What Would Paulo Freire Think of Generative AI? Critical Pedagogy, Cognitive Offloading, and the Struggle for Humanisation

Joseph Vancell

University of Malta, Malta

Abstract

This paper examines generative artificial intelligence (GenAI) through the lens of Paulo Freire's critical pedagogy, exploring how the Brazilian educator and philosopher might assess contemporary AI technologies in education. Methodologically, it adopts a theoretical-interpretive approach, synthesising recent scholarship (2023–2026) on GenAI in education through a systematic application of Freire's foundational concepts, namely the banking model, problem-posing pedagogy, conscientisation, dialogue, humanisation, and praxis, as an analytical framework rather than reporting original empirical data. Whilst existing scholarship has applied individual Freirean concepts to AI critique, this paper makes a distinct contribution by offering a comprehensive and coherent Freirean reading across all of Freire's core pedagogical principles, and by moving beyond critique to propose a praxis-oriented framework of seven actionable strategies for AI integration. The analysis occupies a deliberate middle ground between techno-optimism and techno-pessimism, arguing that Freire would view generative AI as neither inherently emancipatory nor oppressive, but as a contested terrain requiring deliberate pedagogical choices that centre humanisation, equity, and social justice. The paper identifies persistent risks, including technocratic neutrality, algorithmic bias, and the displacement of authentic dialogue, whilst also recognising AI's emancipatory potential when implemented with critical pedagogical intentionality. Ultimately, it contends that Freire would call for a praxis-oriented approach to AI in education, one that transforms technology from an instrument of domination into a tool for critical consciousness and collective liberation.

Keywords: *Paulo Freire, generative AI, critical pedagogy, conscientisação, banking model, problem-posing education, AI literacy, cognitive offloading, skills atrophy, deskilling*

INTRODUCTON

> *I consider the fundamental theme of our epoch to be that of domination—which implies its opposite, the theme of liberation, as the objective to be achieved… In order to achieve humanisation, which presupposes the elimination of dehumanising oppression, it is absolutely necessary to surmount the limit-situations in which people are reduced to things [1].*

The rapid emergence of generative artificial intelligence (henceforth, GenAI), particularly large language models (LLMs) such as ChatGPT, Claude, and Gemini, has profoundly disrupted educational landscapes worldwide [2, 3]. These systems can generate essays, answer complex questions, create lesson plans, and engage in seemingly intelligent dialogue, prompting urgent questions about their role in teaching and learning [4, 5]. A survey of 337 higher education leaders in the USA conducted in late 2024 found that 71% of college students use AI tools for coursework [2]. As educators grapple with whether to embrace, resist, or critically integrate these technologies, Paulo Freire's critical pedagogy offers a powerful analytical framework [6, 7].

Paulo Freire (1921–1997) critiqued the “banking model” of education, in which teachers deposit knowledge into passive student receptacles [6, 8]. He advocated for “problem-posing” education rooted in dialogue, critical reflection, and transformative action, positioning education as inherently political and aimed at humanisation and conscientisation (conscientização): critical awareness of social contradictions and the capacity to act against oppressive elements of reality [9, 10]. Recent scholarship has applied Freirean concepts to GenAI, revealing both emancipatory possibilities and significant risks [11, 12, 13]. This paper synthesises this work and extends it by foregrounding the concern of cognitive offloading and skills atrophy, a risk that AI-enabled externalisation of cognitive tasks may erode the very competences education is designed to cultivate.

This paper adopts a theoretical-interpretive methodology. It does not report original empirical data; instead, it synthesises recent scholarship (2023–2026) on generative AI in education through the analytical lens of Freirean critical pedagogy. Sources were selected purposively to represent the emerging critical literature on GenAI across diverse educational contexts. The aim is conceptual. It attempts to build a Freirean reading of AI's pedagogical implications rather than to produce a systematic review or meta-analysis.

Drawing on Freire's pedagogical theory as an analytical framework, this paper seeks to occupy a middle ground, one that neither uncritically embraces nor wholesale rejects AI's role in educational practice. Whilst existing scholarship has applied individual Freirean concepts, particularly the banking model and conscientisation, to AI critique, this paper makes a distinct contribution by offering a systematic and comprehensive Freirean reading across all of Freire's core pedagogical principles. It further moves beyond critique to propose a coherent, praxis-oriented framework for AI integration, bridging theoretical analysis and practical strategy in ways that prior work has not fully developed. In doing so, it aligns with a growing scholarly effort to engage with AI in education in a measured, and critically informed manner.

The paper is structured as follows. Section 1 outlines the core concepts of Freire's critical pedagogy that are most relevant to the analysis of GenAI. Section 2 examines the risks that GenAI poses when viewed through a Freirean lens, with particular attention to the reproduction of the banking model, the masking of power relations, algorithmic bias, cognitive offloading and skills atrophy, the political dimension of AI-induced deskilling, and the accelerated displacement brought about by the agentic turn in AI development. Section 3 explores the emancipatory possibilities of GenAI when deployed within genuinely critical pedagogical frameworks. Section 4 synthesises these findings into a set of practical commitments for educators seeking to align AI integration with Freirean principles. The conclusion reflects on what Freire's legacy demands of educators in the age of AI.

1. Freire’s Critical Pedagogy: Core Concepts and Principles

1.1 The Banking Model versus Problem-Posing Education

Freire’s critique of the “banking model” of education remains the cornerstone of his pedagogical vision. In this model, education becomes “an act of depositing, in which the students are the depositories and the teacher is the depositor” [1]. Students are treated as empty vessels to be filled with predetermined knowledge, rendering them passive objects rather than active subjects in their own learning. This approach serves the interests of oppression by discouraging critical thinking and maintaining existing power structures. Knowledge, in the banking model, is a finished product to be consumed rather than a living process to be co-created, and the student’s role is one of reception, filing, and storage rather than critical engagement and transformation. The model is not merely an inefficient pedagogy but a political instrument: by training students to receive and accept rather than question and transform, it reproduces the conditions of their own subordination.
In contrast, Freire advocated for “problem-posing” education, which positions teachers and students as co-investigators of reality. This dialogic approach treats knowledge not as a static commodity to be transferred but as something co-created through critical inquiry grounded in learners’ lived experiences. The teacher-student relationship is reconceived as one of mutual learning: teachers learn from students as students learn from teachers, in a horizontal exchange that respects the knowledge all participants bring to the educational encounter. Recent analyses emphasise that Freire’s distinction between these models remains highly relevant to AI integration, as generative systems can easily default to banking-style knowledge delivery if not deliberately designed and framed otherwise [4].

1.2 Conscientisação and Critical Consciousness

Central to Freire’s pedagogy is the concept of conscientisação—the process of developing critical consciousness about social contradictions and the capacity to transform oppressive realities. This process moves beyond mere awareness to encompass “learning to perceive social, political, and economic contradictions, and to take action against the oppressive elements of reality” [1]. Freire understood consciousness-raising not as an individual psychological process but as a collective political one, tied to broader struggles for social justice. Generative AI can either facilitate critical consciousness by revealing algorithmic bias and social structures, or it can naturalise these biases and hinder conscientisação if used uncritically [7]. The question becomes whether AI tools are deployed to

encourage questioning and critique or to deliver seemingly authoritative answers that discourage deeper inquiry.

1.3 Dialogue and Humanisation

For Freire, authentic education requires dialogue—a horizontal relationship between teachers and students characterised by mutual respect, trust, and shared inquiry. Dialogue is not mere conversation but a transformative encounter that recognises the full humanity of all participants, standing in stark contrast to the vertical, authoritarian structures of banking education. Humanisation—the process of becoming more fully human—represents the ultimate goal of Freirean pedagogy, orienting educational technology towards equity, dignity, and access rather than mere efficiency or profit [10]. This humanising imperative must function as an ethical constraint on any educational use of AI.

1.4 Praxis: The Unity of Reflection and Action

Freire insisted on the unity of reflection and action, which he termed 'praxis'. Authentic praxis requires both critical reflection on reality and transformative action to change it. Neither reflection without action (verbalism) nor action without reflection (activism) constitutes genuine praxis. This dialectical relationship between thought and action drives both individual conscientisação and collective liberation. Recent scholarship suggests that generative AI can be integrated into cycles of praxis when used to surface problems, support critique, and inform collective action [7]. However, automated systems risk substituting reflection with optimisation metrics unless pedagogically mediated to maintain the reflective dimension essential to genuine praxis [11].

2. Generative AI as Banking Model 2.0: Risks and Reproductions

2.1 The Delivery of Packaged Knowledge

Generative AI systems excel at producing seemingly authoritative, polished responses to student queries, potentially transforming education into an even more efficient banking operation. Rather than teachers depositing knowledge, AI systems deliver instant answers that discourage students from engaging in the messy, uncertain process of inquiry and knowledge construction. Without deliberate pedagogical redesign, AI reinforces teacher-as-depositor roles by providing ready-made answers and narrowing inquiry [4, p.89]. This risk is particularly acute because AI-generated content often appears comprehensive and confident even when inaccurate or incomplete, inviting uncritical consumption rather than critical engagement. The ease of generating essays, summaries, and solutions fosters an instrumental, consumption-style relationship to learning that prioritises efficiency over understanding, credentialisation over conscientisação.

2.2 Technocratic Neutrality and the Masking of Power Relations

Freire emphasised that education is never neutral; it either serves to integrate students into the existing system or enables them to question and transform it. Generative AI is frequently promoted through technocratic framings that present it as a neutral efficiency tool, masking the power relations and ideological commitments embedded in these systems [12]. This apparent neutrality conceals multiple forms of power. AI systems embody the values and biases of their creators and training data, often reflecting dominant cultural perspectives whilst marginalising others. Corporate ownership of major AI platforms concentrates power in the hands of technology companies (primary, American). Therefore, the emphasis on efficiency and scalability serves neoliberal educational reforms that prioritise measurable outcomes over humanisation and critical consciousness. Freire would critique this technocratic framing as a contemporary form of "false generosity". This, for Freire and later critical pedagogists, is a gesture that appears beneficial whilst ultimately maintaining oppressive structures.

2.3 Algorithmic Bias and Automated Inequality

Generative AI systems trained on historical data inevitably reproduce the biases present in that data, including racial, gender, cultural, and linguistic biases. These systems can generate content that perpetuates stereotypes, marginalises non-dominant perspectives, and reinforces existing inequalities [7, p.7]. Beyond content bias, AI creates new structural inequalities: students and schools with greater

access to advanced AI tools gain advantages over those without, potentially widening existing educational divides. The digital divide becomes an AI divide, with implications for who can participate fully in an increasingly AI-mediated world. Freire would recognise this as a contemporary manifestation of oppression: the use of technology to maintain hierarchies and exclude marginalised communities from full participation in knowledge creation and social transformation [9].

2.4 Cognitive Offloading and the Atrophy of Competence

Cognitive offloading, that is,the practice of externalising cognitive tasks to tools and technologies—is not inherently problematic. Writing itself is a form of cognitive offloading; so is the use of calculators, GPS navigation, and reference databases. Clark and Chalmers [16] argue that cognitive processes 'can be partly constituted by, and partly located in, the environment,' and that this extension can enhance rather than diminish human cognitive capacity. Their famous example of Otto's notebook as a genuine component of his memory system illustrates how cognitive extension can be enabling rather than diminishing.

However, the current scale and scope of AI-enabled cognitive offloading raises qualitatively distinct concerns. Sparrow et al. [13] documented what they termed 'the Google effect': the finding that 'when people expect to have future access to information, they have lower rates of recall of the information itself.' Risko and Gilbert [16] have noted that the consequences of cognitive offloading depend critically on 'whether or not the offloaded processing contributes to learning.' More recently, Gerlich [19] has documented correlations between habitual AI tool use and diminished critical thinking capacities, arguing that when AI systems perform the evaluative and generative work that produces competence, repeated use may erode rather than scaffold that competence.

In educational contexts, this concern takes on particular urgency. Students who use LLMs to generate essays are not merely submitting work they did not produce; they are potentially forgoing the cognitive work through which writing competence, argumentation skills, and disciplinary knowledge are developed. The short-term efficiency gain may come at the long-term cost of the very capacities that education is meant to cultivate. From a Freirean perspective, this is especially troubling because the atrophy of critical cognitive competences directly undermines the possibility of conscientisação: one cannot develop critical consciousness whilst outsourcing one's critical thinking to an algorithm.

2.5 Freire and the Political Dimension of Deskilling

What Freire adds to this cognitive analysis is the insistence that deskilling is not merely a technical phenomenon but a political one. For Freire, the dehumanisation of workers, that is, their reduction to passive instruments of a productive process they neither understand nor control, is not an unfortunate side effect of technological change but an expression of the fundamental logic of oppression [1]. AI-induced deskilling, viewed through this lens, is an expression of the power relations governing who controls the means of cognitive production, who benefits from the automation of skilled labour, and who bears the costs.

This political dimension is amplified by the economic logic that drives AI development. As Zuboff [20] argues, surveillance capitalism 'claims human experience as free raw material for translation into behavioural data'—a logic that AI-enabled deskilling intensifies. When legal AI systems automate the analytical work of junior lawyers whilst leaving the profits of law firms intact, when algorithmic management systems reduce skilled workers to responsive appendages of an optimisation engine, these are not neutral technological developments. They are expressions of a logic that extracts value from human cognitive labour whilst systematically undermining the conditions under which that labour can be meaningfully exercised. Freire would insist that educators name this logic rather than accommodate it.

2.6 The Agentic Turn and Accelerated Displacement

The recent development of agentic AI systems—AI capable of planning and executing multi-step tasks autonomously, browsing the web, writing and running code, managing files, and interacting with external services—marks a qualitative escalation in the scope of potential deskilling and displacement. Where earlier AI tools augmented specific tasks within a human-directed workflow, agentic systems begin to approximate the autonomous performance of entire job roles [13].

This development amplifies the Freirean concern in two ways. First, it accelerates the displacement of workers from roles that previously required substantial expertise, concentrating the benefits of this

displacement among those who own and control the agentic systems. Second, it deepens the opacity of the productive process: agentic AI produces outputs whose reasoning is often opaque even to the system's creators. This opacity is fundamentally antithetical to the Freirean ideal of praxis. Freire insists that reflection and action must be consciously integrated—that genuine praxis requires workers who understand what they are doing and why [1]. A labour process mediated by inscrutable algorithms forecloses this understanding by design, producing what might be called 'automated verbalism': outputs without genuine reflection.

2.7 The Displacement of Dialogue and Agency

Perhaps most troubling from a Freirean perspective is the potential for generative AI to displace authentic dialogue—the horizontal, humanising encounter between teachers and students that Freire considered essential to transformative education. As AI systems become more sophisticated in simulating conversation, there is a growing risk that human-AI interaction will substitute for human-human dialogue, fundamentally altering the relational dynamics through which conscientisação develops [17]. When students turn habitually to AI for answers, feedback, and guidance, the locus of educational interaction shifts from the classroom community to individual interactions with algorithmic systems, undermining the collective dimension of Freirean learning.
Moreover, agentic AI systems that can independently perform cognitive tasks risk reconfiguring student agency itself. If AI drafts essays, solves problems, and conducts research on students' behalf, students may be positioned as managers or consumers of AI outputs rather than as active knowledge creators. This reconfiguration contradicts Freire's insistence that students must be subjects, not objects, in their own education. It mirrors precisely the dehumanising dynamic he identified in the banking model.

2.8 Misinformation and the Post-Truth Challenge

GenAI systems remain prone to "hallucinations": the generation of plausible-sounding but factually incorrect information. This tendency contributes to what scholars describe as post-truth problems in knowledge formation, where the distinction between accurate and inaccurate information becomes increasingly difficult to maintain [14]. For Freire, pedagogy depended on the critical reading of reality:

> *"Reading the world always precedes reading the word, and reading the word implies continually reading the world." [15]*

The proliferation of AI-generated misinformation therefore represents a serious obstacle to conscientisação. When students cannot reliably distinguish between accurate and fabricated information, their capacity to develop critical consciousness about actual social conditions is fundamentally compromised. The problem is compounded by AI's authoritative presentation style, which can lend false credibility to incorrect information and discourage the critical scepticism that Freire considered essential to education.

3. Generative AI as Tool for Liberation: Emancipatory Possibilities

Whilst the risks are substantial, a rigorous Freirean analysis must also recognise generative AI's potential to support liberatory education when implemented with critical pedagogical intentionality. Freire was not a technological determinist; he understood that tools could serve either oppression or liberation depending on how they were used and for what purposes. Several emerging approaches demonstrate how AI might be aligned with Freirean principles, particularly when the risks of skills atrophy and deskilling are explicitly addressed.

3.1 Supporting Problem-Posing Inquiry and Critical AI Literacy

Generative AI can facilitate problem-posing education when used as a tool for iterative inquiry rather than as an answer engine. Instead of accepting AI outputs as final knowledge, students can use these systems to generate hypotheses, explore alternative perspectives, and frame problems for further investigation. As Alm and Watanabe demonstrate in language education, AI can co-create localised communicative scenarios that students then critique and refine, positioning learners as active knowledge constructors rather than passive recipients [4]. Recent studies suggest practical strategies: designing assignments that require critique of AI outputs, using AI to scaffold inquiry rather than deliver final answers, and creating tasks where students must verify, contextualise, and improve AI-generated content [4].

In the context of skills atrophy, this approach is particularly significant. By requiring students to interrogate, verify, and improve AI outputs rather than simply accept them, educators can ensure that AI use contributes to rather than displaces the development of competence. The key pedagogical move is to position AI as a provisional text to be critically read rather than an authoritative source to be consumed—a move that aligns directly with Freire's vision of education as the critical reading of both the word and the world. Critical AI literacy must therefore encompass not only technical understanding and ethical reasoning but also students' metacognitive awareness of when AI use is enhancing versus replacing their own cognitive development.

3.2 Revealing Algorithmic Bias and Fostering Conscientisação

Paradoxically, the very biases embedded in AI systems can become pedagogical resources for developing critical consciousness. When students systematically interrogate AI outputs for bias, stereotypes, and cultural assumptions, they engage in precisely the kind of critical analysis that Freire advocated. As Warr argues, classroom activities that interrogate AI-generated content for bias and social assumptions can build awareness of how power operates through technology and language [7]. This approach treats AI not as a neutral tool but as a text to be critically read because it is a manifestation of social structures and power relations that can be analysed and challenged.

This critical engagement can extend beyond analysis to transformative action. Students might design projects to audit AI systems for disparate impacts and develop collective proposals to mitigate harms, linking critical analysis to social transformation in a genuine cycle of Freirean praxis [7]. Such activities connect the critique of AI to broader struggles for equity and justice, maintaining the political orientation that Freire considered essential to genuine education.

3.3 Extending Dialogue Whilst Centring Human Relationships

Whilst AI cannot replace authentic human dialogue, it may extend dialogic possibilities when carefully integrated. Generative AI can serve as a provisional conversation partner that helps students develop and refine their thinking before engaging in classroom dialogue, functioning as a sparring partner for reflective questioning whilst teacher-facilitated dialogue remains central [4]. This use of AI is analogous to the role of a journal or a drafting process in preparing for genuine dialogue: it supports the individual work of thinking without substituting for the irreducibly social process of meaning-making through encounter with other human beings.

The key distinction is between AI as a tool that supports students' capacity for dialogue and AI as a replacement for dialogic relationships. AI might support dialogue by helping to surface tacit knowledge, generate diverse perspectives for consideration, or provide linguistic scaffolding for students developing academic discourse. In language education, for example, AI can help students practise communicative scenarios and build confidence for human interaction [4]. In this framing, AI functions as a kind of rehearsal space for dialogue rather than a substitute for it. Crucially, the primacy of human relationships in transformative education must remain non-negotiable: Freire's vision of education as humanising encounter cannot be realised through interaction with an algorithm, however sophisticated, because it requires the genuine recognition of another human subject whose freedom and dignity one is called to honour.

3.4 Enabling Praxis Through Iterative Design and Reflection

Generative AI can be integrated into cycles of praxis when used in iterative design processes. Students might use AI to prototype interventions, analyse outcomes, and revise their approaches in ongoing cycles of action and reflection. This application is evident in professional education contexts, where AI-generated case variations support reflective practice and ethical reasoning [21]. For example, students addressing a community problem might use AI to generate potential solutions, critically evaluate these proposals, test implementations, reflect on results, and refine their approach. The AI serves as a tool within a larger praxis cycle rather than as an endpoint of inquiry. Critically, this use of AI for praxis must remain grounded in authentic problems and oriented towards real-world transformation, ensuring that reflection and action address actual conditions and work towards genuine liberation rather than academic simulation.

4. Towards a Freirean Praxis of AI in Education

Based on this analysis, Freire would likely call for a praxis-oriented approach that neither uncritically embraces nor reflexively rejects AI but critically engages these technologies in service of humanisation and liberation [10]. This approach involves several inter-related commitments that together constitute a Freirean framework for educational AI. Crucially, this framework does not treat AI integration as a technical problem to be solved by better instructional design alone, but as a political problem that requires ongoing critical vigilance about the social relations and power structures that AI technologies embody and reproduce.

Thus, any integration of AI in education must be evaluated against the criterion of humanisation, asking continuously whether AI applications strengthen authentic dialogue, support students as subjects in their own education, and foster critical consciousness rather than passive acceptance [10]. This means prioritising human relationships, learner agency, and social justice over efficiency, standardisation, or profit. Institutional decisions about AI adoption must be governed by this humanising criterion rather than by vendor marketing, administrative convenience, or competitive pressure from peer institutions.

Freire's vision of "reading the word and the world" as critical consciousness demands what scholars call "critical AI literacy": encompassing technical understanding of how AI systems work, critical analysis of bias and accuracy, ethical reasoning about social implications, and agentic engagement with AI outputs [4, 7]. Crucially, given the risks of skills atrophy, this literacy must also include metacognitive awareness of cognitive offloading: students should learn not just how to use AI tools but how to monitor whether that use is building or eroding their own competences. Educators must design learning environments in which AI use is structured to ensure that the offloaded processing contributes to learning, in the terms established by Risko and Gilbert [16], rather than short-circuiting it.

Given AI's potential to displace authentic human dialogue and reconfigure agency in dehumanising directions, Freirean approaches must deliberately protect dialogic relationships. Classroom communities must remain spaces of horizontal, humanising encounter, with teacher-facilitated discussion as the primary site of meaning-making and conscientisação. AI supplements rather than supplants the irreducibly relational character of critical pedagogy.

AI integration should support problem-posing pedagogy rather than defaulting to banking-model knowledge delivery. This requires framing AI as a tool for generating questions and hypotheses rather than providing final answers, and positioning AI outputs as provisional texts to be read critically. Instructional design must ensure that working with AI actively engages rather than bypasses the cognitive processes that develop competence, guarding against the deskilling dynamic that Gerlich [19] and Sparrow et al. [17] have empirically documented.

Moveover, given AI's tendency to reproduce existing biases and inequalities, Freirean approaches must actively identify and counteract these tendencies—auditing AI-generated content for bias, ensuring equitable access, involving diverse communities in implementation decisions, and monitoring AI's differential impacts on marginalised students. The goal is not merely to mitigate harm but to actively advance educational equity, consistent with Freire's commitment to the preferential option for the oppressed.

Finally, Freire's insistence on democratic and participatory education requires transparency about how AI systems work and genuine community participation in decisions about AI adoption and use. This challenges the technocratic model of AI implementation, resisting proprietary black-box systems and corporate dominance and insisting on democratic control of educational AI [2, 8]. The decisions about which AI tools are adopted, how they are used, and whose interests they serve must not be delegated to technology companies but made through processes of democratic deliberation that include educators, students, families, and communities.

Conclusion: AI as Contested Terrain

What would Paulo Freire think of generative AI? This analysis suggests he would view it as contested terrain: a site of struggle where competing visions of education, human development, and social justice collide. The question is not whether to use AI in education, but how, for what purposes, and in whose interests.

Freire would recognise that GenAI can easily reproduce and intensify the banking model of education, delivering packaged knowledge to passive students whilst masking power relations through claims of neutrality and efficiency. He would expose how corporate control, algorithmic bias, and the displacement of dialogue threaten to deepen educational oppression. The risks of cognitive offloading and skills atrophy add a further Freirean dimension that goes beyond earlier analyses: AI-enabled deskilling is not a neutral technical phenomenon but an expression of the power relations governing cognitive production and reproduction. The Google effect, extended mind research, and emerging evidence on AI's impacts

on critical thinking together suggest that without deliberate pedagogical counter-measures, widespread AI use in education risks producing what Freire would recognise as a new and particularly insidious form of the banking model—one in which students are not only positioned as passive recipients of deposited knowledge but are systematically stripped of the cognitive capacities that would enable them to resist that position.

The agentic turn amplifies these risks further, displacing skilled cognitive work whilst foreclosing the reflective understanding essential to praxis. When AI systems execute multi-step reasoning processes autonomously and opaquely, the conditions for genuine conscientisação—the capacity to perceive, reflect upon, and act against oppressive structures—are eroded at their root.

At the same time, Freire's pedagogy was never about rejecting tools but about transforming their use in service of liberation. AI can support problem-posing inquiry, reveal algorithmic bias as a site of conscientisação, extend dialogic possibilities, and enable praxis—provided that its integration is designed to preserve and develop learner agency and competence rather than displace them. The emancipatory possibilities of AI are real, but they require what Freire would call intentional praxis: ongoing critical engagement that refuses to accept technology on its own terms and insists on subjecting it to the criterion of humanisation.

Freire would call for an approach to AI in education that centres humanisation and social justice, develops critical AI literacy inclusive of metacognitive awareness of deskilling risks, preserves authentic dialogic relationships, designs for problem-posing rather than banking-model delivery, actively addresses bias and pursues equity, integrates AI into cycles of genuine reflection and action, and maintains democratic control over technological decision-making. These are not optional refinements to AI integration but its necessary preconditions if that integration is to serve rather than subvert the purposes of genuinely humanising education.

As educators navigate the AI epoch, Freire's pedagogy offers this essential guidance: education is never neutral, technology serves human purposes, liberation requires critical consciousness and collective action, and authentic transformation emerges from dialogue, praxis, and the ongoing struggle for humanisation. In the age of GenAI, these principles remain as urgent and indispensable as ever. The emergence of generative AI presents educators with a choice that Freire would find deeply familiar: will we use this technology to integrate students more efficiently into existing oppressive structures, or will we transform it into a tool for critical consciousness and collective liberation? The answer depends not on the technology itself but on the pedagogical, political, and ethical commitments that guide its implementation. Freire would remind us that this choice is not made once and for all but must be enacted continuously through praxis—the ongoing, vigilant unity of critical reflection and transformative action.

REFERENCES

[1] Freire, P. (2018). Pedagogy of the Oppressed (50th Anniversary Ed.). Bloomsbury Academic.
[2] Anderson, et al. (2025). AI in Higher Education: The 2024 Survey Results. Journal of Educational Technology.
[3] Jensen, et al. (2025). Disrupting the Classroom: The Rise of LLMs. Global Education Review.
[4] Alm, A., & Watanabe, Y. (2023). Integrating ChatGPT in language education: A Freirean perspective. International Journal of TESOL & Education, 3(2), 85–98.
[5] Xiaoyu, et al. (2025). Generative AI and the Future of Lesson Planning. International Journal of Learning.
[6] Freire, P. (1970). Pedagogy of the Oppressed. Herder and Herder.
[7] Warr, M. (2024). Blending generative AI, critical pedagogy, and teacher education to expose and challenge automated inequality. Research in Higher Education Teaching and Training, 2(1), 1–18.
[8] Mayo, P. (2020). Critical pedagogy in difficult times. In S. Macrine (Ed.), Critical Pedagogy in Uncertain Times (pp. 39–57). Palgrave Macmillan.
[9] Darder, A., Mayo, P., & Paraskeva, J. (2017). The internationalisation of critical pedagogy: An introduction. In International Critical Pedagogy Reader (pp. 1–24). Routledge.
[10] Kahn, R., & Kellner, D. (2007). Paulo Freire and Ivan Illich: Technology, politics and the reconstruction of education. Policy Futures in Education, 5(4), 431–448
[11] Kildė, L. (2025). Humanising digital language education 5.0 through a Freirean lens. Acta Paedagogica Vilnensia, 54, 105–120.
[12] Lynch, et al. (2025). The Efficiency Trap: GenAI and the Erosion of Reasoning. Critical Studies in Education.
[13] Chan, R. (2025). Critical education, generative artificial intelligence and the tyranny of freedom. Technology, Pedagogy and Education, 34(1), 1–15.

[14] Thompson, J., Martinez, S., & Chen, L. (2025). Algorithmically manufactured minds: Generative and agentic AI in a time of post-truth. Open Praxis, 17(2), 145–162.
[15] Freire, P., & Macedo, D. (1987). Literacy: Reading the word and the world. Bergin & Garvey.
[16] Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7–19.
[17] 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.
[18] Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688.
[19] Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies, 15(1), 6.
[20] Zuboff, S. (2019). The Age of Surveillance Capitalism. PublicAffairs.
[21] Iheduru-Anderson, et al. (2025). Freirean Principles in Clinical Nursing Education. Nursing Outlook.