **balanced, evidence-informed approach** to AI integration—one that safeguards critical thinking while harnessing the transformative potential of emerging technologies.

METHODOLOGY

Research Design

This study employed a qualitative, interpretive research design that integrates comparative discourse analysis and synthetic document analysis to examine how artificial intelligence (AI) influences critical thinking in educational contexts. The design is grounded in the assumption that knowledge about AI's cognitive impact is socially constructed through discourse, particularly as expressed in publicly accessible media and scholarly publications (Couldry & Hepp, 2017). By analyzing both popular and academic sources, the study recognizes that meanings about AI are shaped not only by empirical evidence but also by competing narratives and interpretations.

The study utilized a **purposive sampling strategy** to select data sources that represent the polarized spectrum of the AI discourse.

- **Video 1 (Sovorel):** Selected as a "defensive-pedagogical" benchmark. It was chosen for its focus on AI literacy and the shift from "task execution" to "task stewardship".

- **Video 2 (Al Jazeera):** Selected as a "cautionary-investigative" benchmark. It provides neuroscientific evidence regarding "cognitive offloading" and neural connectivity risks.

- **Scholarly Context:** The Bantugan (2026) paper was selected to provide a localized, multisectoral governance perspective grounded in Catholic social teachings.

The primary data for comparative analysis consisted of two purposively selected YouTube video transcripts. These videos were chosen based on their explicit focus on the impact of AI on critical thinking, their presentation of opposing viewpoints, and their accessibility to a broad public audience. One video advanced the argument that AI reconfigures and enhances critical thinking by shifting cognitive effort toward higher-order processes, while the other emphasized the risks of diminished cognitive engagement and dependency associated with AI use. This purposive sampling strategy aligns with qualitative research practices that prioritize information-rich cases relevant to the research problem (Patton, 2002).

The second data source was a scholarly article by Brian Bantugan (2026), which provided theoretical and empirical grounding for understanding AI's implications in education. The article was selected due to its

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

relevance to AI governance, critical thinking development, and educational policy. Its inclusion enabled the study to contextualize and synthesize the contrasting claims found in public discourse, thereby supporting the development of a balanced analytical framework.

The third data sources, the national learning outcomes for basic education and the program outcomes for the Communication Programs in St. Paul University Manila, as found in online sources and in all syllabi of the said program, respectively, served as references to arrive at the synthesis of the data gained from the first two data sources in the form of a contextualized AI governance policy, particularly in basic education, in general, and the university's communication program at the tertiary level, in particular. This facilitated the contextualization of the AI use policy framework for the communication program given the proper scaffolding in basic education.

Specifically, the study utilized three complementary investigations: a comparative analysis of two contradicting YouTube video transcripts, a synthetic analysis of a published academic paper by Brian Bantugan (2026), and review of the current learning outcomes of the basic education and the communication programs in St. Paul University Manila. This triadic approach enabled triangulation across different forms of discourse, thereby enhancing the analytical depth and credibility of the findings (Denzin, 1978; Patton, 2002), and enabling contextualization of AI governance policy in the university. The integration of these data sources allowed the study to move beyond binary positions and develop a more nuanced, evidence-informed understanding of AI's role in shaping critical thinking.

Data analysis proceeded in three phases. First, the YouTube transcripts were examined through comparative discourse analysis, focusing on how each source constructs the relationship between AI and critical thinking. The process began with transcription and familiarization, followed by open coding to identify recurring themes such as definitions of critical thinking, perceived benefits and risks of AI, the role of the learner, and the nature of cognitive engagement. These codes were then organized through axial coding into broader thematic categories, allowing for the identification of patterns, points of convergence and divergence, and underlying assumptions within each discourse (Strauss & Corbin, 1998). Comparative interpretation of the two transcripts revealed contrasting paradigms of AI as **cognitive augmentation** and AI as **cognitive substitution**, reflecting broader debates in the literature (Risko & Gilbert, 2016).

The second phase involved synthetic document analysis of Bantugan's (2026) work. Relevant sections of the text were identified and coded in relation to the themes derived from the comparative analysis, particularly those concerning AI governance, cognitive implications, and educational applications. Through a process of synthesis and integration, the scholarly insights were used to validate, refine, or challenge claims from the YouTube transcripts. This iterative comparison enabled the study to bridge opposing perspectives and construct a theoretically grounded, stage-based framework for AI use in education, aligned with communication-related learning outcomes.

The analysis was guided by an integrative theoretical framework drawing on cognitive hierarchy theory (Bloom et al., 1956; Anderson & Krathwohl, 2001), sociocultural theory (Vygotsky, 1978), cognitive offloading theory (Risko & Gilbert, 2016), and constructivist learning theory (Bruner, 1960). These theoretical lenses informed both the coding process and the interpretation of findings, ensuring conceptual coherence and alignment with established scholarship on learning and cognition.

The study followed a three-phase **Comparative Discourse Analysis**:

1. **Phase 1: Open and Axial Coding:** Transcripts were coded for recurring themes such as "cognitive atrophy" vs. "cognitive augmentation".

2. **Phase 2: Synthetic Document Analysis:** Scholarly insights from Bantugan (2026) were used to triangulate media claims against institutional policy needs.

3. **Phase 3: Theoretical Synthesis:** Findings were interpreted through the lens of the **Zone of Proximal Development (ZPD)** and **Cognitive Offloading Theory** to determine the "timing" of AI integration.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

To ensure trustworthiness and rigor, the study employed multiple strategies consistent with qualitative research standards. Triangulation was achieved through the use of diverse data sources, combining popular media and academic literature (Denzin, 1978). Comparative validation was conducted by cross-examining opposing claims across the two transcripts, while theoretical anchoring ensured that interpretations were grounded in established frameworks. Additionally, thick description was used through the inclusion of direct quotations to preserve the integrity and context of the original arguments (Geertz, 1973).

Despite these strengths, the study has limitations. The use of only two YouTube videos may not fully capture the diversity of public discourse on AI, and the purposive selection of sources may introduce bias. Furthermore, the reliance on secondary data limits the ability to observe actual learner behavior. However, these limitations are mitigated by the inclusion of a scholarly source and a systematic analytical approach, which together enhance the robustness of the findings.

Overall, this methodology enabled a multi-layered analysis of AI's impact on critical thinking by comparing contradictory public narratives, synthesizing scholarly insights, and grounding interpretations in established theoretical frameworks. Through this approach, the study moves beyond polarized views and contributes to a more nuanced, evidence-based understanding of how AI can be effectively and responsibly integrated into educational contexts.

RESULTS

How do the two videos (for and against AI use) define critical thinking?

Critical Thinking as a Hierarchical Cognitive Process. In the first dataset, critical thinking is framed through an educational-cognitive lens, defined as **"a hierarchical taxonomy that characterizes student learning objectives into six types: knowledge… comprehension… application… analysis… synthesis… and evaluation."** This perspective views cognition as a structured progression of measurable skills where **"critical thinking in knowledge work involves a range of cognitive activities such as analysis, synthesis and evaluation."** Rather than viewing technology as a threat, this dataset suggests that critical thinking is an adaptive competency. With the integration of AI, the nature of intellectual labor evolves because **"the focus shifts from information gathering to information verification,"** effectively reallocating human effort toward higher-level oversight.

Critical Thinking as Active, Effortful, and Self-Driven Cognition. Conversely, the second dataset adopts a neurocognitive-behavioral paradigm, emphasizing the experiential and effort-dependent nature of thought. Here, critical thinking is described as a process where **"you're presented with a bunch of information that you need to sift through on your own, which still requires that you're doing a lot of your own thinking."** There is a pronounced concern that AI mediates this process by turning the learner **"from an active seeker of information into a passive consumer of information."** Under this view, critical thinking is a fragile human attribute—**"something that feels… pretty central to humanness"**—that risks being lost to cognitive offloading. The dataset warns that **"outsourcing thoughts to AI leaves people's minds 'atrophied and unprepared'"** because **"our brains learn to devote fewer resources towards those skills."**

Synthesis of Divergent Paradigms. The core tension between these definitions lies in whether critical thinking is a resilient structure or vulnerable behavior. While the first dataset argues that **"effort shifts from task execution to task stewardship,"** suggesting a professional evolution, the second dataset contends that this reduction in mental friction leads to cognitive decline. Ultimately, these perspectives suggest that critical thinking is both a system of competencies and an active engagement with information. A comprehensive understanding requires integrating both views: while AI may provide the tools for higher-level synthesis, the structural capability of the mind only remains intact when the learner resists passivity and maintains sustained mental effort.

The following table maps the contrasting perspectives of the two datasets against the revised hierarchical levels of Bloom's Taxonomy, illustrating how AI can either support cognitive growth or lead to cognitive atrophy.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Table 1 Maps of the contrasting perspectives of the two datasets against the revised hierarchical levels of Bloom's Taxonomy

Bloom's Level	Educational-Cognitive Paradigm (Video 1)	Neurocognitive-Behavioral Paradigm (Video 2)	AI's Role in Critical Thinking
Creating	Generating original frameworks or models with AI-assisted brainstorming.	Risk of losing "central humanness" and original creative spark.	**Scaffold:** Co-creation. **Substitute:** Automated generation.
Evaluating	**"Information verification"** as the new primary cognitive task.	**"Passive consumer"** who accepts AI output without judgment.	**Scaffold:** Critique of AI. **Substitute:** Uncritical trust.
Analyzing	**"Task stewardship"** and oversight of complex data relationships.	Outsourcing analysis leaves the mind **"atrophied and unprepared."**	**Scaffold:** Pattern recognition. **Substitute:** Data outsourcing.
Applying	Using AI to solve complex, real-world problems in new contexts.	**"Devoting fewer resources"** to the mechanics of problem-solving.	**Scaffold:** Scenario simulation. **Substitute:** Process bypassing.
Understanding	Explaining AI-generated summaries in the learner's own words.	Shift from an **"active seeker"** to a recipient of pre-processed ideas.	**Scaffold:** Concept mapping. **Substitute:** Passive reading.
Remembering	Retrieval of facts is deprioritized to favor higher-order verification.	Sustained **"cognitive offloading"** weakens foundational memory.	**Scaffold:** Spaced repetition. **Substitute:** External storage.

The data reveal that **Video 1** views the upper levels of Bloom's Taxonomy (Analyze, Evaluate, Create) as the new "home" for human effort, while AI handles the lower-order tasks. In contrast, **Video 2** warns that bypassing the lower levels (Remembering, Understanding) through AI substitution prevents the brain from developing the "mental muscle" required to perform at the higher levels.

How does AI affect critical thinking based on the assertions of the two videos?

AI as a Tool that Shifts and Redistributes Critical Thinking. The first dataset argues that AI does not inherently diminish critical thinking; rather, it fundamentally reconfigures where and how human intellect is applied. According to this view, there is a distinct transformation in cognitive focus where **"the focus shifts from information gathering to information verification."** Instead of eliminating thought, AI reorients it toward high-level oversight, suggesting that **"effort shifts from task execution to task stewardship."** The impact of the technology is largely determined by the user's existing skills and intent, as **"knowledge workers who already engage in critical thinking… are likely to continue doing so even when using generative AI tools."** From this perspective, the risk—often framed as the fear that **"if you use AI you're going to lose all of your critical thinking skills"**—is treated as an avoidable outcome contingent on proper training and intentional engagement rather than an inevitable result of the technology itself.

AI as a Tool that Reduces Cognitive Engagement and Weakens Thinking. In contrast, the second dataset presents a cautionary narrative, suggesting that AI can actively erode critical thinking by fostering passivity. The primary concern is that AI **"kind of turns you from an active seeker of information into a passive consumer of information,"** leading to measurable declines in cognitive ownership. Evidence of this "cognitive offloading" is seen in cases where users **"truly could not remember what they were writing about,"** suggesting that **"outsourcing thoughts to AI leaves people's minds 'atrophied and unprepared.'"** This dataset highlights a significant dependency effect, noting that **"if we hand AI the wheel… our brains learn to devote fewer resources towards those skills."** Neurological observations support this, indicating that individuals who rely on their own cognitive processes show **"higher connectivity… in brain waves associated with attention… memory… and creativity"** compared to those who delegate thinking to the machine.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Synthesis: The Dual Nature of AI Mediation. When viewed together, these transcripts reveal that AI's effect on critical thinking is not uniform but exists on a spectrum between **"active stewardship"** and **"passive consumption."** The tension lies in the nature of cognitive effort: while one paradigm sees effort as being redirected toward verification and evaluation, the other sees it as being dangerously reduced. Ultimately, AI appears to amplify existing cognitive habits. For the skilled and reflective user, the technology acts as a partner in **"verification"**; for the uncritical or dependent user, it becomes a **"cognitive substitute"** that triggers disengagement. Therefore, the decisive factor in whether critical thinking is enhanced or diminished is not the tool itself, but the degree to which the learner maintains active agency over the thinking process.

What are the agendas supporting the arguments in each video?

Agenda of Video 1: Defensive–Pedagogical (Pro-AI Literacy and Skill Adaptation). The first dataset is driven by an agenda that seeks to dismantle alarmist narratives and reposition AI as a manageable educational tool. This corrective approach begins with an **explicit rejection of alarmism**, flatly stating, **"no AI will not take away your critical thinking skills."** By framing current fears as exaggerated, the speaker emphasizes a rigorous **interpretation of research**, warning that **"we can't use that just for confirmation bias… we need to make sure that we're fully reading the article."** The primary goal is to advocate for **AI literacy and skill development**, asserting that **"we have to develop critical thinking within our students so that they have this capability."** Ultimately, this perspective reflects a **human-centered responsibility agenda**, shifting accountability away from the machine and toward the user by concluding that **"it isn't AI that's causing that it's you that's making choices."**

Agenda of Video 2: Cautionary–Investigative (Critical Examination of AI's Cognitive Impact). In contrast, the second dataset operates under a cautionary agenda, focused on interrogating the potential neurological and cognitive costs of AI integration. It frames the technology as a **potential cognitive threat**, opening with the provocative question: **"is it breaking our brains?"** To support this concern, the narrative relies on **empirical and scientific evidence**, such as EEG studies where researchers **"were able to track what their brains were doing… and how much over time."** By highlighting negative outcomes—such as findings that users **"truly could not remember what they were writing about"**—the dataset builds a case that **"outsourcing thoughts to AI leaves people's minds 'atrophied and unprepared.'"** This agenda is rooted in **problematizing passivity and dependency**, warning that AI **"turns you from an active seeker… into a passive consumer."** Rather than offering a solution, it concludes with **open-ended inquiry**, noting that the **"long-term effects… remains to be seen."**

Synthesis: Adaptation vs. Preservation. The divergence between these two datasets reveals a fundamental tension in the AI discourse: a pedagogical agenda focused on **adaptation** versus a cautionary agenda focused on **preservation**. While Data Set 1 is reassuring and solution-oriented, aiming to empower educators to integrate AI responsibly, Data Set 2 is exploratory and problem-oriented, seeking to protect human agency from technological erosion. These perspectives are not necessarily contradictory but are rather normative reflections of different priorities. One seeks to normalize the tool to ensure students are not left behind, while the other seeks to problematize the tool to ensure cognitive integrity is not lost.

Given the arguments for and against the use of AI and the learning outcomes of the Communication Program of St. Paul University Manila, how should it be used in basic and higher education?

Foundational Principle: "Learn First, Augment Later". A synthesis of the research suggests a core rule for pedagogical design: if AI is used before a learner develops foundational core skills, it triggers **"cognitive offloading"** and weakens independent thought; however, if integrated after competence is established, it enhances higher-order thinking such as verification and evaluation. This necessitates a progressive model where Basic Education focuses on **Skill Formation** with restricted AI use, while Higher Education transitions toward **Skill Augmentation**, treating AI as a guided cognitive partner.

AI Use in Basic Education (Primary–Secondary). At this level, learners are developing the cognitive foundations aligned with **Program Outcomes (PO) 1–3**. AI should be used as a "cognitive mirror" rather than a substitute. For **PO1 (Information Needs)**, students must first learn to search and organize information

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

manually, using AI only as a secondary verification tool to identify gaps. Regarding **PO2 and PO3 (Communication)**, all production must be student-generated first to avoid becoming a **"passive consumer of information."** AI may support grammar refinement or feedback simulation, but the rule remains: **no AI-generated first drafts**. In the context of **PO6 (Creativity)** and **PO9 (Ethics)**, AI should serve as an initial idea prompt while students receive explicit instruction on AI bias and misinformation, preventing uncritical reliance before they reach higher levels of education.

AI Use in Higher Education. In Higher Education, where learners tackle **PO4–PO10**, AI functions as a "thinking amplifier." This aligns with the shift described in the data where **"effort shifts from task execution to task stewardship."** For **PO4 and PO5 (Communication Planning and Research)**, AI assists in audience analysis and literature summarization, but students must validate all sources and conclusions. In **PO6 (Media Production)** and **PO7 (Leadership)**, AI acts as a co-creator or decision-support tool, but the learner must demonstrate original intentionality and exercise **"judgment and accountability."** Furthermore, for **PO8 (Entrepreneurship)** and **PO10 (Development Communication)**, AI helps explore market gaps and adapt messages, yet it must never replace ground-level community engagement or the critical assessment of feasibility.

Integrated Instructional Framework and Policy Implications. This stage-based model requires a fundamental shift in policy and curriculum design. Curriculum must move toward an **"AI-last"** approach in early stages and an **"AI-with-critique"** approach in advanced stages. Assessment reform is equally vital; evaluation must shift from **output-based** (the final product) to **process-based** (how the student used AI). Educators should require evidence of independent thinking, such as logs showing what AI suggestions were accepted or rejected.

Synthesis: Active Agency vs. Machine Knowledge. Ultimately, the research datasets converge on the insight that AI is a "double-edged tool" whose impact is determined by timing and application. Across all learning outcomes, AI must be a scaffold in basic education and a partner in higher education, but never a replacement for cognitive responsibility. The goal of this framework is to ensure that as students progress toward professional life, they remain **"active seekers"** and thinkers, rather than becoming passive recipients of machine-generated knowledge.

What revisions should be done on the preliminary framework on generative AI governance initially proposed to St. Paul University Manila?

To improve clarity and practical understanding, the following structured model illustrates the framework proposed for St. Paul University Manila.

Table 2 The "Learn First, Augment Later" Hierarchy

Stage	**Primary Objective**	**AI Role**	**Governance Rule**
Foundational	Skill Acquisition	**Cognitive Scaffold**	AI-generated first drafts are prohibited.
Advanced	Task Stewardship	**Cognitive Partner**	Mandatory critical evaluation and justification of AI outputs.
Professional	Efficiency & Ethics	**Augmentation Tool**	Full integration with final-mile human accountability.

Modified Final Policy Framework (Enhanced with Educational Ai Use Principles)

I. Foundational Policy Enhancement: Developmental AI Use Principle

New Clause (to be inserted under Preamble or Ethical Foundations):

AI shall be integrated in education according to a **developmental progression**, ensuring that it **enhances rather than replaces critical thinking**. Its use must align with the learner's stage of cognitive development, preserving foundational skills while enabling advanced competencies.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

This responds directly to the finding that:

- AI can **"shift effort from task execution to task stewardship"**
- but also risks **passive dependency and reduced cognitive engagement** (from transcript synthesis)

II. Teaching and Learning Governance (REVISED)

The original policy states: "AI may not replace formative human mentorship and evaluation"

Modified Version:

A. Stage-Based AI Use Framework. AI use in teaching shall follow a **two-tier developmental model**:

1. Basic Education (Foundational Stage)

- AI use shall be **limited, guided, and non-substitutive**
- Prohibited uses are: (1) AI-generated first drafts; and (2) Full-task completion by AI
- Permissible uses include: (1) Verification of student-generated work; (2) Reflective feedback support; and (3) Guided exploration

Policy Addition: AI shall function as a **cognitive scaffold**, ensuring that learners remain active producers of knowledge rather than passive recipients.

2. Higher Education (Advanced Stage)

- AI may be used for: (1) Analysis, synthesis, and planning; (2) Research support and data interpretation; (3) Media and communication production

With required condition: All AI outputs must be **critically evaluated, modified, and justified by the learner**.

B. AI Use Transparency in Learning. Expand existing disclosure requirement: "Students must disclose AI assistance where required"

Modified: Students must include: (1) **Type of AI use (incidental, moderate, substantive); (2) Cognitive role of AI (support vs. generation); (3) What was accepted, rejected, and revised**

III. Curriculum Integration (NEW SECTION)

AI-Integrated Learning Outcomes Alignment. To align with PO1–PO10, the policy must explicitly define AI's role:

A. PO1–PO3 (Foundational Communication Skills)

- AI restricted to: (1) Verification and (2) Editing support
- Students must demonstrate: (1) Independent information retrieval; and (2) Original communication outputs

B. PO4–PO7 (Strategic and Analytical Competencies)

- AI allowed for: (1) Planning and modeling; and (2) Scenario generation
- Required: Human validation and contextualization

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

C. PO8–PO10 (Entrepreneurship, Ethics, Development)

- AI used for: (1) Opportunity scanning; and (2) Development communication modeling

With constraint: AI outputs must be evaluated against **ethical, cultural, and developmental contexts**

IV. Assessment Reform (NEW POLICY ADDITION)

To address cognitive offloading risks:

A. Process-Based Assessment Requirement. All AI-assisted outputs must include: (1) Documentation of thinking process; (2) Evidence of independent reasoning; and (3) Reflection on AI contribution

B. AI-Resilient Assessment Design. Institutions shall implement: (1) In-class or controlled assessments; (2) Oral defenses and presentations; (3) Iterative drafts showing thinking evolution

V. Risk Classification (ENHANCED)

The framework already includes tiered AI use: "incidental, moderate, and substantive"

Add Cognitive Risk Dimension:

Table 3 Cognitive Risk Dimensions

AI Use Type	Cognitive Risk	Policy Action
Incidental	Low	Freely permitted
Moderate	Medium	Requires disclosure + reflection
Substantive	High	Requires justification + evaluation evidence

VI. Institutional Accountability (EXPANDED)

Original: "Institutions shall promote critical reflection on AI and human values"

Enhanced Version: Institutions shall:

- Monitor **cognitive impact of AI use on learners**
- Ensure AI does not: (1) Replace foundational learning; and (2) Reduce engagement
- Track **long-term effects on critical thinking development**

VII. AI Literacy and Training (REVISED)

Original: "Regular training in AI ethics and governance is mandatory"

Expanded: Training must include: (1) **Cognitive risks of AI (e.g., dependency, passivity); (2) Strategies for critical engagement with AI outputs; (3) Prompting as a thinking tool, not a shortcut.**

VIII. Governance Model Enhancement

The framework already emphasizes: "multisectoral… participatory… ethical oversight"

Add Educational Stakeholder Role: Include (1) Curriculum specialists; (2) Cognitive psychologists / learning experts; and (3) Students as co-designers of AI use policies

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

IX. Concluding Policy Statement (REVISED)

Original: "technology must remain at the service of the human person"

Modified: Artificial intelligence shall remain at the service of the human person by **strengthening—not substituting—human cognition**. Its use in education must ensure that learners develop as **independent, critical, and ethical thinkers**, capable of engaging with AI as a tool for augmentation rather than dependency.

X. Key Policy Shift (Synthesis)

From: AI as **tool to be regulated**

To: AI as **tool to be developmentally integrated and cognitively governed**

Insight

This modification aligns the framework with the central finding: (1) AI is not inherently harmful or beneficial; and (2) Its impact depends on **when, how, and at what stage it is used**

By embedding a **stage-based, cognition-centered approach**, the policy (1) Preserves **critical thinking in basic education; (2)** Enhances **advanced competencies in higher education; (3)** Ensures alignment with **PO1–PO10; and (3)** Strengthens resistance to **passivity, dependency, and techno-authoritarian risks**

Implementation, Faculty Training, and Challenges

To enhance applicability, the following section expands on the practical realities of institutionalizing the policy.

Faculty Training and Awareness. The policy mandates a shift from output-oriented to cognition-centered evaluation.

Training Focus. Educators must be trained in "Process-Based Assessment," learning how to evaluate "Version History" and "AI Audit Logs" rather than just the final essay.

Literacy. Training must include identifying "cognitive risks" like dependency and passivity, moving beyond simple tool usage.

Implementation Challenges

The Literacy Gap. Students and faculty may have divergent views on what constitutes "legitimate" AI use based on emotional or deterministic media narratives.

Techno-Authoritarian Risks. The framework must actively resist corporate influence and data sovereignty issues by prioritizing open-source AI (e.g., LLAMA) for sensitive research.

Evaluation Fatigue. Implementing "Track Changes" or "Process Portfolios" increases the grading burden on faculty, requiring administrative support and possibly AI-assisted grading of the process itself.

Policy Evaluation Mechanisms. To ensure the policy remains responsive, the AI Governance Committee will:

1. **Monitor Cognitive Impact**. Track long-term effects on students' independent inquiry and memory retention.

2. **Participatory Review**. Engage students as "co-designers" of the policy to ensure it remains relevant to their evolving workflow.

3. **Iterative Updates**. Conduct regular audits of the "Risk Tier" classifications as AI capabilities evolve.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

DISCUSSION

AI as a Tool that Shifts and Redistributes Critical Thinking

The first video argues that AI does not inherently diminish critical thinking; rather, it fundamentally reconfigures where and how human intellect is applied. According to this view, there is a distinct transformation in cognitive focus where **"the focus shifts from information gathering to information verification."** Instead of eliminating thought, AI reorients it toward high-level oversight, suggesting that **"effort shifts from task execution to task stewardship."** The impact of the technology is largely determined by the user's existing skills and intent, as **"knowledge workers who already engage in critical thinking… are likely to continue doing so even when using generative AI tools."** From this perspective, the risk—often framed as the fear that **"if you use AI you're going to lose all of your critical thinking skills"**—is treated as an avoidable outcome contingent on proper training and intentional engagement rather than an inevitable result of the technology itself.

AI as a Tool that Reduces Cognitive Engagement and Weakens Thinking

In contrast, the second video presents a cautionary narrative, suggesting that AI can actively erode critical thinking by fostering passivity. The primary concern is that AI **"kind of turns you from an active seeker of information into a passive consumer of information,"** leading to measurable declines in cognitive ownership. Evidence of this "cognitive offloading" is seen in cases where users **"truly could not remember what they were writing about,"** suggesting that **"outsourcing thoughts to AI leaves people's minds 'atrophied and unprepared.'"** This dataset highlights a significant dependency effect, noting that **"if we hand AI the wheel… our brains learn to devote fewer resources towards those skills."** Neurological observations support this, indicating that individuals who rely on their own cognitive processes show **"higher connectivity… in brain waves associated with attention… memory… and creativity"** compared to those who delegate thinking to the machine.

Synthesis: The Dual Nature of AI Mediation

When viewed together, these videos reveal that AI's effect on critical thinking is not uniform but exists on a spectrum between **"active stewardship"** and **"passive consumption."** The tension lies in the nature of cognitive effort: while one paradigm sees effort as being redirected toward verification and evaluation, the other sees it as being dangerously reduced. Ultimately, AI appears to amplify existing cognitive habits. For the skilled and reflective user, the technology acts as a partner in **"verification"**; for the uncritical or dependent user, it becomes a **"cognitive substitute"** that triggers disengagement. Therefore, the decisive factor in whether critical thinking is enhanced or diminished is not the tool itself, but the degree to which the learner maintains active agency over the thinking process.

The following table presents a **Stage-Based Integration Model**, balancing the **Pedagogical Agenda** (which encourages adaptation and literacy) with the **Cautionary Agenda** (which protects against cognitive atrophy).

This model follows a "Scaffold-to-Stewardship" progression, ensuring students build the necessary "mental muscle" before delegating tasks to AI.

Table 4 Stage-Based AI Integration Model

Educational Level	Primary Objective	AI Integration Strategy	Cognitive Safeguard (The "Cautionary" Check)
Foundational (Freshman/Early Major)	**Skill Acquisition** & Brain Connectivity	**AI-Prohibited Zones:** Focus on handwritten drafts, oral exams, and "unplugged" synthesis.	Ensures students do not become **"atrophied and unprepared"** by building foundational memory.
Intermediate (Sophomore/Junior)	**Collaborative Literacy**	**AI as a "Tutor":** Using AI to explain complex concepts or generate counter-arguments for a student-written thesis.	Prevents becoming a **"passive consumer"** by requiring students to lead the inquiry.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

Advanced (Senior/Capstone)	**Task Stewardship**	**AI as a "Processor":** Using AI to handle large data sets or initial drafts, focusing on **"information verification."**	Shifts focus to high-level **"stewardship"**; students must defend every AI-assisted choice.
Professional (Post-Grad/Workforce)	**Efficiency & Ethics**	**Full Integration:** AI handles "task execution" while the human provides the ethical and logical oversight.	Maintains **"central humanness"** through final-mile accountability and creative direction.

Implementation Insights

- **For the Freshman Level:** The goal is to prevent **"handing AI the wheel"** too early. By restricting AI, we ensure the **"higher connectivity"** in brain waves mentioned in the research is established.
- **For the Senior Level:** The focus moves to **Video 1's** vision. Since the student already has a developed cognitive hierarchy, they can safely move from **"execution"** to **"stewardship"** without losing their ability to think critically.

Process-based Evaluation

A **Process-Based Rubric** is designed to shift the focus from the final written product to the **"task stewardship"** and **"information verification"** processes identified in the research. It rewards students for actively challenging AI rather than passively accepting its output.

Table 5 Rubric for AI-Assisted Critical Thinking

Criteria	Level 1: Passive Consumer (Emerging)	Level 2: Guided User (Developing)	Level 3: Active Steward (Proficient)
Initial Inquiry (PO 1)	Relies on AI for the primary search without prior manual research.	Uses AI to find sources but performs limited independent verification.	Conducts independent research first; uses AI to **"identify gaps"** or check for completeness.
Verification & Accuracy (PO 5)	Accepts AI-generated facts, citations, and logic without checking for hallucinations.	Checks major facts but misses subtle logical inconsistencies or "fake" citations.	Systematically audits all AI claims; provides evidence of cross-referencing with scholarly databases.
Original Intent & Voice (PO 2-3)	Final output reads like a standard AI response; lacks personal perspective or local context.	Integrates some personal voice but the structure remains heavily dictated by the AI draft.	AI is used for **"alternative phrasing"** only; final output demonstrates clear, original intent and **"central humanness."**
Metacognitive Reflection (PO 9)	Cannot explain why specific AI suggestions were included; lacks an "AI log."	Provides a basic log of prompts but offers little reflection on why certain outputs were rejected.	Provides a detailed **"Process Portfolio"** justifying every edit, rejection, and modification of AI suggestions.
Ethical Disclosure (PO 9)	No mention of AI use or vague acknowledgment at the end.	Mentions AI tools used but does not specify the extent of the machine's contribution.	Full transparency; clearly delineates which parts were AI-assisted and how bias was mitigated.

The table presents a developmental continuum of AI use in learning, organized across three levels: Passive Consumer (Emerging), Guided User (Developing), and Active Steward (Proficient). These levels illustrate how learners progressively move from dependency on AI toward critical, reflective, and ethical engagement, aligned with key program outcomes (PO1, PO2–3, PO5, and PO9).

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

At the most basic level, the Passive Consumer (Emerging) is characterized by a high degree of reliance on AI with minimal cognitive engagement. In terms of initial inquiry (PO1), learners depend on AI for primary searches without conducting prior manual research, indicating weak information-seeking skills. This pattern extends to verification and accuracy (PO5), where AI-generated facts and citations are accepted without scrutiny, exposing learners to risks such as misinformation and hallucinated content. Outputs at this level (PO2–3) tend to mirror generic AI responses, lacking originality, personal voice, or contextual grounding. Furthermore, learners demonstrate limited metacognitive awareness (PO9), as they are unable to explain their use of AI or justify the inclusion of specific outputs. Ethical considerations are also minimal, with little to no disclosure of AI use, reflecting a lack of accountability and transparency.

At the intermediate level, the Guided User (Developing) demonstrates partial engagement with AI while beginning to exercise independent judgment. For initial inquiry (PO1), learners use AI to locate sources but engage in only limited verification, suggesting emerging but inconsistent research skills. In terms of verification (PO5), they check obvious facts but may overlook deeper inconsistencies or fabricated references. Their outputs (PO2–3) show some integration of personal voice; however, the overall structure remains heavily influenced by AI-generated drafts. Metacognitively (PO9), learners begin documenting their AI use through basic prompt logs, although reflection on decision-making remains superficial. Ethical disclosure improves at this stage, as learners acknowledge the use of AI tools, but they often fail to specify the extent or nature of AI contributions. This level reflects a transitional phase where learners are beginning to shift from passive reliance to more intentional use, yet still lack full critical control.

At the highest level, the Active Steward (Proficient) represents a mature and critical engagement with AI, where the learner maintains full cognitive agency. In terms of initial inquiry (PO1), learners prioritize independent research and use AI strategically to identify gaps or enhance completeness. Verification practices (PO5) are rigorous and systematic, involving cross-referencing AI-generated information with credible scholarly sources. Outputs (PO2–3) at this level are distinctly original, with AI serving only as a tool for refinement, such as suggesting alternative phrasing, while the overall structure and intent remain clearly human-driven. Metacognitive reflection (PO9) is highly developed, as learners provide detailed process portfolios that justify every interaction with AI, including edits, rejections, and modifications. Ethical practice is also fully realized, with transparent disclosure of AI use and explicit efforts to address potential biases. At this level, AI is no longer a substitute for thinking but a strategic partner, supporting and extending the learner's critical and creative capacities.

Overall, the framework illustrates progression from passive consumption to active stewardship, emphasizing that effective AI use in education is not defined by access to technology but by the learner's ability to engage with it critically, reflectively, and ethically.

Implementation Strategy: The "Track Changes" Method

To make this rubric effective, teachers should ask students to submit their work using a "Track Changes" or "Version History" format:

1. **Version 1:** The student's initial, unassisted brain-map or outline.

2. **Version 2:** The raw AI-generated response based on their prompts.

3. **Version 3:** The final draft, showing the student's heavy edits, strikethroughs, and factual corrections.

From Cognitive Substitution to Intellectual Stewardship

This study concludes that AI is neither a neutral tool nor a deterministic force; rather, it is a **mediating variable** whose impact is defined by the timing of its integration and the agency of the learner. The risk of **"cognitive offloading"** and the potential for the human mind to become **"atrophied and unprepared"** are real consequences of premature or unregulated AI use. However, these risks are mitigated when AI is introduced through a **Stage-Based Integration Model** that prioritizes the development of foundational "mental muscles" before allowing for technological augmentation. By establishing "AI-Free Zones" in basic education and

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

transitioning to "Task Stewardship" in higher education, the pedagogical framework ensures that learners build the **"higher connectivity"** in brain waves associated with attention and memory necessary for true critical thinking.

The implementation of a **Process-Based Rubric** further serves as a practical safeguard against passivity. By shifting the focus from the final artifact to the **"information verification"** and critical editing of AI-generated content, educators can transform AI from a **"cognitive substitute"** into a **"thinking amplifier."** This shift ensures that even when AI performs lower-order tasks, the learner remains the active constructor of knowledge, preserving the **"central humanness"** of the intellectual process. Ultimately, the goal of modern education is not to shield students from AI, but to cultivate a generation of **"active seekers"** who possess the resilience to navigate a machine-mediated world without sacrificing their independent capacity for inquiry, reflection, and judgment.

REFERENCES

1. Al Jazeera English. (2025, July 4). Is ChatGPT hurting our critical thinking skills? | The Take. https://www.youtube.com/watch?v=qz_ZYa3oRSE&t=1s
2. Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. Longman.
3. Bantugan, B. (2026, February 3). A Multisectoral and Democratized AI Governance Policy for St. Paul University Manila Countering Global Techno-Authoritarianism and Abuse. International Journal of Research and Scientific Innovation, 13(1), 1040-1079
4. Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 610–623.
5. Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Longmans.
6. Bruner, J. S. (1960). The process of education. Harvard University Press.
7. Carr, N. (2010). The shallows: What the Internet is doing to our brains. W. W. Norton.
8. Couldry, N., & Hepp, A. (2017). The mediated construction of reality. Polity Press.
9. Denzin, N. K. (1978). The research act: A theoretical introduction to sociological methods (2nd ed.). McGraw-Hill.
10. Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. American Philosophical Association.
11. Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., … Vayena, E. (2018). AI4People—An ethical framework for a good AI society. Minds and Machines, 28(4), 689–707.
12. Geertz, C. (1973). The interpretation of cultures. Basic Books.Hobbs, R. (2010). Digital and media literacy: A plan of action. Aspen Institute.
13. Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.
14. Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). Intelligence unleashed: An argument for AI in education. Pearson.
15. Patton, M. Q. (2002). Qualitative research and evaluation methods (3rd ed.). Sage.
16. Paul, R., & Elder, L. (2014). Critical thinking: Tools for taking charge of your professional and personal life (2nd ed.). Pearson.
17. Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688.
18. Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. International Journal of Artificial Intelligence in Education, 26(2), 582–599.
19. Sovorel. (2025, ZFFebruary 24). NO, AI Will Not Take Away Your Critical Thinking Skills. https://www.youtube.com/watch?v=l8EAHrxuAeQ
20. Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. Science, 333(6043), 776–778.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

21. Strauss, A., & Corbin, J. (1998). Basics of qualitative research: Techniques and procedures for developing grounded theory (2nd ed.). Sage.
22. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
23. Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(39). https://doi.org/10.1186/s41239-019-0171-0

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

APPENDIX

AI Governance Policy Manual: Bachelor of Arts Programs

St. Paul University Manila

1. Introduction and Philosophy

This policy manual outlines the governance and ethical use of Artificial Intelligence (AI) within the Communication Program at St. Paul University Manila. As a Paulinian Catholic institution, our approach to AI is grounded in human dignity, social responsibility, and the formation of ethical leaders. We recognize AI not as a neutral tool, but as a socially embedded technology that must be governed to resist techno-authoritarianism and cognitive atrophy.

2. Core Governance Principles

The Communication Program adheres to the following university-wide governance pillars:

- **Human-in-the-Loop (HITL):** Human judgment and accountability remain the final authority in all academic and administrative decisions.
- **Multisectoral Participation:** AI policies are shaped by a democratic process involving faculty, students, administrators, and community stakeholders.
- **Transparency and Disclosure:** All substantive uses of AI in research, teaching, and production must be clearly disclosed.
- **Mission Alignment:** AI integration must advance the common good and align with Catholic social teachings.

3. Pedagogical Model: "Learn First, Augment Later"

To protect the development of critical thinking, the Communication Program adopts a staged approach to AI integration:

3.1. Foundational Phase (Cognitive Scaffolding)

In early coursework (e.g., Introduction to Communication, Basic Writing), AI is strictly limited to serve as a "cognitive scaffold".

- **Objective:** To ensure students master foundational skills—such as information retrieval, basic synthesis, and independent writing—before using AI.
- **Restriction:** Direct AI generation of primary content is prohibited during this phase.

3.2. Advanced Phase (Cognitive Partnership)

In higher-level courses (e.g., Communication Research, Media Production, Strategic Planning), AI transitions into a "cognitive partner".

- **Objective:** To augment intellectual work through data analysis, pattern recognition, and professional-grade synthesis.
- **Requirement:** Students must demonstrate active agency by critically verifying and interrogating all AI-generated outputs.

ISSN No. 2454-6194 | DOI: 10.51584/IJRIAS |Volume XI Issue V May 2026

4. Guidelines for AI Use in Communication Competencies

AI integration is mapped directly to the Program Outcomes (PO) of the Communication Program:

- **Information Literacy (PO1, PO5):** AI may be used to organize research data, but students must independently verify sources to counter "hallucinations" and bias.
- **Knowledge Production & Media Creation (PO2, PO3, PO6):** AI tools for editing, formatting, or image generation are permitted in advanced years, provided they are disclosed and do not replace the core creative process.
- **Ethics and Leadership (PO7, PO9):** Students must analyze the ethical implications of AI, including its role in surveillance and the digital divide.

5. Risk Management and Classification

The program utilizes a tiered risk management system for AI applications:

Risk Tier	Type of Contribution	Governance Requirement
Incidental	Spell-check, grammar correction, basic formatting.	No formal disclosure required.
Moderate	Brainstorming, outlining, or summarizing existing texts.	Mandatory disclosure in footnotes or acknowledgments.
Substantive	Significant data analysis or drafting of complex sections.	Requires prior faculty approval and an AI-use disclosure form.

6. Technical Governance: Open-Source vs. For-Profit

The program encourages a balanced use of AI technologies:

- **Open-Source AI (e.g., LLAMA, Stable Diffusion):** Prioritized for research requiring methodological transparency, data sovereignty, and local control.
- **For-Profit AI (e.g., ChatGPT, Claude):** Used for instructional scaffolding, editorial support, and professional efficiency, subject to strict boundary management.

7. Implementation Instruments

To operationalize this policy, the following instruments are mandated:

1. **AI-Use Disclosure Forms:** Standardized forms for students to document which AI tools were used and for what purpose.
2. **Syllabus Clauses:** Every course syllabus must explicitly state the allowed level of AI use (Scaffolding vs. Partnership).
3. **Process-Based Rubrics:** Faculty will shift from evaluating final outputs to assessing the process of cognition, including prompt logs and reflection papers.
4. **AI Audit Logs:** For advanced research projects, students must maintain a log of AI interactions for auditability.

This manual is subject to continuous review and adaptation by the Communication Program's AI Governance Committee to remain responsive to rapid technological shifts