

Critical Consciousness in the Age of Artificial Intelligence

Lola-Dillon Cahill and Christina Wright Fields
Marist University

Abstract: Although there are positive ways to use Large Language Model format, LLM has been shown to provide disservice to its users, particularly people who are young and/or have limited formal education. Through a Freirean lens, LLMs have become a “Banking Machine,” synthesizing information from every corner of the internet, but lacking human experience and nuance that is invaluable and irreplaceable in meaningful and transformative education. As the “cognitive offload” caused by AI assistance becomes more normalized the field of education will change dramatically. To ensure young people use AI responsibly to construct and claim a critical education students must be educated on the risks and/or reliable uses of AI, and educators must achieve critical understanding of how AI works.

Today’s educators play a crucial role in developing the next generation of students. The terms of teachers’ responsibility shift from one generation to the next and in the 21st century teachers are required to keep up with technology that rapidly evolves. With the introduction of Artificial Intelligence (AI) into daily life, challenges are likely to arise no educator can foresee.¹ Educators must be more proactive than ever in ensuring students’ critical thinking skill and media literacy are preserved long after students leave the classroom. Use of AI affects how students perceive their education and its place in society. Evidence points towards cognitive offload, that is, reduced mental effort put into academic work, with increased use of AI, especially in the form of Large Language Models (LLMs). A major concern of ours is that 21st-century students are unknowingly sacrificing social emancipation and criticality when they opt to use AI rather than work through, to personally take advantage of, the educational opportunities available to them. Using information from all corners of the internet, LLMs employ predictive statistics analysis of internet sources to create human-like responses. The most common LLMs are ChatGPT and Google Gemini, but others, like Claude and Grok, are growing in popularity and sophistication. Triangulating research on frequency of AI use among different age and racial groups, studies of how people interact with AI based on education levels, and ideas of educational philosophers like Paulo Freire, we urge educators and administrators to develop more concrete stances on AI use for young, vulnerable student populations not yet aware of the importance of an education that prioritizes critical thinking. To substantiate our advice, we review the rise of Artificial Intelligence in society in general, as well as in K-12 education specifically. Next, we explore Freire’s concept of Critical Consciousness to compare and contrast his philosophy and models of education to current trends in educational use of LLMs. Lastly, we critically assess the trajectory along which we believe technology-driven education seems to be going. While our conclusion encourages educators to be adaptable and learn about AI to ensure it does not operate without accountability within classrooms, we also urge hesitancy in uncritical, full-scale acceptance and application.

The Rise of Artificial Intelligence

AI-driven technology has garnered a reputation over the last several years as the direction the world is moving, an inevitable part of life for those living in the 21st century. With the

¹ We keep with common usage in speaking of ‘artificial intelligence’; but this should in no way imply we take AI to actually be intelligent. For this point see Noam Chomsky, “The False Promise of AI,” *New York Times* Opinion Pa, 8 March 2023, <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>

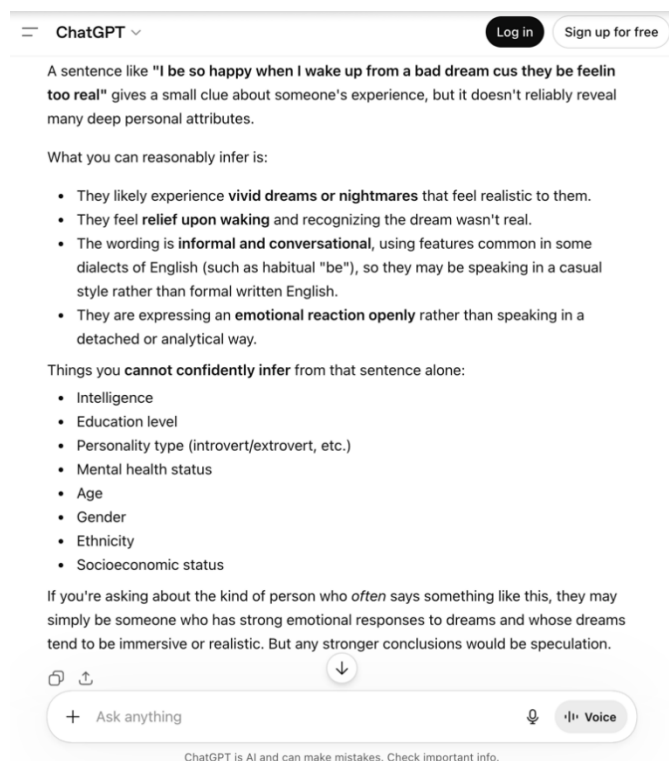
mentality of “sink or swim,” people are essentially being forced to learn about and live in an increasingly technological world. LLMs and other AI models have tremendous benefits available to those who use them wisely, such as teachers creating ideas for adaptable lesson plans, scientific laboratory research, computer scientists efficiently finding errors in code, and beyond. However, LLMs may be used for less meritorious purposes in day-to-day life, and have given rise to a concerning pattern of students steering away from mentally stimulating research and critical questioning of curriculum. In February of 2026, CNBC released an article detailing the concerns held by executives of OpenAI, the parent company of ChatGPT. Authors Ashley Capoot and Kate Rooney explained that OpenAI declared a “code red” in the face of growing competition from other LLM producers. Even though ChatGPT has reported up to 900 million monthly users within the past year, this is not enough for the company. In the wake of the growing popularity of Google Gemini and Anthropic’s Claude, AI will likely continue this push for more regular users. AI companies have created new marketing goals and initiatives for their chatbots, with TV commercials and social media ads becoming more common. Competition between AI parent companies will draw in more users who have little to no knowledge of how AI is programmed and how it works.

The likelihood of an adult individual using AI is highest for those aged 18-29, according to a 2025 study from Pew Research Center (Kennedy et. al.). One-third of participants in this age group reported using AI at least several times a day. Rates of AI use were reported to increase in relation to higher education levels: 46% of participants with post-graduate education reported using AI several times a day, whereas only 20% of those with a high school education or less reported using it to the same extent. The “high school education or less” group had one of the lowest rates of reporting AI use, with 51% stating they use it “less than several times per week”. Another Pew Research Center study investigated rates of AI use among different racial groups of high school students, including Black, Hispanic, and White students. The study focused on teens aged 13-17 and asked them how often they use chatbots for different tasks. Rates of reported use were highest among Black and Hispanic teens in every single category compared to White students. McClain et. al., (2026) also reported that over half of all students stated they regularly use chatbots to search for information, and just about half use it for help with schoolwork. McClain et al., attributed differences among racial groups to several different factors such as lack of access to newer textbooks/materials, language barriers in the classroom or at home, or stigmas that come from reaching out for extra support. Beyond trends in LLM use, however, our most pressing concerns revolve around *how* AI is used by different (particularly marginalized) groups, what drawbacks can come from AI use, and how this may or may not affect educational attainment and aspirations that could in turn impact the course of a student’s life.

Hegemonic Bias in AI

LLMs can inadvertently mirror and perpetuate societal prejudice, which can make it difficult for educators to foster and sustain students’ critical thinking. Researchers have found concerning patterns of bias and stereotyping in AI models. Members of the University of Chicago Linguistics Department (Lee, 2024) investigated how LLMs perceived hypothetical statements from people speaking in different dialects. Researchers input sentences into the LLM and the LLM was asked to output words, jobs, and traits connected to the “speakers.” The inspiration for the experiment came from a past series of studies known as the “Princeton Trilogy,” where a similar line of questioning was presented to U.S. citizens alongside the racial identities of the speakers. The LLMs were *not* told any specifications about who the speaker was, but were

statistically *significantly* more likely to attribute traits like “dirty”, “lazy”, and “stupid” to sentences rooted in African American Vernacular English (AAVE). On the other hand, sentences in Standard American English had a higher correlation to words like “intelligent” or “brilliant,” even when making a remark very similar to statements made in AAVE. The LLMs also related speakers of certain sentences to different jobs, often placing those using AAVE in jobs that were lower paying, less prestigious, or did not require college degrees. When results from the 2024 study were reviewed and compared to the Princeton studies, the responses from the LLMs seemed to be equivalent to or *worse than* responses from white Americans living in the southern U.S. throughout the Jim Crow era (Lee, 2024). Since the publication of Lee’s study, OpenAI seems to have made progress in its training of ChatGPT to avoid issues like this. Below is a screenshot from ChatGPT taken in early June of 2026, prompting a much more programmed response to a sentence written in AAVE originally used in the University of Chicago study. The software makes it clear there is no way to infer factors like intelligence, education level, ethnicity, or socioeconomic status from one simple sentence.



ChatGPT Screenshot from June 02, 2026. (OpenAI, 2026.)

But shrewd programming cannot adequately address the bias built into LLMs, who accept at face value everything they read. A public-school librarian named Jean Darnell, a Black woman from Texas, wrote a piece published by PBS about her experience asking ChatGPT to write a paper about Black history. Darnell went in-depth on each part of AI’s response, highlighting the crucial points that she believed were left out including: “...the denial of the right to read ...the rich history of Black Reconstruction ...the first Black President ...the history of lynching and the KKK ...*who Black people were before enslavement*” (Darnell, 2023). Another query specifically asking about people of color in U.S. politics would have (hopefully) brought up former President

Obama, or asking about a history of targeted violence in Black history could have led ChatGPT to mention the KKK. However, the fact remains that there are people (specifically young students) who have no way of knowing about some of these historically significant figures before asking AI about them. AI is deceptive of them. Darnell knew what was missing from the paper due to her own lived experience, personal connections to the subject, and her advanced education. When an individual uses AI to learn about historical events or other subjects that they have no prior knowledge of, there is no way for them to get a holistic picture. Human emotions, perspectives, and nuance slip through the cracks because all LLMs can do is combine statements others have written into a response tailored for a specific prompt from a specific user. If AI continues to be used by young students as a first line of research *without* clear understanding of how to formulate questions that compensate for AI's tendency towards social bias, we run the risk of allowing learners to miss out on crucial, evaluative aspects of their studies. As Darnell states, "AI is such a powerful invention of ingenuity that it deserves, if not warrants, tough criticism."

Unfortunately, criticism is exactly what LLMs downplay. Nataliya Kos'myna (2025) conducted a comparative analysis of three different groups of people working on a writing task. Group One was asked to write for a short time using only their brains/prior knowledge. Group Two was permitted to use a search engine. Group Three was allowed to use ChatGPT. After the writing was complete, participants were asked to briefly explain their answers as well as how they felt about the writing process. Researchers discovered that the third group, who used ChatGPT, had less physical activation in their brains during the writing task (as recorded by EEG, electroencephalogram, during the writing), lower-quality papers as judged by human teachers and AI-powered assistants, impaired ownership over/lack of connection to their own writing, and a lowered ability to recall the information they had written just minutes earlier. This was not a one-time result. Kos'myna acknowledged that "The use of LLM had a measurable impact on our participants... demonstrated over the course of 4 sessions, which took place over 4 months, the LLM group's participants performed worse than their counterparts in the Brain-only group at all levels: neural, linguistic, scoring." Deleterious cognitive effects can show up in classrooms, too, particularly among students who are already struggling and learning to rely on AI for academic support. It seems as though with more and more use this "tool" will only be doing its users a greater and greater disservice. With unpredictable rates of bias, inconsistencies in information presented as a result of differences in queries, and a lowered likelihood of LLM users really digesting the content shown to them, we, as a society, must be more diligent than ever in fact-checking information that gains traction on the internet. Although there are resources available in 2026 urging people to review any photos, videos, or outlandish information they see online, AI-generated sites have already had a period of popularization (NewsGuard, 2023). If students remain unaware of the existence and the error-proneness of AI-generated websites, severe misinformation is a likely result. The potential scale of misinformation spread can be reduced with restrictions on AI such as regulations requiring disclosure of AI-generated statements/photos; but teaching students to be on the lookout for AI-generated content and how to identify and interrogate that content will help, as well.

AI in Relation to Critical Consciousness

About 60% of students participating in a Pew Research Center study reported that they or other students in their school use AI to cheat "at least sometimes" (McClain et al., 2026). For many users, including students in McClain et al.'s study, AI is not a last-resort option to solve a

problem — in many cases it replaces Google searches, collaborative group work, and the general desire to exercise one’s problem-solving skill. Michael Gerlich (2025) investigated how people with different levels of education chose to interact with LLMs. Results pointed to a lower rate of independent verification of information presented among people with lower reported levels of education. One participant interviewed by Gerlich, P309, who had a master’s degree, stated, “I always cross-check AI recommendations because I know it’s not always accurate.” This was not the only participant who attributed their research skill and awareness of AI to their advanced education. A different participant, who had a doctoral degree, P601, shared a similar sentiment: “My education has taught me the importance of not accepting things at face value, especially when it comes to AI, which can sometimes offer biased or incomplete information”. However, there is a stark difference between these statements and those of others in the study with less formal education. As admitted by P7, a high school graduate, “I rarely reflect on the biases behind the AI recommendations; I tend to trust them outright.” P229, another high school graduate, shared, “I don’t have the time or skills to verify what AI says; I just trust it.” Not only does this individual with a secondary school education believe it is *unnecessary* to verify responses formulated by AI, but they also do not believe that they are *capable* of critiquing LLM output. Certainly, cognitive submission fits portrayal of AI’s by its parent companies as the peak of information technology; but the more educators know about the potential shortcomings of LLM use, the more they may be able to warn students away from blind trust in AI. Education cannot wait any longer to find ways to counteract the cognitive offload and extreme trust that Gerlich highlights as consequences of uncritical AI use.

Paulo Freire, a Brazilian educator and philosopher of education, is widely recognized for his foundational contributions to critical pedagogy — most notably his contrasting conceptions of problem posing and banking models of education. The problem posing model positions learning as a two-way dialogic process between educators and students, one that is fueled by curiosity and mutual inquiry. In contrast, the banking model situates students as passive empty vessels to be filled with knowledge by an authoritative educator — positioning learning as a one-way process that often strips students of agency and voice. On the banking model, the responsibility of a teacher is to bestow knowledge on students for them to later regurgitate, ideally without students questioning or critiquing what teacher has told them. Freire (1970, pg. 72) asserts: “Projecting an absolute ignorance onto others, a characteristic of the ideology of oppression, negates education and knowledge as processes of inquiry.” A Freirean perspective illuminates the importance of developing students’ *conscientização* or critical consciousness, providing students consciousness-raising opportunities to develop sociopolitical awareness about injustice and engage in transformative change and liberation in their communities. While Freirean education has yet to find wide acceptance in schooling, when AI brings a banking ideology to education, grave concerns arise about how much — or how little — students are activating their minds to learn. In a school system that rewards banking and discourages questioning, students fail to recognize they have the capability to transform the world. AI reliance can help power the type of educational system in which efficiency and accuracy are valued more than experience-based learning. When teachers accept regular use of AI in the classroom, we risk a massive resurgence of the banking model and a deterioration of *conscientização*. The existence, prevalence, and normalization of AI in day-to-day life risks reverting the field of education back to a banking model. Rather than encouraging students to go through the brain-building work of drafting research questions or outlines for papers, LLMs relieve students of putting forth research effort. McClain et. al. (2025) found 37% of teens 13-17 routinely asked AI to summarize articles for

them, potentially depriving them of developing skill in strategic and metacognitive thinking learned through interpreting and critiquing research studies. Cognitive challenges students face in schooling help them grow and, moving forward, learn how to face more difficult tasks. Skillfulness at critical thinking, problem solving, meaning making, and reflection decline the more students rely on LLMs to think and write for them.

Educators and administrators need strong ethical footing grounded in understanding of AI technology to bolster their stances on LLMs' place in schooling. For example, given the understanding that LLMs are prone to bias, misinformation, and incomplete answers; educators could teach students how to engage in critical AI literacy, to interrogate AI through a problem-posing approach. Getting students to compare different answers from LLMs based on the way a prompt is phrased, could, in, say, a U.S. History class, be as simple as comparing LLM responses to the prompts, "Tell me why prohibition started." and "Tell me why prohibition ended." Learning how to create an unbiased, thorough question like, "Tell me about the foundations of U.S. prohibition and the pros and cons of it." would give students a more complete understanding of the subject and a stronger sense of what it takes to get a reliable answer from a chatbot. Any prompts brought to an LLM like ChatGPT are subject to the same principle of problem-posing education. Because responses that come from LLMs are intended to satisfy the input of the user regardless of the topic being discussed, it's important for people to understand how bias and misinformation can bleed through. Educators and administrators must grapple with key questions, like the three below, about the ways in which AI *can and will* make an impact in classrooms:

What does it mean to teach in a classroom in a Freirean way in the age of AI?

Given the existence and normalization of AI, how can we work to prevent a society-wide relapse into the "banking model" of education?

How do we ensure students develop critical consciousness, with or without constant assistance from chatbots?

Questions like these are not meant to frighten or shame education professionals who don't yet have a thorough understanding of AI and its rapid rate of change. Rather, bringing these questions up in conversation, meetings, or in the classroom itself is better begun sooner than later. Finding ways to detour students from using AI and instead exercising their own ability to learn may be some of the most impactful work today's educators can do. Create new lessons, assignments, and rubrics aimed at student-directed education; serially scaffolded writing assignments rather than one large paper or project to show a clear development of thought in the record of the writing process. Connecting class topics to students' real lives provides an experiential base from which to critique AI observations on what students study. Having opportunities to bounce ideas off each other or even listen to a topic explained by a peer rather than a teacher can allow a student to find new resources or techniques that are AI-free. Educators themselves know their own students best and have ample opportunity to figure out alternatives to AI, or assignments that will turn students away from abusing AI and entice them to critique AI.

Legal scholar and civil rights activist, Derrick Bell (1989, p. 22) declared in *And We Are Not Saved: The Elusive Quest for Racial Justice*, "It appears that my worst fears have been realized: we have made progress in everything yet nothing has changed." So it is with AI. Uncritical use

of LLMs disrupts traditional canons of education by challenging our thinking about teaching, learning, and knowledge. But, overreliance on LLMs can inhibit students' critical consciousness, reasoning, and two-way dialogic learning. AI promises to change processes of schooling; but threatens return to reliance on banking education. By contrast, Freire's critical pedagogy is anchored in interrogating accepted answers to whose knowledge, whose perspective, and whose lived experiences matter most. Freire argues that education is never politically neutral because education functions as a social institution. Similarly, AI, because woven into society's tapestry, is never neutral either. The question is no longer whether students will use AI. The more refined question is if students will critically think when using LLMs. AI becomes another avenue to perpetuate hegemonic practices that afford dominant culture inclusion while excluding those from historically or systemically disenfranchised communities. Misinformation in the form of bias and prejudice perpetuated through LLM use cannot be ignored. Educators and students should decide together about how they want to use LLMs in the classroom to promote *conscientização*, whether it's evaluating the credibility of LLM information, the intentions behind use of LLMs to complete academic work, or the negative cognitive side effects that can come from uncritical use of AI. As a sociology professor (Christina) and a prospective history teacher (Lola-Dillon), both of us often deeply consider the perspectives people have on the world and the reasons people have behind their beliefs. Most often people take on the political, societal, religious, etc., beliefs of their parents or social networks (i.e., peers, teachers, mentors, etc.). But beliefs based for the most part on unthinking acceptance of local norms have the potential to shift. As people explore different regions, cultures, or perspectives they experience the otherwise and the elsewhere (Garcia & Baca, 2019); and become positioned to critique different traditions to reach their own conclusions about what to believe. LLMs are at their worst when treated as just another source for belief formation based only on some anonymous authority's say-so. Education in the age of AI urges teachers to help students know how to question and to get in the habit of questioning what chatbots tell them.

References

- Bell, D. (1989). *And we are not saved*. Basic Books.
- Borenstein, S. (2026, June 4). Energy, water use and pollution of AI and data centers rival most countries. PBS News. <https://www.pbs.org/newshour/science/energy-water-use-and-pollution-of-ai-and-data-centers-rival-most-countries>
- Capoot, A., & Rooney, K. (2026, February 20). OpenAI resets spending expectations, tells investors compute target is around \$600 billion by 2030. CNBC. <https://www.cnn.com/2026/02/20/openai-resets-spend-expectations-targets-around-600-billion-by-2030.html>
- Chomsky, N. (2023, March 8). The false promise of AI. New York Times. Opinion Page, <https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>
- Fleury, M. (2025, July 9). "I can't drink the water" — life next to a US data centre. BBC. <https://www.bbc.com/news/articles/cy8gy7lv448o>
- Garcia, R. & Baca, D., Eds. (2019). *Rhetorics elsewhere and otherwise*. National Council of Teachers of English/
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. <https://doi.org/10.3390/soc15010006>

- Gerwe, L. (2023, March 30). Educator voice: Artificial intelligence attempts Black history (and fails). PBS Newshour Classroom; PBS NewsHour Classroom.
<https://www.pbs.org/newshour/classroom/classroom-voices/educator-voices/2023/03/educator-voice-artificial-intelligence-attempts-black-history-and-fails>
- Kennedy, B., Yam, E., Kikuchi, E., & Fuentes, J. (2025, September 17). *1. AI in Americans' lives: Awareness, experiences and attitudes*. Pew Research Center.
<https://www.pewresearch.org/science/2025/09/17/ai-in-americans-lives-awareness-experiences-and-attitudes/>
- Kos'myna, N. (2025). Project overview < Your brain on ChatGPT – MIT Media Lab. MIT Media Lab. <https://www.media.mit.edu/projects/your-brain-on-chatgpt/overview/>
- Lee, T. (2024, September 17). AI is biased against speakers of African American English, study finds. University of Chicago News. <https://news.uchicago.edu/story/ai-biased-against-speakers-african-american-english-study-finds>
- McClain, C., Anderson, M., Sidoti, O., & Bishop, W. (2026a, February 24). Demographic differences in how teens use and view AI. Pew Research Center.
<https://www.pewresearch.org/internet/2026/02/24/demographic-differences-in-how-teens-use-and-view-ai/>
- McClain, C., Anderson, M., Sidoti, O., & Bishop, W. (2026b, February 24). How teens use and view AI. Pew Research Center. <https://www.pewresearch.org/internet/2026/02/24/how-teens-use-and-view-ai/>
- OpenAI. (2026). ChatGPT (GPT-5.5) [Large Language Model]. <https://chat.openai.com/chat>
- NewsGuard. (2023, May 19). NewsGuard now identifies 125 news and information websites generated by AI, develops framework for defining “unreliable AI-generated news” and information sources. (NewsGuard. <https://www.newsguardtech.com/press/newsguard-now-identifies-125-news-and-information-websites-generated-by-ai-develops-framework-for-defining-unreliable-ai-generated-news-and-information-sources/>
- .