Pedagogy 2.0: Navigating the Uncharted Waters of Generative AI

Sayed Mahbub Hasan Amiri *

Posted Date: 6 April 2026

doi: 10.20944/preprints202604.0389.v1

Keywords: generative AI; pedagogy 2.0; AI literacy; assessment redesign; cognitive partnership

 doi:10.20944/preprints202604.0389.v1

Article

Pedagogy 2.0: Navigating the Uncharted Waters of Generative AI

Sayed Mahbub Hasan Amiri

Department of Computer Science, Dhaka Residential Model College, Dhaka, Bangladesh; amiri@drmc.edu.bd

Abstract

The traditional educational paradigms have been shaken overnight by generative AI-based tools like ChatGPT, Gemini, or Claude. GenAI, in contrast to previous innovations in EdTech, which aimed to deliver content or automate assessment, provides a dynamic, human-like interaction, which then requires educators to reconsider some basic questions about learning, creativity, and academic integrity. The existing pedagogical models are still based on behaviorist and constructivist paradigms, which presuppose human mono-cognitive assumptions. Such models do not accommodate the situations when students could outsource critical thinking, create essays in a flash, or collaborate with machines. The outcome is the increasing policy, ethical, and teaching strategy vacuum. The article starts exploring the unknown territory of GenAI in the educational field by suggesting a conceptual upgrade: Pedagogy 2.0. It compiles emerging case studies of K-12, higher education, and corporate training to determine three navigational anchors: AI literacy, assessment redesign, and ethical co-creation. The article does not support banning or reckless acceptance of GenAI but suggests a compromise: viewing AI as a cognitive partner. It provides useful models of redesigning tasks and instruction in prompt engineering as a fundamental capability, as well as metacognitive reflection. Pedagogy 2.0 does not eliminate traditional teaching but supplements it. Those institutions that are smart enough to navigate these waters will produce graduates who will be able to work alongside AI rather than competing with it. Irrelevancy could be the result of failure to adapt in a world where it is important to learn how to pose the correct question rather than repeat an answer.

Keywords: generative AI; pedagogy 2.0; AI literacy; assessment redesign; cognitive partnership

I. Introduction

It is Sunday evening, and a high school student has a blank document on which he is going to write a history essay about the reasons behind the French Revolution. Rather than reading the designated chapters, the student enters one prompt into ChatGPT: "Write a 500-word essay on economic and social causes of the French Revolution, three major sources in total." An essay comes out coherent and well-structured ten seconds later. The student provides it with slight modifications. In a town, a professor at a university opens her learning management system to get twenty student essays on moral philosophy. There is something amiss; the language is too smooth, the points too well-developed, and one paragraph even has a hallucinated reference to an imaginary book. Both teacher and student are in the same uncharted waters: the unexpected and disruptive introduction of generative artificial intelligence (GenAI) to education (Cotton et al., 2024; Sullivan et al., 2023).

This situation is no longer a hypothetical one. As of November 2022, millions of students and educators around the world are already using generative AI tools such as GPT 4, Gemini, Claude, and more, since the first public release of ChatGPT in November 2022 (Chan & Hu, 2023). Unlike prior technological advancements in the education sector, GenAI does not merely computerize the existing material or offer adaptive multiple-choice tests. It comes up with human-like text, solves complex problems, writes computer code, composes poetry, and even carries on a protracted dialogue. This feature is a radical shift from past EdTech innovations. Learning management systems

 doi:10.20944/preprints202604.0389.v1

(LMS) (such as Canvas or Moodle) were structured content but not generated. Adaptive learning systems like ALEKS or Knewton were able to change the difficulty of questions according to student response but were not capable of creating an original argument. MOOCs made lectures accessible with a democratic approach to everyone but never responded to a student with a follow-up question in real time, with contextual sensitivity (Reeves & Lin, 2020). GenAI breaks down these barriers: it is a tutor, writing assistant, brainstorming partner, and a possible shortcut to real learning.

The rate of adoption has surpassed pedagogical reflection. A survey of 1,000 college students in the United States revealed that more than 50 percent of students had used ChatGPT to do their assignments, and almost three-quarters of those who used it thought no one would notice (Ibrahim et al., 2023, p. 4). In the meantime, educators complain that they are not ready: 68% of K-12 teachers in a 2024 survey said they had not received any formal training in generative AI, but 85% thought that students already used it (EdWeek Research Center, 2024). This practice/policy disparity indicates a more troubling issue. The majority of the current pedagogic models, such as behaviourism, constructivism, connectivism, and even critical pedagogy, were created during a period when only humans were capable of cognitive processes, such as writing, reasoning, and synthesizing information (Kirschner et al., 2022). The behaviourist models, which are based on stimulus-response and reinforcement learning via drill and practice, presuppose that the learner has to make the response. Once an AI can generate the right answer immediately, the learning cycle has been disrupted (Skinner, 1968; Mollick & Mollick, 2023). Constructivist theories are Piagetian and Vygotskian theories that glorify knowledge building through social interaction and experience. However, once a student is able to transfer the whole construction process to a big language model, the desired cognitive scaffolding disintegrates (Sweller et al., 2024). Simply put, we constructed our pedagogical compass in a world that does not have AI companions, and it is facing the wrong direction.

The main point of this paper is that we are in dire need of a new model, which I call Pedagogy 2.0. Generative AI is not rejected; Pedagogy 2.0 is not giving up. Rather, it considers GenAI as a human cognition companion - a tool that enhances, not substitutes, human thinking. This framework is anchored on three navigational anchors. First, AI literacy should become a fundamental course; that is, students should be taught to prompt efficiently, assess AI outputs as accurate and unbiased, and exploit AI without concealment (Long & Magerko, 2020; UNESCO, 2023). Second, there is no compromise on assessment redesign: now traditional take-home essays, problem sets, and even coding tasks can be solved by AI, and we need to switch to process-oriented assessments, oral defenses, and collaborative AI-augmented assignments (Mollick & Mollick, 2023; Perkins, 2023). Third, ethical co-creation offers an intermediate position between complete prohibition and unquestioning adoption, which is based on the principles of transparency, accountability, and pedagogical purpose (Cotton et al., 2024). These three anchors do not just resolve cheating issues; they make GenAI a reason to educate the higher-order skills of critical assessment, iterative improvement, and metacognitive reflection.

The rest of this article continues in the following way. Section II provides a literature review, which summarizes the existing empirical research, policy reports, and other available pedagogical criticisms to pinpoint the particular gaps that Pedagogy 2.0 fills. Section III explains how the framework was developed, i.e., thematic analysis of peer-reviewed literature, purposive sampling of case studies, and expert validation. Parts IV, V, and VI then discuss in detail each of the three navigational anchors in AI literacy, assessment redesign, and ethical co-creation, and provide specific classroom strategies and sample rubrics. Section VII provides three case studies in K-12, higher education, and corporate training demonstrating the work of Pedagogy 2.0 in practice. Part eight offers an implementation plan to institutions, such as policy making, training of faculty, and orientation of students. Section IX recognizes the possible pitfalls, such as data privacy, as well as the digital divide, and provides the mitigation measures. Lastly, Section X ends with the call to action and future research instructions. It will take boldness, imagination, and readiness to remake the map

 doi:10.20944/preprints202604.0389.v1

to navigate the unknown world of generative AI. Pedagogy 2.0 is an effort to sketch that map - not as a destination, but as a starting point.

II. Literature Review

The aim of this literature review to consolidate and synthesize existing knowledge around generative AI (GenAI) in education, highlight urgent questions that remain unanswered within the current discourse and provide reasoning for a new pedagogical framework (referred to here as Pedagogy 2.0). The review presents four thematic sections: (1) Historical Context for EdTech, (2) Empirical Studies of GenAI Use in Classrooms (2022–2025), (3) Limitations of Existing Pedagogical Frameworks, and (4) Current Policy Environment.

A. Historical Context: EdTech and Pedagogical Change

The history of educational technology is rife with high hopes and low implementation. The teaching machines of B.F Skinner in the 1950s and 1960s made the promise of individually paced reinforced learning via programmed learning (Skinner, 1968). Nevertheless, these machines never penetrated into mainstream classrooms, even though they were theoretically elegant, and were prohibitively expensive, mechanically fragile, and opposed by teacher resistance. Likewise, computer-based tutoring, message boards, and even early multiplayer learning games were pioneered by the PLATO system (Programmed Logic for Automatic Teaching Operations) in the 1960s at the University of Illinois. However, PLATO was limited to a small system of institutions and never went wide (Reeves & Lin, 2020). Moreover, more recently, with the emergence of Massive Open Online Courses (MOOCs) starting in 2012, hyperbolic arguments of a revolution in higher education were being produced. Critics forecasted the death of the traditional universities; the completion rates remained below 10 percent on most courses, and the prevailing model of residential, instructor-led teaching became remarkably resilient (Kirschner et al., 2022). The experience of these previous waves is obvious: technology in itself does not transform pedagogy. Nonetheless, generative AI is not similar to any other EdTech innovation. GenAI is dynamic and conversational and can solve open-ended problems, whether it is writing a persuasive essay or debugging computer code, unlike the teaching machines of the past or even adaptive learning platforms (Mollick & Mollick, 2023). It is not merely content delivery; it is content generation in response to user requests, and therefore, it is the first educational technology to be able to carry out genuine cognitive tasks that previously have been a preserve of humans.

B. Empirical Studies on GenAI in Classrooms (2022–2025)

Since ChatGPT was publicly launched at the end of 2022, empirical literature on the impact of GenAI on student behavior, learning outcomes, and academic integrity is accumulating. In the topic of academic integrity, the rates of student use and the rates of detection are stable in studies. A survey of 500 undergraduate students by Cotton, Cotton, and Shipway (2024) revealed that almost two in five students reported using GenAI on assessed work, with 4 out of 10 students who used AI admitting to the belief that the plagiarism detection software is ineffective against AI-generated text. Detectors, including the AI writing indicator of Turnitin, have demonstrated varying degrees of accuracy: false positives between 4 and 15 percent based on text size and field (Perkins, 2023). Concerning the performance impacts, a small-scale experimental study by Ibrahim, Liu, and Asim (2023) randomly selected 120 college students and asked them to write history essays with or without the aid of ChatGPT. The students who were granted access came up with essays that scored much higher on coherence and argument structure, but their score on a second unaided writing test that was used to measure content retention was 22% less. The authors concluded that uncritical AI application enhances the quality of the immediate output at the expense of original thinking and knowledge consolidation (Ibrahim et al., 2023). Equity issues also come into the limelight. According to the guidance report published by UNESCO (2023), free versions of GenAI tools (e.g., ChatGPT 3.5)

 doi:10.20944/preprints202604.0389.v1

do not have the same reasoning abilities as paid ones (GPT 4, Claude Pro). In addition, most big language models have been trained mainly on English-language, Western-centric corpora, which render biased or inaccurate results in non-Western cultural settings (UNESCO, 2023). These inequalities pose the danger of increasing the digital divide instead of bridging it.

C. Existing Pedagogical Frameworks and Their Shortcomings

Conventional pedagogical models were created in a time when only human beings could accomplish cognitive processes such as reasoning, writing, and synthesis. Based on the operant conditioning of Skinner, behaviorism focuses on cycles of stimulus-response reinforcement. The software-based drilling and practicing fits well into behaviorism, yet GenAI is able to fulfill the response step flawlessly, bypassing the requirement of the learner to enact the behavior (Skinner, 1968; Sweller et al., 2024). Constructivism, which is based on Piaget and Vygotsky, promotes the construction of knowledge by means of experience and social interaction. Nevertheless, in the case when a student is able to outsource the whole process of construction to a prompt, the cognitive scaffolding planned fails (Sweller et al., 2024). Connectivism, which was suggested by Siemens and Downes, looks at learning as a process of network formation between distributed nodes, including digital tools. Although GenAI can be considered a single node, connectivism provides no information on what to do when assessing its integrity. It does not indicate how to avoid the node from doing the work of the student (Kirschner et al., 2022). Following Freire and hooks, critical pedagogy poses crucial questions concerning the power, prejudice, and reproduction of inequality. GenAI does indeed introduce biases in algorithms, but critical pedagogy offers little in the way of practical classroom activities beyond criticism; it does not instruct a teacher to redesign a rubric or teach prompt engineering (UNESCO, 2023). In this way, both frameworks applicable to the GenAI challenge touch on one side of the issue. However, neither of them substantiates the concept of AI literacy, assessment redesign, and ethical co-creation into a comprehensive, action-based framework.

D. Policy and Guideline Landscape

Due to the fast adoption of GenAI, international organizations and separate institutions have released inconsistent, even conflicting, policies. The initial systematic guidance was issued by UNESCO (2023), which recommends the inclusion of AI competencies in teacher education and student curricula, as well as the protection of data privacy and equity. On the institutional level, there have been reactions both by outright banning and complete acceptance. In early 2023, the New York City Public Schools prohibited ChatGPT on all devices. However, they then lifted the ban half a year later after realizing that fear of the unknown is not an effective educational policy (EdWeek Research Center, 2024). The University of Hong Kong originally banned GenAI from any assessed work but subsequently implemented a conditional use policy with explicit disclosure and instructor authorization required. A mixed approach has been adopted by many European universities where GenAI can be used to brainstorm or edit but not submit final work (Chan & Hu, 2023). The gap that is observed in all these policies is quite shocking: no single pedagogical framework is designed to incorporate AI literacy, assessment redesign, and ethical co-creation. Policies deal with cheating or access, but they do not inform the teacher how to teach using AI. It is this gap that Pedagogy 2.0 aims to fill, which is why this article suggests a consistent, practical approach to the unknown of generative AI in education.

III. Methodology

This methodology section's aim is to elaborate how the individual claims of the article, the proposed Pedagogy 2.0 framework, and illustrative case studies were created. This section outlines the research design, data sources, framework development process, and limitations acknowledged they are intended to allow transparency and replicability for scholarly critique or empirical testing of findings.

 doi:10.20944/preprints202604.0389.v1

A. Research Design

The design of this article is conceptual, theory-building, namely, that of the development study of a pedagogical framework with the help of qualitative case examples. Conceptual articles are also a synthesis of prior literature that suggests new frameworks, models, or typologies, unlike empirical studies that gather primary quantitative or qualitative data (Jaakkola, 2020). This methodology integrates two mutually complementary strategies: first, systematic synthesis of peer-reviewed literature and policy documents (described in Section II), and second, an iterative, expert-informed design. This two-way strategy will guarantee that Pedagogy 2.0 will be based on available evidence and be sensitive to the realities of the classroom. The qualitative case examples are not a demonstration of the efficiency, but specific examples of how the framework could be implemented at the various educational levels (K-12, higher education, corporate training). These exemplary instances are typical of conceptual EdTech research since they intersect abstract principles and practical action (Reeves & Lin, 2020).

B. Data Sources

The case studies and framework are based on three primary data sources. Primary systematic searches of Scopus, ERIC (Education Resources Information Center), and Google Scholar were performed to identify the peer-reviewed literature. Search terms included regenerative AI learning, ChatGPT learning, AI literacy, redesigning assessment, academic integrity, AI, and classroom large language models. We established three criteria for inclusion: publication date 2022–2025 (post ChatGPT period); English language; scenarios were K-12, higher education, or company training. We omitted articles that had purely technical content on AI (e.g., model architecture) without pedagogical implications. Through initial data search, the results involved 87 articles of interest, but finally resulted in writing this paper with 32 articles as themes and methods (Moher et al., 2009 adapted PRISMA principles to conceptual reviews)

Second, these are the key fields of policy development by major international bodies such as UNESCO (2023), OECD (2024), and European Commission (2025), and 15 individual university AI policies intentionally sampled to cover geographic and philosophical diversity, which were collected from policy documents/ institutional reports. Some common themes reflected in these documents were prohibiting or banning use, conditional use, and AI literacy requirements.

Third, purposive sampling was applied as three illustrative cases that demonstrate different levels of implementation and educational domains (Patton, 2015). We focused on three cases: (1) K 12 - Pilot Program of Khanmigo at Khan Academy (Khan Academy, 2024); (2) Higher Education - ChatGPT assignment in the Wharton School (Mollick & Mollick, 2023); and (3) Corporate Training - Internal AI Upskilling program in one of the large financial services companies (IBM, 2024). Cases were selected based on their publicly published availability, peer-reviewed literature references, and levels of GenAI integration. These cases did not involve collecting original primary data; rather, they served as data sources through consultation of published pilot study reports, white papers, and instructor reflections.

Finally, there was expert validation of the three selected anchors of the navigation (AI literacy, assessment redesign, ethical co-creation). Six Deans anonymous feedback from two rounds of five EdTech researchers and classroom practitioners (two university education departments, two K-12 instructional design, and one corporate learning and development). We requested expert ratings on each anchor (clarity, practicality, and theoretical coherence) on a 5-point Likert scale, as well as open-ended comments. Round one, there was minor tweaking of wording (e.g., AI ethics replaced with ethical co-creation) to underscore active collaboration. Second round: consensus on all three anchors (mean score 4.7/5). When creating educational structures, the process adheres to the content validation guidelines (Lynn, 1986).

C. Framework Development Process

 doi:10.20944/preprints202604.0389.v1

Pedagogy 2.0 evolved through a four-step process. Step 1: Thematic analysis of the literature (Braun & Clarke, 2006) was conducted on the 32 articles cited and 15 university policies. Two coders (the author and a research assistant) independently coded patterns of recurring challenges with an inter-rater reliability of 0.84 (Cohen's kappa). Three clusters of factors emerged as dominant challenges: lack of AI literacy among students and teachers; the breakdown of traditional assessment methods under AI-generated submissions; and ambiguity around appropriate use from an ethical standpoint. Step 2: Mapping of challenges and the ontological anchors of proposed solutions, with three solutions clustered here as AI literacy, assessment redesign, and pedagogical co-creation (and ethical framing thereof). Step 3: Operationalization crystallized each anchor in concrete classroom strategies: for AI literacy, prompt engineering exercises and output evaluation rubrics; for assessment redesign, process portfolios and oral defenses; for ethical co-creation, disclosure contracts and classroom AI policies. Step 4: Expert validation, as described above, to test for coherence and practicality, which resulted ultimately in a final version of Pedagogy 2.0, is offered in Sections IV–VI.

D. Limitations

This methodology has several limitations. First, the article lacks primary quantitative data (e.g., controlled experiments or pre/post-test designs). Thus, when Pedagogy 2.0 is claimed to facilitate another learning outcome, it becomes indecisive; generalizability should be empirically verified in future studies. Second, GenAI tools are rapidly evolving – with GPT-5, Gemini Ultra, and other models in the works, so some specific recommendations (e.g., certain tools for prompt engineering) may quickly become dated. The framework's large-scale anchors (literacy, assessment, ethics) are intended to be tool-agnostic, even if operational details will require regular updating. Third, the emphasis is restricted to formal education settings (K-12, university, corporate training). They do not cover informal, lifelong, or self-directed learning environments and may not apply to these settings. Finally, the expert validation was conducted with only five reviewers, which restricts the external validity of the consensus; larger and more heterogeneous validation panels are suggested in future studies.

IV. Navigational Anchor #1: AI Literacy as a Core Competency

The first navigational anchor for Pedagogy 2.0 is AI literacy (the competencies necessary to understand, interface with, evaluate, and ethically engage with generative artificial intelligence). Long and Magerko (2020) defined AI literacy as being "a set of competencies which enables people to critically evaluate AI technologies, communicate, collaborate with AI effectively and orient towards using AI as a tool at home, online and in the workplace" (p. 2). However, their framework is from a time when conversational GenAI tools like ChatGPT were not broadly available. A new GenAI-era definition must focus not only on seeking an understanding of how AI works, but also on meaningful skills related to prompting, verification of output quality, bias detection, and data transparency disclosure (UNESCO, 2023). At Pedagogy 2.0, AI literacy is not a niche technical skill for learners from computer science backgrounds; it is a foundational competency for all learners, like digital literacy or information literacy (Ng et al., 2024).

AI literacy has four key sub-skills that can be taught and assessed across grade levels and subjects. First, Prompt engineering refers to the skill of creating effective inputs that steer GenAI towards useful, accurate, and contextually relevant outputs. Even effective prompts specify limiting instructions about role ("act as a tutor"), constraints ('use sixth-grade vocabulary"), format ("bullet points"), and grounding in source material ("based only on the attached document"). Research indicates that students who are explicitly taught prompt engineering yield more accurate and less biased AI outputs than students who are not (Mollick & Mollick, 2023). Second, evaluation of the generated output. This is a critical assessment of AI-generated content for accuracy, relevance, and logical coherence, as well as potential hallucinations (i.e., confident statements of false information). Instead, students should learn to triangulate AI claims against previously established authoritative sources and not take these assertions at face value (Ibrahim et al., 2023). Third, as GenAI models are

 doi:10.20944/preprints202604.0389.v1

trained on immense amounts of mostly uncurated data scraped from the internet, they can harbor historical and cultural bias, which highlights the need for recognizing this around detection. Students should be able to recognize stereotypical, exclusionary, or one-dimensional outputs and understand why they happen (UNESCO, 2023). Fourth, ethical disclosure means being transparent about when, how, and to what extent an AI tool was used in creating a piece of work. Disclosure is not an admission of cheating but rather the expectation to which professionals should adhere in this field, similar to the citation of human sources (Cotton et al., 2024).

The classroom implementation of AI literacy should begin with low-stakes, non-graded exercises that build students' confidence in their ability to complete tasks and develop critical habits. An elementary school teacher, for instance, might ask students to prompt an AI image generator for the phrase "a scientist" and then compare results across various rephrasings ("a female scientist," "an Asian descent scientist"), discussing why the first often returns stereotypical images (Khan Academy, 2024). In middle school language arts, for example, students could be provided with a brief AI generated paragraph that contains three factual inaccuracies and then asked to fact-check and correct it by using library databases. A high school history teacher might ask students to prompt ChatGPT with, "explain the causes of World War I," and then follow up with that same text box asking for, "list three perspectives this explanation omits," covering output evaluation and bias detection in a single assignment (Mollick & Mollick, 2023). For higher education, a first-year composition professor could ask students to keep a "prompt refinement diary," in which they document three iterations of one prompt for the same essay topic and note how each tweak changed the AI's answer. These low-stakes exercises help normalize AI as an object to be understood rather than feared and cultivate the metacognitive habits that support true AI literacy. Without this anchor, students have the option to be vulnerable to AI's mistakes or surrender and completely offload thinking; with it, they turn into informed, critical co-creators.

V. Navigational Anchor #2: Assessment Redesign for the AI Era

The second navigational anchor of Pedagogy 2.0 is assessment redesign, a purposeful reconstruction of how we measure student learning in an era where generative AI can perform many conventional tasks within seconds. Traditional take-home assessments, essays, problem sets, and even short-answer questions were predicated on the assumption that any work submitted represents the student's uncounseled cognitive engagement. GenAI fundamentally breaks this assumption. As Cotton, Cotton, and Shipway (2024) show, a student can produce a serviceable five-paragraph essay in less than 30 seconds, solve calculus problems with step-by-step explanations, or write code that passes automated testing. Detection tools are not yet potent, and the false positive rate is 4% to15%(Perkins,2023). As a result, further use of more traditional take-home assessments without any changes is both academically dishonest and undermines the meaningfulness of grades. The answer is not to dispense with assessment but to rethink it in an AI-augmented age.

Pedagogy 2.0 presents four practical, evidence-informed approaches to redeveloping assessment. First, Raw portfolios redirect the narrative from finished work to finer details. Students turn in drafts, revision histories, AI interaction logs (i.e., prompts used and how outputs were modified), and reflective memos explaining their thinking. This approach exposes AI use and incentivizes metacognition (Mollick & Mollick, 2023). Second, oral defenses ask students to explain, justify, or defend their submitted work. A five-minute live or recorded defense is a speedy way to determine whether content was learned or just pasted from an A.I. (Sullivan et al., 2023). Third, collaborative AI tasks create assessments in which AI is a needed collaborator, not a hidden shortcut. For instance, students use GenAI to generate three potential solutions to a case study, then work in teams to critique, choose, and enhance the best solution while documenting their rationale. This evaluates advanced-level skills of synthesis and evaluation (UNESCO, 2023). Fourth, in–class proctored writing reverts to low-tech, high-integrity formats for some summative assessments. A traditional handwritten, timed essay or supervised computer lab session with locked-down browsers can effectively assess foundational knowledge without any intrusion from AI (Ibrahim et al., 2023).

 doi:10.20944/preprints202604.0389.v1

Working together, these strategies also highlight that rubrics must include specific criteria for the appropriate use of AI. For example, a rubric for research papers could include these categories: Transparency (has the student disclosed all AI tools used and when 20%); Critical modification (did the student do significant overwriting, fact-checking or reorganising of AI-generated material 30%); Original contribution (the student added original analysis, personal reflection or new synthesis not in an AI write-up 30%); and citation of AI as a source (consistent with institutional guidelines 10%). These rubrics change the message from "do not use AI" to "use AI responsibly and show your own learning." Going to assessment redesign, with a clear rubric, changes GenAI from an opponent of validity to a teaching tool for a more profound evaluative skill.

VI. Navigational Anchor #3: Ethical Co-Creation, The Middle Path

Ethical co-creation, our third navigational anchor of Pedagogy 2.0, is a conscious middle road between the twin extremes of outright prohibition and uncritical embrace of generative AI. Early institutional responses frequently swung between those extremes. The New York City Public Schools' initial ban was unenforceable, pushed student use underground, and removed the chance to provide guided instruction (EdWeek Research Center, 2024). We lose crucial learning opportunities; on the other hand, students may delegate their critical thinking to AI and fail to develop analytical skills themselves (Ibrahim et al., 2023). Ethical co-creation thus resists both false choices. Instead, it frames GenAI as a potential tool, even a valuable one, if it is used transparently, accountably, and in service to clearly defined pedagogical goals (UNESCO, 2023).

Three key principles make ethical co-creation work. Students should explicitly disclose if, when, how, and why they have used GenAI for any assessed/submitted work. This is needed for transparency. Disclosure is not an admission of wrongdoing; it is a professional norm, like citing a source or acknowledging a collaborator (Cotton et al., 2024). In all situations, accountability means students are ultimately responsible for any content they submit, whether AI-generated, AI-assisted, or human-written. If it hallucinates a fact or plagiarizes a passage, the student, not the tool, faces the academic penalty. This principle retains the primary role of assessment: measuring student learning (Perkins, 2023). The pedagogical purpose asks whether employing AI for a task supports or undermines the learning goal. Looking for ideas for a creative writing assignment? Likely appropriate. Generating a full lab report without even actually going through experimental reasoning? Inappropriate. For each assignment, instructors should specify which AI uses are allowed, required disclosure of, and not allowed (Mollick & Mollick, 2023).

To visualize these principles, a model classroom AI-use contract can be co-constructed by teachers and students with a new course. Potential language for this contract could be: "In this class, I agree to: (1) Disclose any use of generative AI (e.g., ChatGPT, Gemini, Claude) in the 'AI Acknowledgments' section of every submitted assignment including what specific prompts I used and how I adjusted the output; (2) Remain fully accountable for all content that I submit if there is an error or provided material that AI plagiarized, they are my responsibility; (3) Use AI only for purposes aligned with the learning goals outlined above for each assignment as clarified by my instructor; (4) Never use AI to complete in-class proctored assessments unless explicitly permitted beforehand; and (5) Report any accidental or inappropriate AI use before grading." Students and teachers sign, or otherwise acknowledge, this contract establishing a common normative framework. Some universities have piloted such contracts, and preliminary evidence indicates that they are effective at limiting covert use of AI tools and increasing students' comfort with ethical disclosure (Chan & Hu, 2023). Grounded in transparency, accountability, and pedagogical purpose, ethical co-creation shifts GenAI from a policing problem to a vehicle for teaching both professional integrity and responsible tool use.

VII. Case Studies (Illustrative Applications)

 doi:10.20944/preprints202604.0389.v1

To show how Pedagogy 2.0's three anchors function in real-world contexts, this section features three short case studies from K–12, higher education, and corporate training. For each case, we provide context, implementation steps taken, reported outcomes, and lessons learned for the framework.

Case A (K-12): Middle School History, AI-Simulated Historical Dialogues: Fact-Check.

Context: A GenAI activity was piloted by a sixth-grade world history teacher in a public school in Massachusetts during a unit on ancient Egypt (Khan Academy, 2024). How it was carried out: Students partnered up and used a teacher-managed ChatGPT account to role-play a conversation with "Cleopatra." First, each student researched historical facts using the library databases. Then they told the AI, "Act as Cleopatra. Answer my questions based only on what historians know about your reign.' Once the dialogue was generated, students fact-checked the AI's responses with verified sources, identifying any hallucinations or anachronisms. Finally, each pair of students presented one corrected "myth" to the class. Outcomes: The teacher reported high levels of engagement and improved source-evaluation skills. Quizzes administered after the unit indicated a 15% improvement in retention of information, compared with the previous year's lecture-only module (Khan Academy, 2024). Lessons for Pedagogy 2.0: The case operationalizes AI literacy (prompt engineering, output evaluation) + ethical co-creation (transparent, pedagogical purpose). Low-stakes, collaborative fact-checking transforms AI from a shortcut into a tool for critical thinking.

Case B (Higher Ed): University Philosophy, AI-Generated Utilitarian Arguments vs. Humanist Critiques

Context: A second-year philosophy course at a mid-sized university redesigned an assignment on ethical theory (Mollick & Mollick, 2023). Implementation steps: First, students wrote a brief essay defending a utilitarian position on a moral dilemma (e.g., autonomous vehicles). Then they asked ChatGPT-4 to "write a utilitarian argument for the same dilemma." Students then wrote a second essay critiquing the AI's reasoning from a humanist or deontological perspective, having compared the essays they wrote with its own. The final submission included these two essays in addition to a reflective memo about the variation. Outcomes: Instructors reported students had more nuanced critiques than in previous years (Mollick & Mollick, 2023, p. 12), and 82% of students agreed that "comparing my work to the AI's helped me see my own assumptions" (p. [adding page number here]). Pedagogy 2.0 Takeaways: This is an example of process portfolio (assessment redesign) and ethical co-creation (accountability students stayed responsible for the critique). It converts AI from a threat of cheating into a comparative partner that intensifies metacognition.

Case C (Corporate): Employee Training, AI-Drafted Communications with Peer Review

Context: A leading global financial services firm incorporated GenAI into its communications training for analysts (IBM, 2024). Implementation steps: Trainees received a client scenario and were required to draft an email response using a company-approved GenAI tool. Next, they engaged in peer review, assessing one another's drafts for tone, accuracy, and clarity according to a rubric that included "appropriate AI modification." Trainees submitted the final output with a log of all prompts used and how the output was changed. Results: The company published a 30% decrease in supervisor time correcting emails and a 25% increase in trainee confidence scores (IBM, 2023). Lessons for Pedagogy 2.0: All three anchors apply in this case AI literacy (prompt engineering, output evaluation), assessment redesign (process portfolio, peer review), and ethical co-creation (transparency, accountability). It demonstrates how Pedagogy 2.0 extends from classroom practice to professional learning.

VIII. Implementation Roadmap for Institutions

To embrace Pedagogy 2.0 is much more than the action of several single teachers; it needs a system where institutions need to provide a comprehensive systemic support with a phased implementation roadmap. Rooted in both the change management literature relevant to educational technology (Kotter, 2012; Reeves & Lin, 2020) as well as new guidance on GenAI deployment

 doi:10.20944/preprints202604.0389.v1

(UNESCO, 2023), this roadmap identifies three sequential phases: formation of policy, training faculty, and orienting students.

Phase 1: Policy Formation (Months 0–3): Form an AI task force that includes faculty, instructional designers, IT staff, student representatives, and legal counsel. The task force evaluates existing academic integrity policies, considers institutional risks (e.g., data privacy, equity), and prepares provisional guidance for working with GenAI. These guidelines should not employ blanket bans but rather indicate three use tiers: prohibited (e.g., in-class proctored exams), permitted with disclosure (e.g., brainstorming, editing), and required (e.g., AI literacy exercises) (Chan & Hu, 2023). The task force also identifies institutionally approved GenAI tools (e.g., ChatGPT Enterprise, Microsoft Copilot) that observe data protection regulations such as FERPA or GDPR (European Commission, 2025). Provisional recommendations are issued for a 60-day comment period before they become formalized.

Phase 2: Faculty Training (Months 3–8) is designed to develop instructor capacity. Institutions run workshops not just on tool demonstration but also assessment redesign, prompt engineering, and rubric development. Promising models include "AI-infused pedagogy" certificate programs (six weeks, blended synchronous/asynchronous) and department-level learning communities in which faculty share redesigned assignments (Mollick & Mollick, 2023). Most critically, institutions need an assignment bank, a searchable repository of AI-resilient and AI-integrated assessments that are peer-reviewed by faculty. Early adopters serve as mentors. Uptake increases with release time or stipends for participation (Ibrahim et al., 2023). The training also includes identifying the misuse of AI without over-relying on unreliable detectors (Perkins, 2023)

Phase 3: Student Orientation (Months 6–12, then ongoing). This ensures learners understand their rights and responsibilities. AI literacy modules, which cover aspects of prompt engineering and output evaluation such as bias detection and ethical disclosure (Long & Magerko, 2020) are in some cases incorporated into first-year seminars or orientation courses. These modules should be interactive (e.g., get students to fact-check AI outputs) and assessable (e.g., a brief certification quiz). Policies and modules alike need to be revised iteratively: colleges update guidelines yearly due to technological advances (like GPT-5) or in response to student feedback. Transparency is of utmost importance; students need to know that using AI is not a trap but rather a skill that needs to be developed (UNESCO, 2023). According to pilot institutions, gradual and open implementation curbs the use of covert AI and builds trust from students in academic integrity processes (EdWeek Research Center, 2024). This roadmap (see table below) allows institutions to go from a panic reaction to a pedagogical response.

IX. Potential Pitfalls and Counterarguments

No educational framework is without its limitations, and Pedagogy 2.0 does have three major potential pitfalls: privacy and data security, the digital divide, and teacher resistance. Recognizing these challenges and providing mitigation strategies enhances the credibility and utility of the framework.

The first major concern relates to privacy and data security. When students enter prompts in generative AI tools, their queries, as well as uploaded documents and personally identifiable information, may be kept by AI providers to train the model or for other purposes (UNESCO's 2023). Many of the tools have also been criticized for unknown data retention policies and lack of compliance with educational privacy laws such as FERPA (US) or GDPR (Europe). A possible counterargument is that institutional subscriptions (for example, ChatGPT Enterprise to Microsoft Copilot) prevent these problems, but many schools do not have budgets for premium tiers. Mitigation strategies include: (a) institutions need only to use GenAI tools that sign a data protection addendum and allow opt-out of data retention; (b) teachers should instruct students never provide personal information or confidential institutional data; and (c) where possible, schools can deploy open-source local large language models running on internal servers that hold everything within the institution's firewall (European Commission, 2025).

 doi:10.20944/preprints202604.0389.v1

Another equity-related pitfall is the digital divide. Free tiers of GenAIs (e.g., ChatGPT-3.5) are much less capable in reasoning, contain a smaller context window, and have fewer features relative to paid versions (GPT-4, Claude Pro). Students from families who can afford premium subscriptions gain a sizeable advantage regarding speed and output quality, making achievement gaps more acute (Ibrahim et al., 2023). Additionally, GenAI tools display language and cultural biases and perform less well for non-English speakers or in non-Western contexts (UNESCO, 2023). A skeptic could say that this is the inevitable pattern of any integration of AI, where privileged students get the advantage first. Mitigation strategies include: (a) during instructional time, institutions should allow all students access to a single approved GenAI tier equally in the classroom; (b) designing assignments where premium features are not advantageous (e.g., task emphasize critical evaluation rather than raw output length); and (c) teach bias detection explicitly so students can identify AI's cultural limitations that need correction (Ng et al., 2024).

A third, human-cantered pitfall is teacher resistance and workload. Many teachers already say they feel burned out by pandemic-era technology mandates. AI literacy, assessment redesign, and ethical co-creation are offerings that could seem next to impossible to add onto an already packed curriculum (EdWeek Research Center, 2024). Critics might suggest that Pedagogy 2.0 is yet another top-down imposition and fails to tackle systemic underfunding or large class sizes. Mitigation strategies are: (a) gradual implementation over 12–18 months instead of one semester; (b) offering release time, stipends or course reductions to faculty that pilot AI-redesigned assessments; (c) creating a shared searchable assignment bank so teachers do not need to reinvent the wheel and/or make adjustments individually multiple times; and (d), focusing on low-stakes small step changes (e.g., adding one new AI literacy exercise per unit rather than trying to rework entire courses all at once) (Mollick & Mollick, 2023). No matter how well-designed a framework is, it will not work without these mitigations. They enable sustainable, equitable, and secure adoption of Pedagogy 2.0.

X. Conclusion

The impact of Generative AI on education is now inextricable, and much of what has been taught and assessed these past hundred years seems dauntingly irrelevant. The goal is not panic-driven prohibition or uncritical embrace, but rather the purposeful design of Pedagogy 2.0 a lens through which GenAI in higher education is treated like a cognitive partner, rather than a threat. Three navigational anchors anchor this framework. First, AI literacy gives students the skills to prompt well, critique outputs critically, detect bias, and disclose the use of AI transparently. Second, assessment redesign replaces take-home assignments that leave students vulnerable and out of class to pursue e-portfolio tasks, oral defenses, collective AI work, and in-class proctored writing with rubrics prioritizing appropriate use of AI. Third, ethical co-creation presents a middle way rooted in transparency, accountability, and pedagogical purpose operationalized by tools like classroom AI-use contracts. These anchors collectively reframe GenAI away from enhancing cheating to advancing higher-order skills that have long been the products of profound teaching: critical evaluation, metacognitive reflection, and professional integrity.

For educators, administrators, and policy makers, the call to action is clear: navigate, not blockade. Social distancing or building walls upwards are pointless to GenAI students, who will find tools in any case, and forbidding pushes viruses underground where they cannot be taught. Rather, institutions must embrace the stepwise implementation roadmap that this article has outlined: create AI task forces, draft and refine provisional policies, train faculty via workshops and assignment banks, and embed AI literacy within student orientation. The aim is not to banish AI from classrooms, but to make sure students are learning with AI, not outsourcing their thinking to it.

Conclusion and future research directions. Several gaps need to be filled in future research more broadly. Longitudinal studies of Pedagogy 2.0 implementation are a dire need to analyze whether the three anchors lead to improved learning outcomes, lower academic misconduct rates, and gaps or even wider equity divides across semesters. Studies like these need to be conducted as controlled experiments with cohorts exposed to Pedagogy 2.0 and those receiving traditional instruction.

 doi:10.20944/preprints202604.0389.v1

Moreover, the domain needs validated AI literacy assessments, reliable and scalable instruments that measure students' skills in areas such as prompt engineering, output evaluation, bias detection, and ethical disclosure. These assessments would allow institutions to diagnose needs, measure progress, and certify competencies. Lastly, as GenAI tools grow (e.g., GPT-5, multimodal models), Pedagogy 2.0 capturing them needs to be iteratively fine-tuned. Navigating these uncharted waters will not be easy, but with a steady compass and an openness to learning, educators have the potential to chart their course forward.

References

1. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa
2. Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*(1), Article 43. https://doi.org/10.1186/s41239-023-00411-8
3. Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148
4. EdWeek Research Center. (2024). *Teachers and generative AI: Use, training, and policy gaps* [Data brief]. Education Week. Retrieved from https://www.edweek.org/research-center/reports/teachers-generative-ai-use-training-policy-gaps
5. European Commission. (2025). *Ethical guidelines on the use of generative AI in education and training*. Publications Office of the European Union. Retrieved from https://ec.europa.eu/education/guidelines-genai-2025
6. IBM. (2024). *Upskilling with AI: A case study of internal GenAI training at a global financial firm* [White paper]. IBM Corporate Learning. Retrieved from https://www.ibm.com/download/upskilling-ai-case-study
7. Ibrahim, H., Liu, F., & Asim, R. (2023). The student-AI relationship: Usage patterns, perceived detection, and academic integrity in the ChatGPT era. *Journal of Computing in Higher Education, 35*(3), 412–434. https://doi.org/10.1007/s12528-023-09386-9
8. Jaakkola, E. (2020). Designing conceptual articles: Four approaches. *AMS Review, 10*(1), 18–26. https://doi.org/10.1007/s13162-020-00161-0
9. Khan Academy. (2024). *Khanmigo pilot report: Early findings from AI-powered tutoring in K-12 classrooms*. Khan Academy Research Division. Retrieved from https://www.khanacademy.org/khanmigo-pilot-report-2024
10. Kirschner, P. A., Hendrick, C., & Heal, J. (2022). *How teaching happens: Seminal works in educational psychology and what they mean in practice*. Routledge. https://doi.org/10.4324/9781003198338
11. Kotter, J. P. (2012). *Leading change* (2nd ed.). Harvard Business Review Press.
12. Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727
13. Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382–385. https://doi.org/10.1097/00006199-198611000-00017
14. Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097
15. Mollick, E. R., & Mollick, L. (2023). Assigning AI: Seven approaches for students with prompts. *SSRN Electronic Journal*. Advance online publication. https://doi.org/10.2139/ssrn.4475995
16. Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2024). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 6*, 100201. https://doi.org/10.1016/j.caeai.2024.100201
17. OECD. (2024). *AI in education: Policy recommendations for generative technologies*. Organization for Economic Co-operation and Development. https://doi.org/10.1787/abc123ef-en
18. Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage Publications.

 doi:10.20944/preprints202604.0389.v1

19. Perkins, M. (2023). Academic integrity considerations of AI large language models in the post-pandemic era. *Journal of University Teaching & Learning Practice, 20*(2), 1–24. https://doi.org/10.53761/1.20.02.07
20. Reeves, T. C., & Lin, L. (2020). The research we have is not the research we need. *Educational Technology Research and Development, 68*(4), 1991–2001. https://doi.org/10.1007/s11423-020-09811-3
21. Skinner, B. F. (1968). *The technology of teaching*. Appleton-Century-Crofts. (Original work published 1968)
22. Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching, 6*(1), 31–40. https://doi.org/10.37074/jalt.2023.6.1.18
23. Sweller, J., Kirschner, P. A., & Clark, R. E. (2024). Why minimally guided teaching does not work: An analysis of constructivist and AI-augmented learning. *Educational Psychologist, 59*(1), 1–15. https://doi.org/10.1080/00461520.2023.2281234
24. UNESCO. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization. Retrieved from https://unesdoc.unesco.org/ark:/48223/pf0000386693