Against Artificial Education: Towards an Ethical Framework for Generative Artificial Intelligence (AI) Use in Education

Andrew Swindell
Asian University for Women, https://orcid.org/0000-0001-9026-690X
Luke Greeley
Rutgers Business School, https://orcid.org/0000-0002-5544-1589
Antony Farag
Rutgers Graduate School of Education, https://orcid.org/0000-0002-0771-2268
Bailey Verdone
Rutgers Graduate School of Education, https://orcid.org/0009-0004-4239-3217

Abstract
The arrival of generative artificial intelligence (AI) is fundamentally different from prior technologies used in educational settings. Educators and researchers of online, blended, and in-person learning are still coming to grips with how to employ current AI technologies in the learning experience, let alone understanding the potential consequences that future and unknown developments in AI will produce. Despite potential risks, AI may revolutionize previous models of teaching and learning, and perhaps create opportunities to realize progressive educational goals. Given the longstanding tradition of using philosophy to examine questions surrounding ethics, ontology, technology, and education, the purpose of this critical reflection paper is to draw from prominent philosophers across these disciplines to address the question: how can AI be employed in future educational contexts in a humanizing and ethical manner? Drawing from the work of Gunther Anders, Michel Foucault, Paolo Freire, Benjamin Bloom, and Hannah Arendt, we propose a framework for assessing the use and ethics of AI in modern education contexts regarding human versus AI generated textual and multimodal content, and the broader political, social, and cultural implications. We conclude with applied examples of the framework and implications for future research and practice.

Keywords: generative artificial intelligence (AI), Bloom's taxonomy, Paolo Freire, philosophy of technology, Gunther Anders

Swindell, A., Greeley, L., Farag, A., & Verdone, B. (2024). Against Artificial Education: Towards an Ethical Framework for Generative Artificial Intelligence (AI) Use in Education. *Online Learning*, *Vol 28*(2), (7-27). DOI: 10.24059/olj.v28i2.4438

Introduction

The arrival of generative[1] artificial intelligence (AI) is fundamentally different from prior technologies used in educational settings. Educators and researchers of online, blended, and in-person learning are still coming to grips with how to employ current AI technologies in the learning experience, let alone understanding the potential consequences that future and unknown developments in AI will produce (Bozkurt, 2024; Şenocak et al., 2024). Despite potential risks, AI may revolutionize previous models of teaching and learning and perhaps create opportunities to realize progressive educational goals. When ChatGPT was released to the general public in November 2022, these dangers and possibilities were immediately made salient as students across the world could instantly use AI to create content for class assignments and educators scrambled to understand if and how AI could fit in their coursework. While some research has explored the ethical implications and imagined futures of AI in education (Akgun & Greenhow, 2022; Nguyen et al., 2023), there is limited consensus or direction for classroom use of AI specifically. Given the longstanding tradition of philosophy to examine questions surrounding ethics, ontology, technology, and education (An & Oliver, 2021), the purpose of this critical reflection paper is to draw from prominent philosophers of these disciplines to address the question: how can AI be employed in future educational contexts in a humanizing and ethical manner? While this is not an empirical study guided by explicit research questions, the overarching problem this paper addresses is how to ethically and practically incorporate generative AI into modern and imagined education systems.

We first review literature from theorists such as Anders and Foucault to bring awareness to the functions of technologies in modern society and their capacity to change the human condition and shape sociopolitical relations. Next, we draw from educational philosophy, where arguments have been made using Paulo Freire that ethical education at its core promotes humanization, conscientization, and radical engagement with the world, while technological adoption in educational practice is not inherently harmful, so long as the technology promotes these humanizing ends (Boyd, 2016; Kahn & Kellner, 2004; Farag et al., 2021). We then synthesize the work of Bloom's (1956) taxonomy and Arendt's (1958) hierarchy of labor, work, and action to construct a lens through which educators and classroom teachers can understand both the types and goals of learning that are necessary given the possible futures of AI in education. With this theoretical grounding, we propose a framework for assessing the use and ethics of AI in modern education contexts regarding human versus AI generated textual and multimodal content, and the broader political, social, and cultural implications. The framework is general in that it is not limited to a specific modality of education delivery, thus it can be applied to online, blended, and in-person settings. We conclude with applied examples of the framework and implications for future research and practice.

[1] While general or generic artificial intelligence are more broadly defined, the subject of this analysis is limited to generative artificial intelligence (AI) systems which are "developed to analyze complex patterns and structures in human language [and] are primarily designed to comprehend and replicate it. Generative AI, when expertly trained, exemplifies a potent tool capable of learning, unlearning, and relearning which makes it a continually adapting [and] evolving entity" (Bozkurt, 2024, p. 2).

The Philosophy of Anders on Technology and Society

After decades of neglect by academics, especially in the English-speaking world, many scholars across disciplines are beginning to study Gunther Anders. Anders, a German-Jewish intellectual, wrote prolifically on "the lasting effect of the industrialism, the psychological and political implications of mass media, the legacy of the Holocaust and the intervention of the atomic bomb for history and memory, the Vietnam War, ecological devastation, and new forms of internationalism and collective political action" (Dawsey, 2014, p. 11). A connecting thread between Anders's interdisciplinary milieu of interests was how technologies were not just reshaping the social and psychological behaviors of humans, but how technologies were actually modifying what it means to be human, and perhaps moving into a space where humanity is irrelevant. The opening to his deepest work on the subject, *The Obsolescence of Man, Volume II*, begins

> It is not enough to change the world. That is all we have ever done. That happens even without us. We also have to interpret this change. And precisely in order to change it. So that the world will not go on changing without us. And so that it is not changed in the end into a world without us. (Anders, 1980/2015, p. 1)

Anders's formulation for interpretation and change rests on a process of critical reflection, in order to assess the current place of humans within the interconnected systems of technologies they have developed, and engagement with imagination in the act of exaggeration, in order to reveal the indeterminate and identify spaces for potential freedom.

As a starting place, Anders argues that we must understand "Artificiality is the nature of man and his essence is instability" (Anders, 1936/2009, p. 146). Humans are able, through abstract thinking, to transform the world and "build over it according to a thousand historical variants" creating a paradoxical situation in which "agency" brought to us through technology simultaneously creates a prison of contingency, whereby given technologies shape and direct our means of production, our communication, and our social structures but cannot be escaped. This paradox was specifically escalated at the point in which humans began "Producing machines, or at least parts of machines, by way of machines" (Anders 1980/2015, p. 6) creating entire systems in which seemingly infinite products can be produced with very little input from humans themselves. "Only at the beginning of these chains of production (as inventors or artisans) and at the end (as users) do men have a place" (Anders, 1980/2015, p. 6). As early as the 1950s Anders formulated what he called the "Promethean discrepancy," which is the gap between what can be produced and what humans can imagine or emotionally grasp in terms of the consequences and implications for society and for the individual being. It is in this space that the method of exaggeration is needed to bring fantasy and anticipation to get a better understanding not of the future potential of technology but its present reality. Catalani (2020, p. 161) makes this point well by summarizing, as an example from Anders's work, a bomb that in pure sight as an object is "understated" but its true meaning can only be grasped by imaging the terrible potential it contains. Anders's explicitly says that when we use our imagination in this way, we increase our "Freedom to Fear," which is a "stirring fear, since it should drive us into the streets instead of under cover" and a "loving fear, not fear of the danger ahead but for the generations to come" (Anders, 1962, p. 498). Anders (1980/2015) argues that humans have adopted as a principle of modernity, "the possible is generally accepted as the compulsory and what can be done as what

must be done"; a principle that confuses technological possibility with moral justification (p. 7). The examples of this confusion are manifest in the creation of weapons that can destroy the world, the overproduction of goods that can destroy the environment, and the development of machines that can do our thinking for us. Consider the Promethean discrepancy found in the widespread adoption of "smart" devices and the correlated social media platforms that operate on them. Such technologies have been rapidly adopted, with little consideration for their consequences, and their use thus far seems to have negative societal impacts on domains such as personal privacy, the development of youth, and psychological health, not to mention the resources that are being used to produce them and their impact on the biosphere. In the following section, the concept of Promethean discrepancy is used to illustrate the vast and tremendous unknown that society must confront regarding AI, and to interrogate its already -realized and imagined impact on global education systems.

Foucault, Power, and Generative AI as an Educational Technology

Given the uncertainty surrounding how AI and society will impact each other, we use a Foucauldian analysis to understand technologies as systems of power, and it allows us to deconstruct the role technology plays in modern epistemologies (Giantini, 2023; Swartz & McElroy, 2023; Megahed et. al., 2023). The most striking example of producing machines by way of machines as described by Anders is through modern AI. While artificial intelligence has impacted society in various forms (Rahm, 2023) its new manifestation, generative AI, through its emulation of human interaction, has penetrated through a metaphysical barrier between machine and man. According to Baido and Ansah, "Generative modeling artificial intelligence (AI) is an unsupervised or partially supervised machine learning framework, which generates manmade relics via the use of statistics, probabilities etc." (2023, p. 53). However, the application of this machine learning in the form of GPT, or Generative Pre-trained Transformer, which can "engage customers in human-like conversation" erodes the boundaries between machine and human developed epistemologies (Baido and Ansah, 2023, p. 53). Rather than users searching the internet and AI providing a group of answers associated, AI generates content acting *as* a human. Generating information in a human-like manner centers the machine as the synthesizer and evaluator of knowledge, rather than the human. With vast amounts of data processed at lightning speed to generate human-like interactions, the user gains tremendous advantages in perceived efficiency but is left exposed to ontological risks where meaningful thought, investigation, and creation are "outsourced" to other actors. Uncritical users of AI, without training, adaptation, and awareness, may be left vulnerable to manipulations by corporations, bad actors, or perhaps the AIs themselves. The core of human-driven epistemology is shaken by the powerful effects AI has on our abilities to synthesize knowledge.

French philosopher Michel Foucault provides a helpful vocabulary for analyzing these shifting power dynamics through his concept "regimes of truth." Lorenzini (2015) analyzes Foucault's concept, stating

> Each society has its regime of truth, and by this expression [he] means: (1) "the types of discourse [society] harbors and causes to function as true"; (2) "the mechanisms and instances which enable one to distinguish true from false statements" and (3) "the way in which each is sanctioned"; (4) "the techniques and procedures which are valorised for obtaining truth"; (5) "the status of those who are charged with saying what counts as

> true" (Foucault 1976, p. 112; 13). Therefore, "truth" is "a system of ordered procedures for the production, regulation, distribution, circulation and functioning of statements"; it is linked "by a circular relation to systems of power which produce it and sustain it, and to effects of power which it induces and which redirect it". (p. 2)

Considering the current arms race amongst trillion-dollar corporations such as Microsoft and Google as they develop respective AI technologies following a decade where algorithm-driven platforms such as Facebook, YouTube, Twitter, and TikTok created an ecosystem of anger, misinformation, and conspiracy thinking, it is easy to imagine the outsized role accelerating AI will have on epistemologies and definitions of truth. These for-profit companies, incentivized to increase user platform "engagement," are seeking ways to more effectively and efficiently manipulate human behavior. Through AI, corporations enhance their abilities to establish regimes of truth leaving human users of these platforms struggling to formulate their own truths.

For educators whose core mission is to disseminate knowledge and assist learners in the comprehension of the world, it is worth considering the rapidly changing information and technological landscape and, more specifically, the competing regimes of truth that AI presents. There have been instances in which similar questions have been asked of competing AIs like OpenAI's ChatGPT and Google's Gemini, which have led to different responses reflecting various data sources and methodologies including learned falsehoods (Marr, 2024; Pacchiardi et al., 2023; Wong, 2023). Further complicating these discrepancies is that the engineers and programmers who design AI systems often cannot explain how a specific response or output was reached given the complexity of the data architecture and decision-making processes of the systems (Dobson, 2023). AIs have been shown to "hallucinate," calling into question what these corporations define as truth and exaggerating the influence of errors. In a study that assessed AI's ability to understand culturally nuanced questions, researchers found that "the evaluated models demonstrated varying performance with controversial topics, those lacking clear scientific consensus and the brain teasers proving more susceptible to GPT hallucinations…their identification necessitates human-centric assessments by interrogators who apply their own cultural perspectives and values" (McIntosh et al., 2023, p. 12). Other limitations include AI's replication of human biases such as racism and sexism (Rountree & Condee, 2021).

These inconsistencies, hallucinations, and biases represent varying regimes of truth and without a human-centric ability to assess the conclusions various GPTs come to, educators and learners will be unable to distinguish between these radically different and competing "truths." The various regimes of truth replicated by AI produce errors that can be normalized in ways that can be harmful to humans. If education is to be a liberating force against oppressive constructs, and AI reinforces those constructs, educators will continually be fighting an uphill battle. Given the influence these corporations have already attained in creating learning management systems of educational institutions (Farag et al., 2021; Ofgang, 2023), these regimes of truth can manipulate educators and learners who are relying on AI. As the primary aim of these competing corporations is increasing profitability by developing and marketing more advanced AI valuing efficiency over accuracy, educators cannot fully depend on the "truths" these corporations provide regardless of how accurate and efficient they advertise themselves to be. Without cautious and critical strategies for use, learners become consumers and data providers to corporations building their ever increasingly powerful AIs.

Freire and AI as Humanizing Education

Regardless of how the corporate and regulatory power dynamics surrounding AI unfold, it is clear that the AI used in educational settings cannot be assumed to always be accurate or have students' or educators' best interests at heart. The debate over ethical use of technology in education is certainly not new, and in educational philosophy, arguments have been made using Paulo Freire as a guide that technological adoption in educational practice is not inherently harmful, so long as the technology promotes conscientization and humanizing ends (Boyd, 2016; Kahn & Kellner, 2007). Here Freire's (1968) method of *conscientization* is especially relevant, where communities engage in critical dialogue, investigation, and action as an educative endeavor, in order to develop a better understanding of the world and engage in the "ontological and historical vocation to be more fully human" (Freire, 1968, p. 37) and thereby provide an effective framework to counter machine generated understandings of the world. Freire argued that conscientization through educational processes and activities is realized when learners (and/or educators) understand not just their place in the world, but how to transform both their reality and the world (Freire, 1965, 1997) . Regarding the role of technology in these processes of conscientization and humanization, Farag et al. (2021) noted that:

> The growing integration of technology into educational settings was viewed by Freire as an area of concern and pedagogical possibility…despite Freire's concerns that increasingly advanced technologies would serve cultural and economic interests of those in power, he viewed technologies as tools which, if used critically and carefully, could benefit the oppressed…Freire identifies the primary goal of humanity as the ontological pursuit of being more fully human, which is done through exercising one's capacities, having one's voice heard, and acting with others to transform the world. (p. 3)

These concepts of conscientization and humanizing education have been used in education philosophy for decades, though research on AI using Freire has only just begun. Freirean analysis has been employed to question the neutrality of AI systems in education (Stapleton et al., 2023) and used to examine how AI can assist students in identifying what matters to themselves in an educational setting (Loftus & Madden, 2020). Freire argued that the technique of dialogue in education amongst students, teachers, and all others involved was vital to pedagogy, and specifically a crucial means to help all parties realize conscientization and better understand their agency in the world. While research on how AI and chatbots in particular is limited on how they can facilitate such meaningful dialogue, Lin (2023) specifically found that:

> ChatGPT's conversational capability enables it to not only play an intermediate role but also act as an independent agent to engage in classroom dialogues with learners and teachers. What is more impressive is that the conversation with active users within a certain classroom context further becomes training data for ChatGPT, which allows it to grow accuracy and sensitivity in conversations by gaining a higher level of context-specific knowledge. (p. 3)

Despite the limited research on the application of Freire to AI in education, it is clear that applying Freire's concept of humanizing education and conscientization to examine the political, economic, social, cultural impacts of AI can help inform how teachers, students, and

administrators can use AI in education in a way that is both humanizing and advances student learning objectives like critical thinking and creativity. AI, at its best, can act as a fellow member in humanization and conscientization, though significant questions remain on how to accomplish this at the classroom level.

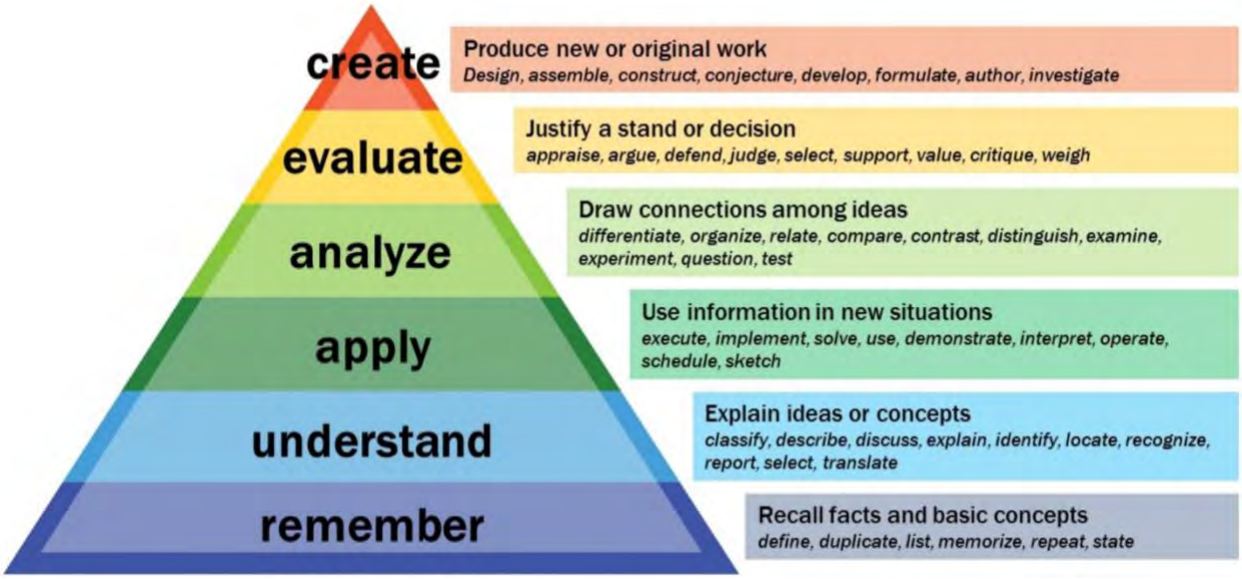
Typologies and Purpose of AI Use in Education

Given this analysis, which demonstrates that educators will need teleological guideposts to employ AI at the classroom level in a humanizing way, we explore how AI fits within Bloom's (1956) taxonomy, a prevalent framework amongst practitioners for thinking about educational activities. We suggest that in the presence of machines, which can perform much of the "labor" of classifying, analyzing, and synthesizing, educators require a new framework that prioritizes the use of technology and adoption of learning toward goals that improve the human condition. For this purpose, we apply Hannah Arendt's (1958) hierarchy of labor, work, and action (terms used by Arendt to classify and value how humans spend their time and energy) to educational contexts to interrogate both the intent and outcomes of education.

Understanding Teaching and Learning with the Revised Bloom's Taxonomy

Bloom's (1956) *Taxonomy of Educational Objectives* and Anderson and Krathwohl's (2001) subsequent *Revised Bloom's Taxonomy* (RBT) have been a foundation in educational preparation programs and classrooms globally for decades. The application of the RBT in research and practice varies significantly, and is understood narrowly in this analysis as a way to classify teaching and learning in the context of AI that is familiar and clear. The cognitive dimension of RBT (Figure 1) depicts a spectrum of cognitive processes in learning that range from the lower tier of remembering and understanding to the higher levels of evaluating and creating. Each of these tiers represents an ascending order of cognitive complexity and its corresponding "action" words that can be used to create lesson plans, educational activities, and assessments in particular. Anderson and Krathwohl (2001) also argued in their development of RBT that there is a knowledge dimension to learning that includes "metacognitive knowledge is knowledge of [one's own] cognition and about oneself in relation to various subject matters . . ." (p. 44)." Accordingly, the cognitive dimension shown in Figure 1 and the metacognitive element of the knowledge dimension illustrate a clear and distinguishable taxonomy and framework that AI can be understood within.

Figure 1
Bloom's Taxonomy
(Armstrong, 2010)



Given the utility that the RBT provides in examining learning processes, research has already been conducted on AI and the RBT. Research on generative AI Large Language Models (LLM), like ChatGPT, have shown these LLMs can complete activities across the RBT, though they perform significantly better at the lower tiers (remember, understand) than the upper (evaluate, create) (Binh Nguyen Thanh et al., 2023; Lourenco et al., 2023). LLMs have also been shown that using keywords from across the spectrum of RBT can challenge "the AI model's capabilities" (Elsayed, 2023, p. 6). Regarding AI usage in educational contexts, Jin et al. (2023) found that LLMs are useful in "supporting metacognitive, cognitive, and behavioral regulation…but not for regulating motivation" (p. 1). More generally, Damaševičius (2023) argues that "the advanced natural language processing capabilities of ChatGPT have led to a shift away from the linear, hierarchical model of Bloom's taxonomy, towards a more dynamic and fluid approach to knowledge acquisition and application" (p. 115) suggesting that AI may unlock new ways of teaching and learning that can go beyond the tiers of RBT. Despite these potential positive use cases, calls have been made to "prioritize ethical AI usage, cultivate AI literacy, and develop frameworks that empower students and educators to safely harness the full potential of these technologies" (Farrelly & Baker, 2023, p. 11). We argue that as a normative framework, RBT has great utility in understanding AI use in education, but that it is not sufficient to interrogate the ethical quandaries that AI presents, particularly surrounding how AI can work towards or against a humanizing education. In other words, given AIs demonstrated ability to operate, at varying levels of competency, across the RBT in educational settings, a new framework is needed that incorporates the impact of AI not just on these tiers of learning, but also centers the human condition, with its various psychological, social, and physical motivations, and the need to act in the world. Because RBT is not prescriptive in a way that relates to the purposes or motivations to which humans apply their learning, we propose supplementing the framework with the work of political philosopher Hannah Arendt. Arendt is specifically valuable in providing a normative directionality for human action, especially in an era of rapid technological change.

The Philosophy of Arendt and Action-oriented Educational Goals

Hannah Arendt's *The Human Condition* (1958/2013) is a careful and thought-provoking exercise in the classification and examination of the most fundamental human activities and provides useful insight into a theory of education that seeks to develop capacities for deliberating in, and acting on, the world. Originally published in 1958 during the Cold War era, when human activities seemed to be leading toward self-destruction, Arendt began the work as a call to critically reflect on the significance and purpose of human activity so that we may "think what we are doing" before it is too late (p. 5).[2] Introducing her book shortly after Sputnik had been launched, Arendt presciently describes a new age where the life-giving qualities of Earth are taken for granted, and where scientific advancements, expressed in new technologies, become so ascendent they think *for* us:

> The trouble concerns the fact that the "truths" of the modern scientific worldview, though they can be demonstrated in mathematical formulas and proved technologically, will no longer lend themselves to normal expression in speech and thought…In this case, it would be as though our brain, which constitutes the physical, material condition of our thoughts, were unable to follow what we do, so that from now on we would indeed need artificial machines to do our thinking and speaking. If it should turn out to be true that knowledge (in the modern sense of know-how) and thought have parted company for good, then we would indeed become the helpless slaves…thoughtless creatures at the mercy of every gadget which is technologically possible, no matter how murderous it is. (p. 3)

Arendt's diagnosis aptly characterizes the new AI age, where the creators of LLMs and similar software cannot explain exactly how their programs are working (known as the "black box" problem, see Dobson, 2023), and where at least 68% of the human population is using a smartphone (Laricchia, 2023) without full comprehension of how they work or where their materials are sourced from. Despite this challenging disconnect, her study of human activities that follows, and how society organizes itself, suggests a space for reclaiming agency.

For Arendt, humankind is defined by its most common activities, which she categorizes into a hierarchy of three groups and labels humans according to activities in which they engage. *Animal laborans* is the label she gives to the most basic and essential human activities: labor and consumption. Arendt calls laboring and consuming activities "metabolic" not only because of their close relation to the life process, but also because laboring and consuming, like the functions of the body, require little thought (p. 98–100). The second label Arendt gives is *homo faber*, which means man who "works upon" or who "makes something" (p. 136). *Homo faber* is defined by the activity of "work." Work is the act of creating and constructing material objects of a lasting durability (p. 143). According to Arendt, it is work that allows humans to build a world that extends across generations and improve the conditions under which we labor and consume.

It is the third activity of "action" that is Arendt's highest classification of human endeavor and representative of the title *zoon politikon.* Arendt says that humans are imbued with

[2] It is perhaps not coincidental Arendt's foundational concerns pertaining to technology share similarities to Anders: they were married from 1929–937 and maintained a lifelong correspondence.

the unique freedom to "set into motion," through deeds and speech, actions that have never been taken before and whose influence is impossible to predict (p. 189). Essential to action is the presence of others—a public sphere, where humans debate, discuss, plan, create, perform, and engage in activities with the purpose of changing the perspective of others and influencing their future behaviors. Arendt states that action "corresponds with the human condition of plurality" and action itself is what constitutes "*the* condition…of all political life" (p. 7). Action is not drawn from the necessity of *animal laborans* or the desire for mastery of *homo faber*, which each seem to follow observable patterns of "standard" human behavior. Arendt says that genuine action possesses a "miraculous" quality because it is immersed in the diversity of interests and values of human life and brings into being new thoughts and relationships amongst the public that had never before existed (p. 178). As such, action is very much an "event," not a process or idea, which takes place at a particular time and place when a human attempts to influence others (p. 259).

Arendt carefully delineates how, over the course of human history, labor, defined as the repetitive and painful activities required to sustain one's life, has ascended from the domestic sphere of life to dominate the public sphere. During the industrial revolution the activities of *homo faber* (i.e., the activities of production) were broken into small unskilled movements, so that the coordinated activity of labor became central to economic and social life (p. 123–124). When the system of production requires any indistinguishable person to constitute its basic movements, humans become defined by their "labor power" and not by their skills or actions, and person labors so that they might "earn a living" (p. 125–128). Of significance to Arendt is that the purpose of the machine, organization, or society is not primarily to build something of lasting significance or value to humanity, but instead to increase consumption that *by definition* cannot have lasting value (p. 145–149). Arendt clearly identifies the negative consequences of laboring for its own sake: destructive tendencies go unquestioned and those caught in the never-ending labor and consumption cycle face existential unhappiness. As a consequence of consumption's ascendency in the public sphere, the activity Arendt labels as the greatest form of human activity, "action," has been subsumed, and its place in the public sphere (and the public sphere itself) has diminished.

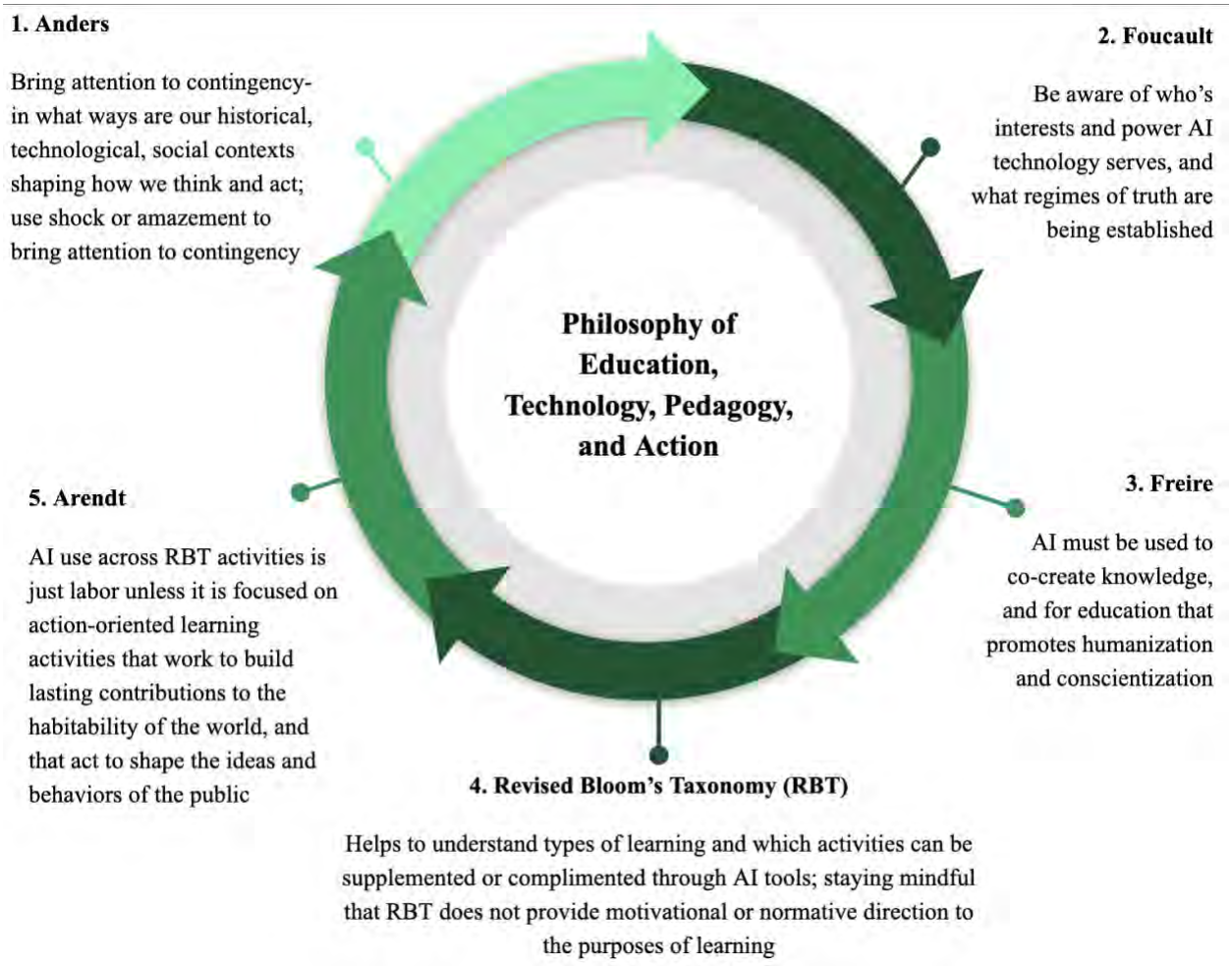
Applying Arendt's framework to the question of AI and its use in educational spaces, it moves the question of technological use from a scientific or pedagogical one (as in the RBT) to an ethical or political one. Let's consider an example of how AI might be used with current GPT technology in a classroom. A journalist, under pressure to produce more consumable content for its struggling publication, uses a GPT to write a story about the benefits and costs of electrical vehicle production and use. A teacher, excited by the labor-saving allure of an AI teaching assistant product called Brisk, uses the software extension to read the news story about electric vehicles and design a 60-minute lesson plan for their students, complete with learning goals, discussion prompts, a presentation activity, and summary quiz about the reading. The students, given carte blanche to use their school-provided Chromebooks, "read" the story using an AI platform like Perplexity, which provides summary analysis and key takeaways for them to use in their discussion and respond to the quiz. Simultaneously, they use Microsoft's AI image generator to create a slide deck for the class to graphically represent their group's ideas. The teacher completes the assessment cycle by having their AI assistant grade the quizzes, provide feedback to the students, and input their scores into a learning management system.

In this scenario, the AI engages in activities of labor and consumption, while all of the parties involved advance nothing of lasting significance, and if debate or critical reflection arise amongst students it is an incidental, rather than planned, outcome of the AI-prescribed lesson. Indeed, the Brisk teaching assistant might be well programmed to incorporate into the lesson features of the RBT such as understanding, evaluating, and creating activities; but unless a human being in this process is attuned to helping learners act in the world and make it a place, using Arendt's (1963/2006) words, "fit for human habitation" (p. 233), the most common educational experience might become, ironically, ones in which humans are unnecessary. As will be discussed in the next section, educational activities guided toward higher-order Arendtian work and action would focus on the tangible outcomes produced through the activity and opportunities provided to learners to critically think, reflect, and influence with peers and the broader community.

Framework for Ethical AI Use in Education

The work of Anders, Foucault, Freire, Bloom, and Arendt serve as the philosophical foundation for understanding how AI ought to be used in education. To summarize, Figure 2 illustrates the various philosophical threads that are discussed above, highlighting the core ideas from each philosopher, and how they relate to both AI use in education and each other. It is important to note the relationship between RBT and Arendt specifically, given that Arendt's concept of action is used to supplement RBT as a way to understand motivation and purpose more fully in the context of AI.

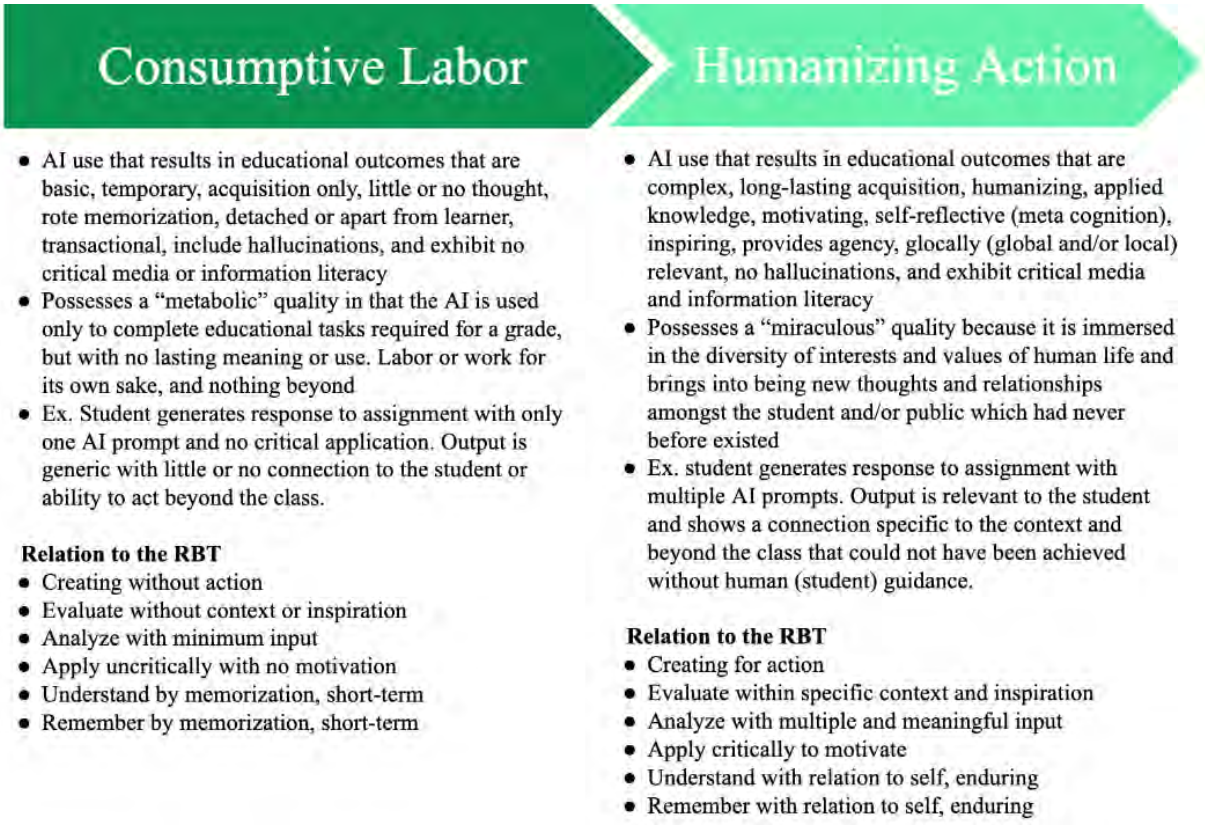
Figure 2
Philosophical Foundation for Ethical AI use in Education



Conceptualizing and imagining the philosophical underpinnings of AI use in education, however, is necessary but not sufficient for educators to confidently and proactively make decisions in the classroom. Recognizing that we easily feel compelled to adopt and use new technologies as they arise, Anders argues that we adopt a spirit of "shame," where we feel that we cannot stop our behaviors and practices "even if we can conceive of it" (Anders, 1936/2009, p. 152). The key for overcoming this shame, and the deterministic qualities of technology, according to Anders, is to use philosophical exaggeration and reflection to reveal the specific contingencies of our historical moment and contrast them with the particular indeterminacy of our individual selves, so we might begin to understand where we are capable of acting (p. 150). Stated another way, to find "the point of specific indetermination that the general determinability of man is possible" (p. 151). Coupled with Freire's concepts of humanization and conscientization, and Arendt's theory of action, and applied to educational practice, a goal for educators is to create public spaces for reflection, dialogue, and critical questioning; where responding to works of art, storytelling, politics, and science specifically designed to elicit a "shocking" space between the given technological immersion and the indeterminant experience of being a unique human, the learning community attempts to develop and interpret personal meaning and advocate for social purpose. Arendt reminds us that in such a public space, where authentic speech and deeds may have unforeseen outcomes, we develop important interpersonal capacities such as promise-making, forgiveness, and an appreciation for plurality.

Accordingly, the central finding of this philosophical analysis is to propose a framework for quickly and practically assessing AI use in educational settings. The framework, illustrated in Figure 3, is centered on student AI use and displays a spectrum from consumptive labor to

Figure 3
Framework for Ethical AI Use that Promotes Humanizing Action



humanizing action that draws from the work of Freire and Arendt specifically. The illustration is directional (from left to right), which exhibits how action-oriented and humanizing use of AI is ethically more appropriate in education than just using AI for labor and consumptive purposes alone. The purpose of this spectrum-based orientation is to allow educators to visualize where certain educational activities or AI-assisted student created content fits, from an ethical perspective. For example, if a student hands in an assignment that uses AI generated textual or multimodal content, it might fall towards the consumptive labor end of the spectrum if it clearly does not relate to that student's specific situation, has errors from potential hallucinations, or exhibits a generally uninspired position towards that activity. Conversely, student work on the humanizing action end will clearly show critical use of AI, relation to that student's perspective, or a demonstrated use for that student to act in the world.

It is important to note that brief examples of RBT are included in Figure 3 for each level of the RBT hierarchy. While early studies have shown that AI is perhaps more capable of performing tasks at the lower end of the RBT (i.e., understanding, applying), each type of these cognitive tasks are important to learning (Binh Nguyen Thanh et al., 2023; Lourenco et al., 2023). Moreover, the purpose of this framework and spectrum is to show how and where AI can

function ethically across the RBT. Indeed, AI may eventually enable students to complete tasks at all levels of the RBT. However, a student creating (the highest order of RBT) content with little thought or lasting impact on themselves or the world means less than if it pushes a student towards action. The framework and spectrum orientation in this case is specifically intended to serve as a guidepost or compass of sorts that educators can aim towards when assessing if a specific AI use case is appropriate and ethical. To clarify the use of this framework with RBT explicitly included, examples are provided below to illustrate both its intended use and function in educational settings.

Applications

While the framework for student use of AI is certainly a large part of the ethical equation, these actions do not happen in a vacuum. The learning objectives, design, activities, and assessments must also be considered in real-world teaching scenarios. Given the fact that AI in education is here to stay, our roles as educators must undergo a paradigm shift. This change necessitates a kaleidoscopic approach, one that adapts to the ever-shifting capabilities and structures of AI, especially in imagining that the primarily text-based functionality will soon (perhaps by the time of publication!) be multimodal. Here we provide a few imagined use examples of applying the framework to educational contexts allowing for the educator to be both adaptive and kaleidoscopic in reimagining their pedagogy. The examples below are provided to operationalize the framework of AI use that promotes humanizing action in imagined classroom contexts. We start with guiding questions, that can ideally serve as guideposts for educators to formulate curricula, unit, and lesson plans that use the capabilities of AI:

1. In what ways are our historical, technological, social contexts shaping how we think and act; what activity or experience can shock learners into appreciating their contingency?
2. Will the technologies we are going to use advance humanizing ends? In what ways can the technology enhance or harm the co-creation of knowledge?
3. How can we design learning activities that have benefits beyond their own sake; how are the learning activities helping students to act in the world?
4. In what ways can AI reduce the burdens of teaching and learning labor while increasing the capacity to act in the world?

In the following, we provide examples of activities educators across all modalities (i.e., online, blended, and in-person) might implement in order to effectively use the powerful AI "regimes of truth" by centering humanization and moving towards action-oriented pedagogies. A modified lesson plan format is used given its familiarity with educators generally and to help make the application clear for classroom use. Accordingly, each example begins with the overarching purpose, followed by a summary of the learning activity, learning objectives, ethical use cases for AI, and possible assessments. Actionable learning objective words are borrowed from the RBT throughout, and all activities are designed to be adapted for students aged 10 and above.

Community-Based Political Action

Purpose

Allow students to see themselves as civic actors and participants with agency in their local communities.

Summary

This activity will help students identify and evaluate local political activities, apply their understanding to their own lives, and create a plan to become involved in a meaningful and individualized way. Student learning is focused partially on knowledge acquisition but, more importantly, on civic engagement and community participation. In the classroom exercise, students engage in a metacognitive exercise through collaboration with AI to learn about their local political landscape. Educators have students identify local political issues within their communities and allow the use of AI to facilitate researching related and applicable political platforms and candidates and evaluate platforms and candidates from their own unique student perspective. In this example, skillfully and critically using AI allows students to more deeply understand their own social, economic, and political contexts. Leveraging AI, students undertake the task of identifying local political representatives (town council, board members, etc.) and evaluate their respective platforms centering their own life experience and agency. A multifaceted activity like this serves as a catalyst for the development of critical thinking skills, as students navigate the complexities of evaluating platforms, juxtaposing their evaluations of said platforms within their own educational experiences, and articulating personal assessments of their effectiveness.

By encouraging students to assess the efficacy of platforms, the assignment transcends the realm of automated tasks, labor, compelling students to engage in a wholly human endeavor—conducting a nuanced and uniquely individual logical sequence. The ultimate outcome of this practice is a manifestation of contextual critical thinking, where students craft their own evaluations grounded in personal opinions regarding the effectiveness of identified platforms, thereby fostering a deeper understanding of their own educational experiences and their place within the broader context of community-driven educational paradigms.

Students will be able to

1. Identify and understand local political issues.
2. Understands themselves as part of a larger collective by reading and understanding local political discourse issues.
3. Evaluate political candidates and platforms.
4. Articulate students' interpretation and preference of political platforms and candidates.
5. Create an individual plan of action to become politically involved.

Humanizing Action-Oriented AI Use

1. AI streamlines research activities and allows students to focus their mental faculties around evaluating representatives and identifying biases, all while positioning themselves within their respective contexts, rather than simply learning about them and recalling that information.

2. Students use AI to identify what political dialogue exists in their local community (i.e., community-based discussion forum)
3. Students choose a specific political issue/debate that exists in their local community.
4. Students use AI to understand what political candidates or policies are involved in said issue.
5. Students evaluate candidates and policies according to their own perspective as a student/community member who is directly involved.

Assessment

1. Students write a position paper on their issue of choice using their evaluations and personal perspective as resources. Students are permitted to use and cite AI for information gathering while their positions are wholly their own. Evaluations of platforms are assessed based on foundational understanding and articulation of argument.
2. Students discuss and debate positions of similar topics. Students engage with one another, modify and amend their positions through collaborative thought and consensus.

Socratic Dialogue for Individualized Research

Purpose

Help students conceptualize what is important to them, ask exploratory and critical questions about the topic, and apply their self-directed research toward a meaningful outcome.

Summary

A central component of being human is to understand what matters to oneself, direct these interests toward knowledge acquisition, and discover ways to act in a meaningful and impactful way. In this example, students are tasked with researching a topic of their choosing both to learn about it and apply this knowledge to their own context. To facilitate this endeavor, AI acts as an agent of Socratic dialogue and questioning for the student, helping students generate research idea topics that will be specifically catered towards student interests. AI will be equipped to ask students questions regarding their level of interests and commitment, suggest other topics of potential interest based on specific student response in addition to refine students' thinking regarding logical sequencing of topic selection and eventually argument. This personalized approach allows them to analyze how these topics manifest in their own lives and communities, gaining valuable insights.

Equipped with these findings, students use AI to curate a comprehensive list of research questions and potential projects. By actively creating engagement plans for community outreach related to their chosen research topics, students' learning therefore transcends the classroom and students themselves become proactive contributors. This distinctive approach positions students as researchers and educators. They can share their findings with community members, fostering dialogue and promoting knowledge exchange. By using AI as a research tool and not just an information aggregator, students emerge as catalysts for community-based action and learning. In this example, the humanity of activity rests in an action-oriented interaction with a mentor whereas the labor of developing a research topic can be facilitated by AI. Within the ecosystem of AI-student-mentor, AI serves not as a replacement but as an enabler, enhancing the efficacy of human interaction by streamlining the initial stages of research ideation. This symbiotic collaboration between students and AI epitomizes the thoughtful integration of technology into

the learning process, nurturing a pedagogical environment where exploration and experimentation harmonize with the guidance and expertise provided by human mentors; thus, promoting more human ends and metacognition.

This multifaceted process exemplifies action-oriented, community-driven research, where students actively shape their learning journey while contributing to their communities.

Students will be able to

1. Identify a topic of interest to them.
2. Engage in dialogue discussing their own interests.
3. Develop a research topic and questions.
4. Conduct research to address questions.
5. Synthesize information and communicate their findings.
6. Create a community outreach plan based on findings.

Humanizing Action-Oriented AI Use

1. Student uses AI as a Socratic questioner to help students identify a research topic of interest.
2. Student identifies a specific research topic idea and uses AI to help guide research process.
3. Student uses AI to generate community outreach ideas regarding said research.

Assessment

1. Student presents research findings to the class.
2. Student engages in community-based outreach based on research findings.

Conclusion

Pandora's box is open and AI will forever be a part of our society and education systems. Educators must adapt strategies that continuously address and reconsider the Promethean discrepancy surrounding how AI is used, and toward what ends. In this critical analysis, we contend that the philosophical traditions on technology, education, and society can help guide how AI is used, similar to other technologies of the past. Anders and Foucault bring awareness of the relationships between technology and power, while Freire and Arendt ground education as a humanizing force that must build towards action. Taken together, we argue that a framework for ethical AI use in education across all online, blended, and in-person modalities, must not just a means to complete the temporary and metabolic purposes of previous educative goals, but rather a new way to help students and teachers alike realize their full humanity and agency to act in relevant and meaningful ways in the world.

While this framework is designed around student use of AI, we argue that teachers and instructional designers must also use similar guideposts for using AI to create and implement teaching activities. That is, AI use in education demands not just a rethinking of what is appropriate given our current educative goals, but also a reimagining and redesigning of our pedagogy more generally. To this end, we provide imagined use examples, and argue that the framework itself is useful to be applied across varied educational contexts. Educators will need

continuous creativity to reframe and rethink teaching, learning, and assessment as AI continues to develop. As is often the case with any ethical debates, the real tests will come in the gray areas of AI use, and we call for future research to be conducted in these spaces. While the functionality of AI might be outdated by the time this is published, we argue these ethical guideposts will not. Humanizing education that helps foster agency in the world is a timeless goal—it grounds humanity in an ever-changing world.

Conflict of Interests Statement

The authors declare that they have no competing interests.

References

An, T., & Oliver, M. (2021). What in the world is educational technology? Rethinking the field from the perspective of the philosophy of technology. *Learning, Media and Technology*, *46*(1), 6–19.

Akgun, S., & Greenhow, C. (2022). Artificial intelligence in education: Addressing ethical challenges in K–12 settings. *AI and Ethics*, *2*(3), 431–440. https://doi.org/10.1007/s43681-021-00096-7

Anders, G. (1956/2002*). Die Antiquiertheit des Menschen: Über die Seele im Zeitalter der zweiten industriellen Revolution*. C.H. Beck.

Anders, G. (1962). Theses for the Atomic Age. *The Massachusetts Review*, *3*(3), 493–505. http://www.jstor.org/stable/25086864

Anders, G. (2015). *The obsolescence of man, volume II: On the destruction of life in the epoch of the third industrial revolution.* (Recluse, A., Trans.) Libcom.org. (Original work published 1980). Accessed October 30, 2023, at: https://libcom.org/book/export/html/51647

Anders, G. (2009). The pathology of freedom. (Wolfe, K., Trans). In Bischof, G., Dawsey, J., & Fetz, B. (Eds.) (2014). *The life and work of Günther Anders: émigré, iconoclast, philosopher, man of letters*. Studien Verlag, Austria. (Original work published in 1936).

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Arendt, H. (1958/2013). *The human condition* (2nd ed.) University of Chicago Press.

Arendt, H. (1963/2006). *Eichmann in Jerusalem*. Penguin Publishing Group.

Armstrong, P. (2010). *Bloom's Taxonomy.* Vanderbilt University Center for Teaching. https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52–62.

Thanh, B. N., Vo, D. T. H., Nhat, M. N., Pham, T. T. T., Trung, H. T., & Xuan, S. H. (2023). Race with the machines: Assessing the capability of generative AI in solving authentic assessments. *Australasian Journal of Educational Technology*, *39*(5), 59–81. https://doi.org/10.14742/ajet.8902

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. Longmans.

Boyd, D. (2016). What would Paulo Freire think of Blackboard. *International Journal of Critical Pedagogy*, *7*(1), 22.

Bozkurt, A. (2024). GenAI et al.: Cocreation, authorship, ownership, academic ethics and integrity in a time of generative AI. *Open Praxis*, *16*(1), 1–10. https://doi.org/10.55982/openpraxis.16.1.654

Catalani, F. (2020). Anticipation as critique: Objective phantasy from Ernst Bloch to Günther Anders. *Praktyka teoretyczna*, *35*, 149–166.

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, *66*(4), 616-630.

Crawford, J., Cowling, M., & Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, *20*(3), 02.

Crompton, H., Burke, D., & Lin, Y.-C. (2019). Mobile learning and student cognition: A systematic review of PK-12 research using Bloom's Taxonomy. *British Journal of Educational Technology*, *50*(2), 684–701. https://doi.org/10.1111/bjet.12674

Damaševičius, R. (2023). The rise of ChatGPT and the demise of Bloom's Taxonomy of Learning Stages. In *Creative AI Tools and Ethical Implications in Teaching and Learning* (pp. 115–134). IGI Global. https://doi.org/10.4018/979-8-3693-0205-7.ch006

Dawsey, J. (2013). *The limits of the human in the age of technological revolution: Günther Anders, post-Marxism, and the emergence of technology critique*. (Publication No. 3568370) [Doctoral dissertation, University of Chicago]. ProQuest Dissertations & Theses Global. https://www.proquest.com/openview/97326c30c7b44c5e6e7684e17262cbe2/1?pq-origsite=gscholar&cbl=18750

Dobson, J. E. (2023). On reading and interpreting black box deep neural networks. *International Journal of Digital Humanities*, *5*(2), 431–449.

Elsayed, S. (2023). *Towards mitigating ChatGPT's negative impact on education: Optimizing question design through Bloom's Taxonomy* (arXiv:2304.08176). arXiv. http://arxiv.org/abs/2304.08176

Farag, A., Greeley, L., & Swindell, A. (2021). Freire 2.0: Pedagogy of the digitally oppressed. *Educational Philosophy and Theory*, 1–14. https://doi.org/10.1080/00131857.2021.2010541

Farrelly, T., & Baker, N. (2023). Generative Artificial Intelligence: Implications and considerations for higher education practice. *Education Sciences*, *13*(11), 1109. https://doi.org/10.3390/educsci13111109

Freire, P. (1965). *Education for critical consciousness*. Bloomsbury Academic.

Freire, P. (1968). *Pedagogy of the Oppressed: 30th Anniversary Edition*. Bloomsbury Publishing USA.

Freire, P. (1997). *Politics and education*. UCLA Latin American Center.

Freire, P. (2005). *Education for critical consciousness*. Continuum.

Fuchs, C. (2017). Günther Anders' undiscovered critical theory of technology in the age of big data capitalism. *tripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society*, *15*(2), 582–611.

Giantini, G. (2023). The sophistry of the neutral tool: Weaponizing artificial intelligence and big data into threats toward social exclusion. *AI and Ethics*, *3*(4), 1049–1061. https://doi.org/10.1007/s43681-023-00311-7

Jin, S.-H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, *20*(1), 1–21. https://doi.org/10.1186/s41239-023-00406-5

Kahn, R., & Kellner, D. (2004). New media and internet activism: From the "Battle of Seattle" to blogging. *New Media & Society*, *6*(1), 87–95. https://doi.org/10.1177/1461444804039908

Laricchia, F. (2023) *Smartphone penetration worldwide as share of global population 2016–2022*. Statista. Retrieved January, 6 2024, from https://www-statista-com.proxy.libraries.rutgers.edu/statistics/203734/global-smartphone-penetration-per-capita-since-2005/

Lin, Jingjing, ChatGPT and Moodle Walk into a Bar: A Demonstration of AI's Mind-blowing Impact on E-Learning (March 20, 2023). Available at SSRN: https://ssrn.com/abstract=4393445 or http://dx.doi.org/10.2139/ssrn.4393445

Loftus, M., & Madden, M. G. (2020). A pedagogy of data and artificial intelligence for student subjectification. *Teaching in Higher Education*, *25*(4), 456–475. https://doi.org/10.1080/13562517.2020.1748593

Lourenco, A. P., Slanetz, P. J., & Baird, G. L. (2023). Rise of ChatGPT: It may be time to reassess how we teach and test radiology residents. *Radiology*, *307*(5), e231053. https://doi.org/10.1148/radiol.231053

Lorenzini, D. (2015). What is a "Regime of Truth"? *Le foucaldien*. *1*(1). doi:10.16995/lefou.2

Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., ... & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers and Education: Artificial Intelligence*, *3*, 100056.

Marr, B. (2024, February 13). AI Showdown: ChatGPT vs. Google's Gemini—Which reigns supreme? *Forbes*. https://www.forbes.com/sites/bernardmarr/2024/02/13/ai-showdown-chatgpt-vs-googles-gemini-which-reigns-supreme/?sh=518ed7746072

Masapanta-Carrión, S., & Velázquez-Iturbide, J. Á. (2018). A systematic review of the use of Bloom's Taxonomy in computer science education. *Proceedings of the 49th ACM Technical Symposium on Computer Science Education*, 441–446. https://doi.org/10.1145/3159450.3159491

McIntosh, T. R., Liu, T., Susnjak, T., Watters, P., Ng, A., & Halgamuge, M. N. (2023). A culturally sensitive test to evaluate nuanced GPT hallucination. *IEEE Transactions on Artificial Intelligence*, 1–13. https://doi.org/10.1109/TAI.2023.3332837

Megahed, F. M., Chen, Y.-J., Ferris, J. A., Knoth, S., & Jones-Farmer, L. A. (2023). How generative AI models such as ChatGPT can be (mis)used in SPC practice, education, and research? An exploratory study. Quality Engineering, *36*(2), 1–29. https://doi.org/10.1080/08982112.2023.2206479

Müller, C. J. (2016). *Prometheanism: Technology, digital culture and human obsolescence*. Rowman & Littlefield.

Müller, C. J. (2019). From radioactivity to data mining: Günther Anders in the Anthropocene. *Thesis Eleven*, *153*(1), 9–23.

Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B.-P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, *28*(4), 4221–4241. https://doi.org/10.1007/s10639-022-11316-w

Ofgang, E. (2023, December 20). Gemini: Teaching with Google's latest AI. *TechLearningMagazine*. https://www.techlearning.com/news/gemini-teaching-with-googles-latest-ai

Schraube, E. (2005). "Torturing things until they confess": Günther Anders' critique of technology. *Science as Culture*, *14*(1), 77–85.

Şenocak, D., Bozkurt, A., & Koçdar, S. (2024). Exploring the ethical principles for the implementation of artificial intelligence in education: Towards a future agenda. In R. Sharma & A. Bozkurt (Eds.), *Transforming education with generative AI: Prompt engineering and synthetic content creation* (pp. 200–213). IGI Global. https://doi.org/10.4018/979-8-3693-1351-0.ch010

Stapleton, L., Taylor, J., Fox, S., Wu, T., & Zhu, H. (2023). *Seeing seeds beyond weeds: Green teaming generative AI for beneficial uses* (arXiv:2306.03097). arXiv. http://arxiv.org/abs/2306.03097

Swartz, M., & McElroy, K. (2023). The "Academicon": AI and surveillance in higher education. *Surveillance & Society*, *21*(3), 276–281. https://doi.org/10.24908/ss.v21i3.16105

Rahm, L. (2023). Education, automation and AI: A genealogy of alternative futures. *Learning, Media and Technology*, *48*(1), 6–24.

Rountree, B., & Condee, W. (2021). The nonmaterial mirror: Performing vibrant abstractions in AI networks. *Theatre Journal*, *73*(3), 299–318.

Wong, M. (2023, June 23). The chatbots may poison themselves. *The Atlantic*. https://www.theatlantic.com/technology/archive/2023/06/generative-ai-future-training-models/674478/