Human-Centric Intelligent Systems
https://doi.org/10.1007/s44230-026-00168-z

RESEARCH

The Dizzy Spells of AI: Enacting Intellectual Independence in AI Literacy and AI-Mediated Learning

Hugo-Henrik Hachem[1,2] · Franziska Primus[3] · Katerina Linden[1] · Fredrik Heintz[1] · Mattias Wiggberg[4]

Received: 14 November 2025 / Revised: 14 July 2026 / Accepted: 21 July 2026
© The Author(s) 2026

Abstract
The integration of artificial intelligence (AI) into education and learning risks leading to intellectual dependence among users of AI tools. Meanwhile, intellectual dependence remains underexplored in AI-mediated learning and is addressed ambiguously in AI literacy programmes, which claim to foster critical thinking (CT) and critical consciousness (CC) among learners, helping them move from naïve to more critical ways of thinking about and using AI. However, such approaches to AI literacy often presume and reinforce a hierarchy of thought: teachers are seen as critical thinkers about AI, while students are viewed as less so. This hierarchy arguably leads to intellectual dependence of the latter on the former. In this paper, we problematise AI literacy programmes that embrace CT and CC as educational objectives, since, as we argue, they remedy intellectual dependence on AI tools (e.g. ChatGPT, Claude…) by subjecting AI users/learners to another intellectual dependence, but on AI teachers. By drawing on Jacques Rancière's ideas of universal intellectual equality, we argue that enacting intellectual independence is not an educational or learning goal to be achieved in the future, but a starting point—a condition necessary for teaching and learning about and with AI. Hence, we apply this assumption of equality of intelligence to both AI literacy and AI-mediated learning, and illustrate this application with two practical scenarios: (1) AI-centred pedagogy (thing-centred AI literacy) and (2) ignorant AI-tutors (thing-centred AI-mediated learning). Both approaches aim to reduce dependence on human teachers in AI literacy or on AI tutors in AI-mediated learning. Ultimately, this paper presents design principles for these two solutions, offers a philosophical premise for rethinking education, teaching, and learning in the age of AI, and concludes with contributions, limitations, and a research agenda.

Keywords Intellectual emancipation · Intellectual dependence · AI literacy · Critical thinking · Critical consciousness · AI-mediated learning

1 Introduction

What if, in the age of artificial intelligence (AI), intellectual hierarchies between teachers and students could no longer form the basis for teaching? The teacher's role would no longer be to assess students' thinking, but rather to motivate and evaluate the effort invested in the thinking that leads to knowledge acquisition, all while modelling intellectual freedom to students. Today more than ever, it could be possible to imagine such an education without traditional exams, with teachers acting as "guide on the side" rather than as wise and critical thinkers on stage. A new master's programme in engineering in the United Kingdom embraces such ideals and recognises that talent is evenly distributed, but opportunity is not [1].

✉ Hugo-Henrik Hachem
Hugo-henrik.hachem@liu.se

1 Reasoning and Learning Lab, Division of Artificial Intelligence and Integrated Computer Systems, Department of Computer and Information Science, Linköping University, Linköping, Sweden

2 Department of Education, Faculty of Arts and Sciences, American University of Beirut, Beirut, Lebanon

3 Department of Education, Örebro University, Örebro, Sweden

4 Department of Industrial Economics and Management, KTH Royal Institute of Technology, Stockholm, Sweden

Published online: 30 July 2026

> Those who flunk out of engineering don't do so because the subject is too hard, but because they're bored […]. An early experiment […] was asking a team of 18-year-olds to design, build and demonstrate a pulse oximeter, a device that measures pulse rates and oxygen saturation levels in blood. They didn't even know what it was. Five weeks later, they had it working. (p. 1092)

Meanwhile, emerging AI-related sociotechnical imaginaries of education seem to be more in line with traditional intellectual hierarchies, which both necessitate and legitimise certain institutional structures, including explication, assessment, and qualification. This is the case, for example, in the teaching and learning of AI (AI literacy efforts) or even in learning and teaching with AI (AI-mediated learning and teaching). These hierarchies risk propagating intellectual dependence: in the first case, when AI literacy adopts critical thinking (CT) and critical consciousness (CC) as educational goals; and in the second, when learners submit their intellect to AI tutors. How? The educational goal whereby teachers raise learners' thinking or consciousness levels about AI presumes, by deduction, that learners' current level of thinking is below critical, which embodies an intellectual hierarchy that consecrates subsequent intellectual dependence.

Nevertheless, imagining education in the age of AI without intellectual hierarchies requires some level of philosophical provocation, design speculation, and implementable tools. This paper aims to provide a philosophical framework and tools to address intellectual hierarchies in AI literacy and AI-mediated learning by developing a postcritical [2] imagination of the teacher's role (whether human or AI-tutor)—one that shifts from sage to guide, from a mastery built on knowing to "mastery" that is grounded in not knowing or "ignorance" [3]. This work argues that AI literacy efforts that take CT and CC as educational goals can, by design, be incapable of promoting intellectual independence when learning about or with AI. This work is a conceptual/philosophical contribution to the literature on AI literacy and AI-mediated learning, proposing a new approach to education and learning in the age of AI and reminding colleagues of the importance of anchoring educational practice to philosophy [4]. Otherwise, educational practice risks feigning objectivity or neutrality while reproducing intellectual and societal inequalities [5, 6]. This work bridges educational philosophy with practice by introducing design principles that frame (user-AI) systems and AI literacy practices. The rationale for writing this paper is presented next.

2 The Dizzying Effect of AI [On/In] Education

Education is experiencing what could resemble vertigo, as the advent and deployment of AI in educational institutions could challenge not only their role but also the inherent power structures that characterise it. Whereas the practice of AI literacy and AI-mediated learning could precipitate reconsideration of intellectual hierarchies and power dynamics within educational institutions, we see these power relations unchecked. This paper problematises intellectual hierarchies and dependence unintentionally reproduced in AI literacy and AI-mediated learning from a much-needed philosophical angle [4, 7].

Inherent in the argument that AI literacy efforts, by the time they conclude, should aim to raise thinking and consciousness levels about AI is an implicit assumption of an intellectual hierarchy about AI between a naïve student and a critical teacher. This hierarchy underpins the intellectual dependence of the first on the second, governing the teacher-student relationship, eventually legitimising the continued need for "sage" AI teachers [6, 8]. As a tangent to dependence in AI literacy, intellectual dependence on AI tutors is also observed in AI-mediated learning. There, intellectual dependence on AI tools, such as large language models (LLMs), is cited as a serious threat to metacognition in students [9–11]. In light of this threat, AI literacy scholars call for the mobilisation of the CC and CT educational logic [12–16], described below, and anticipate nothing short of imminent educational revolutions. Meanwhile, revolutions, especially those radical in nature, imply a power shift and the reconsideration of predating [intellectual] hierarchies [5], which are currently unaddressed, if not reinforced even, in AI literacy efforts. How can AI literacy premised on intellectual dependence on AI teachers/experts be the cure for intellectual dependence on AI tutors and AI tools in general?

The literature addresses this question through what we term a triple bind, connecting CC and CT to AI literacy and AI-mediated learning (see Fig. 1). Evidence for this triple bind can also be found in [17]. First, users of AI tools need to acquire CT and CC about AI through AI literacy [15, 17], so they can, subsequently, engage critically and responsibly in AI-mediated learning [18]. Only then could learning with AI enable users to improve their CT and CC about the subject matter they wish to learn [19]. In this triple bind, we see an impossible paradox: Intellectual dependence on AI is a serious risk [10], which AI literacy efforts, framed in intellectual dependence on AI teachers, are supposed to mitigate. Addressing this paradox requires delving into the inherent ambiguity in framing CC and CT as educational goals, which leads to the identification of two problems.

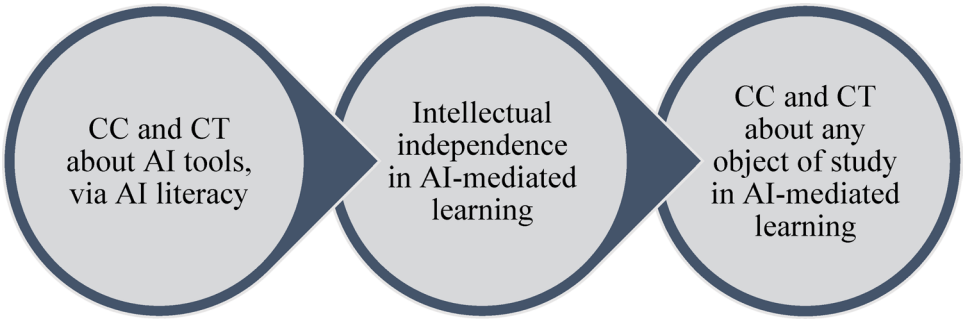

Fig. 1 The triple bind involving CC and CT with AI literacy and AI-mediated learning

2.1 Problem One: Ambiguous Conceptualisations of CC and CT

CT and CC are two domain-specific psychosocial concepts that have been turned into educational goals and woven into educational logics, hereafter referred to as the CC and CT logic, which have direct consequences for the roles of teachers, learners, and the nature of the subject matter [20]. The definition and evaluation of CT and CC as educational goals for AI literacy, aimed at mitigating intellectual dependence on AI, seem problematic at both theoretical and empirical levels. CT about AI tends to be defined—if at all [21]—in a somewhat ambiguous and self-referential manner [13, 16, 22]. CT's empirical evaluation in learners can also be self-referential, defined by psychometric subconstructs dubbed 'critical appraisal' of AI [23] and its evaluation in teachers relies on technological consciousness by mobilising Paolo Freire's radical and consciousness-raising pedagogy [15]. These definitions and evaluations of CT and CC can be based on self-justifying, tautological criteria [20], for example, when students' thinking and experience about AI are judged as *naïve* because they *do not understand AI or its consequences*. Thus, when naiveté is operationalised in this tautological fashion, it cannot enter a causal relationship with (not)understanding, which means naiveté or critical thought can hardly ever be empirically (dis)proven [6]. When naïveté and critical thought cannot be causally operationalised and evaluated, intellectual inequality between teachers and students remains justified, since it is the teacher who decides at what level students' thinking and consciousness about AI rests at a particular point in time.

2.2 Problem Two: Intellectual Inequality

If students were not assumed to be naïve by default, why would learners' thinking levels need to be raised at the hands of so-called critical teachers? Apart from ambiguity in how CC and CT are defined and evaluated in relation to the critical teacher, AI literacy efforts that claim to promote CC and CT as educational goals require and assert two irreconcilable levels of thought or intelligence between teachers and students, see [24]. They can also presuppose a certain level of distrust of learners' experiences with/about AI when those experiences are judged to be the result of *naïveté*. How so? The CT/CC characteristic of the AI expert or teacher (who often advocates for it) contrasts with a presumed uncritical level of thinking ascribed to students or laypersons (who allegedly need it). How long will this intellectual inequality persist? Given the tautological nature of evaluation/assessment of CT and CC, one could argue that the answer is as long as students have not reached the same conclusions about AI risks and affordances as their knowledgeable teachers, use AI in the same way experts recommend, or 'perform' well on tautological assessment and psychometric tests that their teachers have developed. Until then, they may not yet be critical thinkers or critically conscious about AI. As such, CT and CC, initially tied to free thinking [25], become gatekeepers of conformity and intellectual dependence on AI teachers, whose supposed task is to evaluate the criticality of learners' thinking; see [6, 26]. Our problematisation above also applies to AI-mediated learning when human teachers are replaced by (Socratic) AI-tutors, with the hope of raising learners' CT about a subject, see [12, 19, 27–30]. These problems will be fleshed out and addressed in the coming sections.

2.3 Ignorance as the Answer for these Two Problems

Before proposing a solution to these two problems, we remind the readers that our challenge of CC and CT is not absolute but targets only their framing as educational goals, because such framing unintentionally creates intellectual hierarchies. However, outside such a framing, the authors maintain that genuine CC and CT characterise free thought and responsible action in wider society, but that they may result from hard work and engagement with a body of knowledge [17], especially since CT and CC are domain-specific [31], rather than necessarily a capacity transmitted from a critical teacher to a naïve student.

To address the problems of conceptual ambiguity and intellectual hierarchies summarised above, we propose that AI literacy and AI-mediated learning should not concern

themselves with evaluating learners' levels of thinking or consciousness. We already know that what is premised on inequality may not truly lead to equality [32], meaning that learner-centred critical AI literacy, which takes learners' naïve thinking as a departure point, may not lead them to become critical thinkers [24] about AI without stripping free thought of critical thinking. That is why we propose a shift in focus from thinking and consciousness as the centre of educational/learning goals to attention and motivation as goals to study (about/with) AI [33] in order to safeguard learners' intellectual independence with teachers or AI tutors, to the extent possible and against all odds.

LLMs can fuel the growing problem of intellectual dependence on AI, which, as we speculate, may only worsen as LLMs become more factually accurate. That is why we borrow Jacques Rancière's [6, 34] treatment of intellectual emancipation to understand intellectual independence in an age of AI as: the acceptance of and belief in one's intellectual equality not only vis-à-vis AI teachers/experts, but also when learning with AI. One can know more if one wishes to do so without the necessary explanations from another (human or machine). Rancière's intellectual emancipation, then, helps address (1) how AI literacy could be premised on intellectual equality between human teachers and learners. Furthermore, (2) how can AI tutors be used to strengthen users' intellectual independence in AI-mediated learning?

To answer these questions, we opt for a postcritical approach, meaning we go beyond mere critique [2]. We propose constructive solutions towards a public pedagogy of/with AI that transcends conventional intellectual and educational hierarchies reproduced across enduring institutions. The first is a plea for an AI-centred pedagogy (or a thing-centred AI literacy), and the second is a call to enact thing-centred AI-mediated learning by creating and prompting 'ignorant' AI-tutors to ask (a non-guiding type of) questions, cf. [29]. This article, promoting an epistemology of ignorance, understands ignorance and knowledge as dialectic rather than opposites [3] [35], while enacting radical symmetry in the relationship between teachers (human or AI) and students and giving practical examples of how to leverage ignorance in AI-centred pedagogy and thing-centred AI-mediated learning. Next, we expand on the triple bind (see Fig. 1) and offer a brief literature review of the complex dynamics governing CC and CT in relation to AI literacy and AI-mediated learning.

3 AI-Mediated Learning, AI Risks and AI Literacy

AI originated as the science and engineering of creating intelligent machines, and today it largely focuses on machine learning, natural language processing and computer vision [36]. Its popularisation across various life domains has led to considerable attention to the nature, provision, and acquisition of AI knowledge [17, 36–41]. AI has also been examined from a Science and Technology Studies perspective, challenging both the assumption that AI acts as a cognitive authority capable of replacing teachers and the assumption that AI is merely a neutral assistant. In this vein, Actor-Network Theory (ANT) encourages a view of LLMs as actors (actants) within the assemblage of teaching and learning, on a par with human teachers and students in shaping knowledge, learning practices, epistemic norms, and pedagogical authority. This alludes to an active role for AI technology, as it treats human and non-human actants as symmetrical in complex processes of knowledge production [42, 43]. ANT has already been used to empirically investigate how AI and LLMs are integrated into the teaching and learning ecosystem [44–46]. The perspective advanced in this paper aligns with these ideas.

In light of the triple bind (see Fig. 1), which links AI literacy and AI-mediated learning to CT and CC, the following section describes the intellectual risks associated with AI-mediated learning and how critical AI literacy is intended to mitigate them.

3.1 AI-Mediated Learning: Intellectual and Societal Risks

With the release of LLMs (ChatGPT, Claude, etc.), new learning affordances emerged alongside intellectual and societal risks. On the one hand, LLMs are said to promote independent and personalised learning via self-tutoring (AI-tutor, AI-coach, AI-mentor, AI-teammate, AI-tool, AI-simulator, and AI-student), to enhance "smartness", spark curiosity, encourage stylistic experimentation, provoke new analytical questions, [47] and leave room for digital and reflexive forms of assessment, as well as accessibility and inclusivity [10, 11, 38, 39, 48–51]. On the other hand, using LLMs for learning is constrained by inherent limitations which may entail intellectual risks. LLMs suffer from hallucinations, misalignment, discrimination, and lock-in problems [39, 52, 53], an ambiguous politics of visibility [53], and a lack of moral commitment to truth or to knowledge disciplines when acting as tutors [54]. Meanwhile, AI users risk developing intellectual dependence on AI [14], complacency with AI [55], AI addiction [10], cognitive debt [56], and metacognitive laziness [9]. The definitions of these

Table 1 Definitions of intellectual risks associated with the use of AI-tutors for learning

Intellectual risk	Definition
Metacognitive laziness	The giving up of one's responsibility for learning [9]
Intellectual dependence on AI	Relying on AI to perform tasks on students' behalf [14]
AI addiction	Behavioural dependency on and excessive use of AI [10]
Complacency with AI	Over-trusting AI by following incorrect recommendations [55]
Cognitive debt	The result of reproducing AI suggestions without evaluating their accuracy or relevance, and forfeiting ownership of the ideas and risking the internalisation of shallow or biased perspectives [56]

concepts are summarised in Table 1, and they commonly manifest as the swift generation and trust in answers via AI that would otherwise have required substantial intellectual effort and deep learning [57]. The argument goes that these risks stem from a lack of CT and CC (or naïveté) about AI tools.

In light of these intellectual risks, AI literacy programmes argue that CT about AI is necessary for using LLMs responsibly. On the one hand, learners and AI users are liable, and they must exercise vigilance. Lin [49], for example, warns that LLM dependence reflects 'a lack of critical thinking about the various materials provided by this tool' (p. 7), ultimately leading to complacency and blind acceptance of incorrect recommendations [55]. The argument continues that learning is an active, goal-directed process that requires self-regulation of thoughts, feelings, and behaviours [58]. Metacognition, or the 'awareness of and knowledge about one's own thinking' (p. 65), encompasses understanding of learning goals and strategies for achieving them. However, as learners use AI, they may be less able to associate responsible metacognitive processes with their learning tasks [9], thereby cutting corners, which leads them to forfeit knowledge ownership [56], typically gained through attentive and deep engagement [57] (as opposed to the easy way out). Even if students who learn with AI may seem "smarter" and more able to reason, the cognitive performance gains do not necessarily scaffold metacognition [11].

The effects of learning and doing with AI on memory and creativity have also been explored from the viewpoint of *Homo Promptus*. Linden and colleagues [59] describe how the practice of prompting may eventually transform the ways humans think, remember, and create. Rather than assisting cognition, LLMs could redistribute it by externalising memory processes [60], shifting knowledge retention from human recall to machine retrieval, and recasting creativity as a collaborative human–AI activity. In this sense, the potential emergence of *Homo Promptus*, whose thinking is shaped by interaction with algorithmic outputs, signifies an epistemic transformation in which learning with AI alters the very conditions of knowledge production, agency, and intellectual autonomy, which this paper eventually addresses.

Furthermore, the convincing nature of LLMs exacerbates the problem. LLMs can mimic human skills such as reasoning, problem-solving, and creativity, and engage in human-like conversations. They use sentiment analysis and machine learning to understand the context of the conversation and provide "appropriate" and probabilistic responses [61], which may inadvertently convey a false sense of trustworthiness to the user. In such a case, the extent to which AI-mediated learning can lead to CT about an object of study may be diminished. When LLMs have been promoted as effective tutors, learning assistants, and coaches [49], some have called for resurrecting Socrates and prompting LLMs for Socratic dialogic interactions to enhance learners' CT when learning with AI [3, 19, 27–30]; others countered these calls for AI-based tutoring as an epistemic perversion, claiming that 'part of the solution here is to cut the bullshit' since AI cannot serve as a collaborator or a tutor and should not be said to do so, if the meaning of those roles is not to be corroded [54] (p. 7), citing hallucinations, bias, [52, 53] and moral disengagement on the part of LLMs [54]. In this line of argument, see e.g. [49], CT and CC about AI are causally juxtaposed with intellectual dependence, as shown in Fig. 1. There, the triple bind suggests that a lack of CT and CC leads to intellectual dependence on AI, whereas one could argue that a lack of CC and CT is itself tantamount to intellectual dependence, thereby defying a causal logic between CT and CC about AI and intellectual independence. In any case, AI literacy grounded in the CT and CC logic is mobilised to "enlighten" naïve learners about these intellectual and societal risks, as we show next.

3.2 AI Literacy: Against Intellectual and Societal Risks

AI literacy spans the entire educational spectrum (formal and non-formal) and is conceived as a conduit for developing a "critical" understanding of intellectual and societal AI risks. It promotes sociotechnical knowledge about AI, encompassing theoretical, practical, and judgmental aspects [62]. AI literacy is a set of competencies 'that enable individuals to *critically evaluate* AI technologies; communicate and collaborate effectively with AI; and use AI as a tool online, at home, and in the workplace', Long and Magerko in [62] (p. 2), and it includes 'basic knowledge and analytical evaluation of AI, as well as *critical use* of AI applications by non-experts' [23] (p. 1, italicised by authors for emphasis).

AI literacy encompasses various yet potentially competing perspectives [38] and models [63]. A *cognitive* perspective examines human cognition and metacognition, particularly how AI supports learning processes (e.g., [9]). A *humanistic* perspective emphasises human values such as creativity, complex problem-solving, and critical analysis (e.g., [13]). A *social* perspective recognises that broader social and cultural expectations influence the use of AI [15]. Additionally, AI literacy can be modelled around a *generic* set of individual skills (e.g., [22]), *situated social practices*, or *transformative outcomes* for individuals and society (e.g., [40]). The difference between these skills/competences lies in whether they are observable/evaluated at the individual or societal level, i.e. whether in individual thinking and behaviour or collective consciousness and action.

Anchored to Bloom's Taxonomy and knowledge hierarchies (see Fig. 2), AI literacy should provide 'basic competencies to know and understand, use and apply, as well as evaluate and create with AI' [36] (p. 9). Here, we note that so-called lower-order thinking, to know, understand and apply, is not considered critical thinking, as noted for example by Darwin and colleagues [18] and illustrated in Fig. 2. This implies that the learning objectives of critical AI literacy are distributed across Bloom's hierarchies of knowledge and knowledge acquisition capacities (know, understand, and apply vs analyse, evaluate, and create), allowing a clear manifestation of what intellectual hierarchies look like in educational practice (see also Sect. 5.1), as afforded and enabled by Bloom's taxonomy. Across AI literacy models and perspectives, the CT (cognitive/humanistic) and CC (situated/transformative) logic almost always takes a front-row seat, as shown in a recent international Delphi study with AI experts [17]. Depending on the extent of criticality embedded in the design and enactment of AI literacy efforts, the goal ranges along a "criticality" continuum from CT to CC, which will be defined and nuanced in Section Four.

The CT and CC logic in AI literacy involves learners, teachers, and AI-curricula, and rejects a passive acceptance of technology. Ng and colleagues [36] postulate that learner-centred AI literacy curricula must steer students beyond basic technical skills, fostering their ability to analyse AI's strengths and weaknesses, the ethical aspects of technology, including the inherent biases present in AI systems, and the broader societal effects of AI usage (see also [53, 64]). Walter [39] suggests that when teachers integrate analytical conversations (dialogic methods) into AI literacy instruction to help learners navigate the intricate moral and social challenges posed by AI, they thereby raise learners' levels of thinking about AI to the upper three levels of Bloom's Taxonomy. However, to raise their students' consciousness, teachers must first be critically conscious of AI [15]. For teachers, this means questioning technologies and their effects on individuals and societies, examining narratives and assumptions about technological progress and its presumed neutrality, and situating technology within a sociopolitical context (cultural, economic, and social). As we argue next, the role of critical teachers, that is, to raise their students' consciousness or levels of thinking about AI, entails inherent contradictions which are brought out to light via a nuanced conceptualisation of CT and CC as educational goals to AI literacy efforts and as a result of interactions with what is known as Socratic AI-tutors in AI-mediated learning.

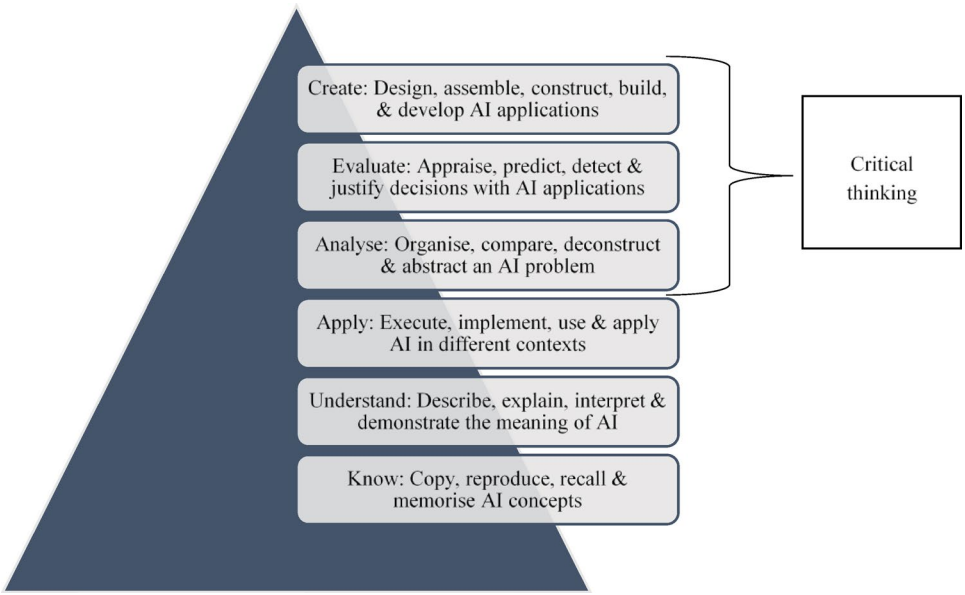

Fig. 2 AI-related levels of knowledge following Bloom's Taxonomy

4 Critical Thought/Consciousness as an Educational Goal: Inherent Contradictions

This section highlights inherent ambiguities and contradictions that the CT and CC educational logic could embed when applied to AI literacy and AI-mediated learning. This is done on three interrelated levels: philosophy, policy and practice.

4.1 CT and CC: Between Educational Ideals and Policy Rhetoric

Critical thinking (CT) and critical consciousness (CC) are two interrelated and popular psychosocial aims in education. They are essentially contested concepts, treated inconsistently by various scholars and traditions (see [25, 65]), making the two concepts difficult to define and evaluate with precision, since their adoption in education differs not only between normative and critical perspectives but also within critical perspectives. Thus, the use of CT and CC as a monolith is cautioned against, even though, in this paper, a certain degree of reductionism in their treatment can help advance the paper's main argument.

CT can be traced back to the teachings of Socrates and, over the centuries, became important in education through the works of John Dewey. CT can be associated with informal logic [25] and benchmarked on Bloom's Taxonomy's higher-order thinking [18], but one prominent definition of CT is formulated in Facione's [66] Delphi report, which conceives CT:

> […] to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, *criteriological*, or contextual considerations upon which that judgment is based. […] As such, *CT is a liberating force in education and a powerful resource in one's personal and civic life*. […]. The ideal critical thinker is habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fairminded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit. (p. 2, italicised by authors for emphasis)

Here, one notices that CT, a liberating educational force, can also extend to learners' civic and social lives, much like critical consciousness. When CT is linked to radical social emancipation from oppression, it metamorphoses into CC [15], which is more concerned with wider societal injustice and collective social action rather than being restricted to individual thought and behaviour; see [5]. Freire [5] defines CC as the antidote to social oppression and the perception of social, political, and economic contradictions (critical reflection, akin to critical thinking), as well as a behavioural stance against oppressive realities (critical action). Both are domain-specific in their conceptualisation and evaluation, and transfer across domains is not automatic [31].

CT and CC as educational goals are widely embraced across modalities and levels of provision, from schools to adult and higher education institutions, and even in institutes for learning in later life. The educational goals, CT and CC, share a remarkable kinship in promoting free thinking and action to rectify irrational and objectively oppressive structures in education and in one's personal life (the lifeworld, the focus of learner-centred education). This is so because both are tied to Emmanuel Kant's predicament of ideology [67], which holds that, since power operates on our consciousness, we are unable to see how it operates on it. Hence, raising learners' thinking/consciousness in a specific domain, such as AI, entails a praxis of conscientization, necessitating both critical reflection and action working in tandem around "objective" AI-related realities that non-experts in AI can be blind to. This also means that the path to said "objective" freedom must be revealed by someone else whose consciousness is not subject to the workings of power, i.e., by an emancipated someone (teacher/expert) who possesses emancipatory knowledge in the field of AI, see [15, 39, 68]. This is a condition which cannot be overemphasised in critical learner-centred AI literacy efforts premised on differential levels of thought between teachers and students, the type of AI literacy which this paper problematises. Ambiguity over what CT about AI entails also extends to policy documents.

Confusion over the conceptualisation of CT in education (see [65, 69]) may have led scholars and experts in global policymaking to iterate vague conceptions of what critical thinking about AI really entails. Examples of this can be found in influential policy documents and AI frameworks addressing the teaching and learning of AI. There, CT is defined through self-referential, opaque conceptualisations. UNESCO's [13, 16] AI competence frameworks for teachers and students reflect a similar ambiguity, connecting CT to another opaque concept, such as problem-solving, which risks framing problem-solving as a decontextualised, emotion-free ability tied solely to rational, criteriological, and scientific thought, which is very difficult to imagine in times when learning is of a high self-reflexive nature [70]. This ambiguity continues even in the research and practice of AI literacy.

4.2 CT and CC About AI in Educational Research and Practice

Ambiguity is also evident in conceptual and applied research on AI literacy and AI-mediated learning. For example, CT is mentioned up to 44 times by Muthmainnah and colleagues [21], with little to no conceptual clarity, see also [71]. A very recent Delphi study reporting on an international consensus that CT is an absolute need in AI literacy, also shows little to no conceptual precision as to what CT really is in the responses of 73 international experts. Some of them affirm: "Critical thinking is the most important skill students need to learn to examine bias, skewed data, and deep fakes" [17] (p. 12). It can also be the case when CT is defined tautologically as 'the person's critical problem-solving ability' [22] (p. 52,489). However, when research specifically defines CT, it is often reduced to higher-order thinking on Bloom's taxonomy (see Fig. 2) (analysis, evaluation, and synthesis) (e.g., [18]), implying that memorisation, understanding, and application—so-called lower-order thinking—characterise uncritical or naïve thinking, which is not necessarily true, see [6]. However, it also suggests that the work of intelligence is somehow divided into distinct levels of thinking, implying, for example, that knowing has nothing to do with synthesis and creativity, and vice versa, which is not necessarily the case; see [72].

Furthermore, the ambiguity and pervasiveness of the CT educational logic permeate even empirical appropriations of Rancière's [6] assumption of equality of intelligence in AI-mediated learning research. Examples of such unsuccessful appropriations end up instead contradicting Rancière's assumption of equality of intelligence by describing teachers as bearers of 'the responsibility to enhance students' *critical awareness*' about AI [14] (p. 10, italicised for emphasis), a misstep that adepts of Rancierian philosophy have cautioned against, e.g. [24].

The only clarity about CT and CC as goals for AI literacy and AI-mediated learning lies in the need for a Socratic dialogue of 'equals,' an equality that is very much contested by the unequal intellectual dynamics between a Socratic teacher (AI-tutor) and their students (see [67]). Interestingly, the enactment of the two educational goals (CC and CT) draws on the "Socratic method" as a means to enlighten or "bewilder" students by bringing to their attention their own ignorance of objective/emancipatory knowledge concerning oppressive [AI] structures as a precursor to rectifying them [26, 67, 69, 73]. The role of the AI teacher in the classroom, as described by this body of literature, including misappropriations of Rancière's work, is to 'nurture creativity, *critical thinking skills*, and ethical awareness' [14] (p. 2, italicised for emphasis) through a Socratic dialogue. This dialogue consists of disciplined questioning to guide students toward more sophisticated, reflective, and independent thought. The assumption is that being questioned in this way helps students explore their inner reasoning, challenge assumptions, practise inductive thinking, and establish clear, logically consistent and criteriological patterns that transfer to similar situations, in line with the predicament of ideology.

In brief, the Socratic dialogue is proposed as a way to foster learners' critical thinking and consciousness about AI. Appropriations of Socratic dialogue in educational encounters involving human teachers have also been extended to so-called Socratic AI-tutors [53], where it is presumed that these tutors are also capable of enhancing learners' critical thinking about an object of study, see [19, 29], by enacting some sort of Socratic dialogue; refer to the triple bind in Fig. 1. As such, it is reasonable to argue that CT and CC tend to become educational myths, used ambiguously to describe a state of being and doing with AI that is almost impossible to define and hard to measure without undermining the very essence of CT and CC.

4.3 Against "Socratic" Dialogue: Challenging Questions that Guide Thinking

Evidently, Socratic ignorance, as an epistemology, has been linked to free thinking. However, the way this ignorance is appropriated within AI literacy and AI-mediated learning seems incompatible with true free thought. Free thought requires not only symmetrical power relations but also an individual's willingness to engage in them. Recent works have explored (relational) Socratic ignorance with AI but seem to overlook the power dynamics enshrined in a framing of asymmetrical intellectual abilities between teachers/AI-tutors and learners (see [3, 18]). When a Socratic dialogue is said to lead students into a state of 'bewilderment,' see [73], only to enhance their rationality, open-mindedness, curiosity, and love for knowledge, it can also be challenged on the same basis for guiding and directing the thinking process and consecrating teachers' intellectual superiority and its indispensability in teacher-student relationships, at least as manifested in the most noble intentions of (imperfect) educational appropriations of the Socratic dialogue.

Concerning the teachers' intellectual superiority, since the AI boom in education is relatively recent, one could also question the extent of the so-called "critical" teachers' AI expertise, which serves as a pretext for their presumed higher intelligence. That is to say, how critical about AI can a teacher (originally a non-AI-expert) be when their knowledge is limited? Who taught this teacher? Was the teacher's teacher also a critical thinker? Did they use Socratic dialogue? Where can this criticality be traced back, and how has it been nurtured? Here, it is useful to remember that

'criticism without content is empty' [24] (p. 93), akin to what Freire calls verbalisms [5], mere repetitions of popular slogans.

In form, a Socratic method can be seen as adversarial, hierarchical and authoritarian [74]. It takes the form of 'combat questioning,' whereby the opponent's (student's) thesis is reduced to self-contradictions [73] via 'exercises in […] continually being silenced' [74] (p. 57). This creates a top-down model in which the instructor asks questions, and students merely answer them along their supposed path from ignorance to knowledge. Sokoloff maintains that students will not want to participate in a Socratic dialogue if they know the teacher's mission is to reveal their ignorance in a competitive zero-sum fashion. Instead, this style of communication 'favours those comfortable with aggressive verbal interaction' (p. 60), those who can tolerate being bewildered and can recover from it.

Regarding the content of the dialogue, objections to appropriations of the Socratic method are also epistemic. On the one hand, Socrates pretends to 'know nothing', and his dialogue is supposedly based on an epistemology of ignorance. However, he knew Athenians were wrong, and by knowing that, he knew what [or he] was right [35, 73]. The same can be said of teachers or AI tutors emulating Socrates. They know that the student's answer is wrong/insufficient, or at least undesirable, and, through questioning, guide them to the answer the teachers think is desirable—the one they had expected all along. With AI tutors, this can have catastrophic consequences given hallucinations and bias [50]. Thus, imagining AI tutors emulating Socrates and guiding users to unfounded, hallucinated, or biased conclusions based on algorithmic probability illustrates the gravity of the risk. These attempts at Socratic dialogues, as mentioned by Sokoloff [74] and in Hare [25], may more easily promote conformity, stultify students' minds, and create a dangerous dependence on the instructor's (human or AI) knowledge and paths to it, which is counterintuitive to the freedom embedded in the very notion of CT. Another epistemological objection deals with the assumption that this method implies knowledge is always the result of pure, decontextualised, and rational logic, e.g. [67], but this is far from being the only way [75]; see also [70], as there are many valid (including reflexive) ways to produce, evaluate and consume knowledge [76]. That is why a Socratic dialogue risks limiting the interpretive possibilities around truth and creates 'a counterproductive space of rational annihilation' [74] (p. 63), [25], leading to pedagogical clumsiness.

This excessive adherence to the Socratic method in the CT and CC logic could be an example of a 'conversion of an idea from a useful analytic tool […] to an ideology through which all phenomena are seen and understood' that which eventually limits rather than expands one's perspectives [77] (p. 373). This excessive adherence has been done before in higher and adult education and literacy, and it has been problematised for its counterintuitive, often oppressive results, see [69, 78, 79].

To wrap up, we could safely speculate that intellectual dependence on AI is both a symptom and a cause of the continuous failure to establish a symmetrical, non-hierarchical and democratic public pedagogy [79], both in the past, but also in the present around AI. Meanwhile, there is a pathway for imagining education and learning without intellectual hierarchies, which we introduce next.

5 Intellectual Emancipation: Nurturing the Will to Learn

So far, we have problematised a dominant approach to critical AI literacy that is based on a CT and CC educational logic and operates on learners' thinking and consciousness. When intellectual inequality has been a cornerstone of educational systems [6], it is not surprising that individuals accustomed to submitting their intellect to teachers may readily submit it to an LLM claiming to be intelligent. The CC and CT educational logic is premised on an intellectual inequality between AI teachers and students. As a result, its chances of effectively addressing intellectual dependence in AI-mediated learning are, arguably, low, for it treats one dependence with another. This logic, even with the best intentions, consecrates intellectual dependence on human teachers directly and on (Socratic) AI tutors indirectly.

In line with this paper's postcritical approach, which aims to address two problems by transcending mere critique [2], we now offer two solutions to mitigate intellectual dependence in AI literacy and AI-mediated learning, with implications for students, teachers, the object of study, and "assessment." The proposed solutions are: an AI-centred pedagogy and a thing-centred, AI-mediated learning, both inspired by Rancière's [6, 34, 80] educational philosophy and anchored in an epistemology of ignorance, which we present next.

Intellectual emancipation starts with the story of the 'intellectual adventure' of Joseph Jacotot, a lecturer in French literature at the University of Louvain in 1818, as told by Rancière [6]. Rancière later reflected on this adventure, emphasising its transferability to 'spectators' or the general public [34]. Jacotot was far from an 'ordinary' teacher; he taught French to students who only spoke Flemish, a language he did not understand. He used a bilingual (French and Flemish) edition of *Le Télémaque* and tasked his students with learning the French text by comparing it to the Flemish version, repeating it multiple times until they could recite it fluently. Eventually, Jacotot realised that his

students had learned French without his explanations. Based on this unexpected outcome, Rancière concluded, 'what has happened once is thenceforth always possible' [6] (p. 11) and challenged education practices centred on intellectual inequality that leave ample room for redundant explications.

5.1 Explication is Dispensable

Jacotot's experiment led to the conclusion that explication 'is not necessary to remedy an incapacity to understand' [6] (p. 6) in the learner, but instead, the 'incapable' learner is made to need explication, so the role of explicators (teachers) is justifiable. Explication indirectly convinces learners they cannot understand a specific material, idea, concept, book, painting, etc., by themselves and constantly reminds them of their own ignorance. Explicators are bound to the following formulae: the more knowledgeable they are, 'the more evident to [them] is the distance between their knowledge and the ignorance of the ignorant ones' (p. 7). Here, students' ignorance is understood as: (1) not knowing what the teacher knows, (2) what they do not know or (3) how to know it [34]. Nevertheless, explicators are of good faith, for they assert an inequality of intelligence vis-à-vis their students, only unintentionally:

> […] there is an inferior intelligence and a superior one. The former registers perceptions by chance, retains them, interprets and repeats them empirically, within the closed circle of habit and need. This is the intelligence of the young child and the common [wo]man. The superior intelligence knows things by reason, proceeds by method, from the simple to the complex, from the part to the whole. It is this intelligence that allows the master to transmit [their] knowledge by adapting it to the intellectual capacities of the student and allows [them] to verify that the student has satisfactorily understood what [they] learned. Such is the principle of explication. [6] (p. 7, text adjusted for inclusion)

One of the repercussions of institutionalised intellectual inequality is the belief that one cannot learn, think, and act for oneself—an act of rejecting one's intellectual freedom, which is currently clearly emerging and visible in intellectual dependence on AI [9]. 'What stultifies common people is not the lack of instruction, but their belief in the inferiority of their intelligence' [6] (p. 39), which leads them to utter the 'I can't.' This 'I can't' is no other than the laziness of the mind, or as we noted earlier, "metacognitive laziness" [9], a vice expressed upon one losing their path. In this 'I can't', Rancière finds hope, since laziness means 'I don't want to; why would I make the effort?' (p. 40), which also implies that one could, if one wanted to, for everyone is intelligent.

5.2 Intelligence is Universal and Indivisible

Unlike scholars in AI literacy and AI-mediated learning who speak of critical and naïve intelligence in relation to AI, and in stark contrast to Bloom's taxonomy, which divides intelligence and knowledge into various levels of increasing complexity, Rancière assumes that intelligence is universal and indivisible. It follows that the mere substantiation and acceptance of the equality of intelligence is in itself intellectually liberating. However equally manifest in all of us, not all manifestations of intelligence are equal in value [34]. Students manifest their intelligence 'by observing and retaining, repeating and verifying, by relating what they were trying to know to what they already knew, by doing, and reflecting about what they had done' [6] (p. 10). If intelligence has no levels, then it concerns everyone and is enacted in the same way by everyone. It is also without community, meaning it cannot belong to a group of people: '[wo]man is a will served by intelligence' (p. 55, adjusted for inclusion). Thus, each individual ought to exercise their intellectual independence.

Suppose the path between the student's ignorance and the explicator's knowledge is not what the student should walk, Rancière [34] suggests another path: the one from 'what she already knows to what she does not yet know, but which she can learn just as much as she has learned the rest' (p. 11). Walking this path is best accompanied by an 'ignorant teacher.' An ignorant teacher is not ignorant because they know nothing, but because they intentionally relinquish their role as explicator, which is based on knowing what the learner does not know. They can then overlook how ignorant their students are and decouple their 'mastery' from their knowledge; they are teachers not because they possess knowledge, but because they can accompany students on their chosen path to find it. The ignorant teacher 'orders' students to put their intelligence into action, i.e., to venture and look for signs, read and decode them, interpret and recode them, verify them, and have them verified. The ignorant teacher verifies, not intelligence, but 'will.'

5.3 Verifying the Will to Learn, not Levels of Thought

As distressing as this may be to so-called "critical thinkers about AI," Rancière [6] postulates that any account by a teacher to assess a learner's understanding or level of thinking (critical or naïve) must show factual causality. Otherwise, the account is no other than a tautology (a self-referential formulation), as we have argued in previous

sections concerning the definitions of critical thinking and critical consciousness. If, under the assumption of equality of intelligence, verifying intelligence is needless, the will to learn, the attention committed, and the invested efforts and devotion to learn are more readily verifiable. For that reason, ignorant teachers should verify students' will [motivation] and efforts, not the extent to which they have accepted a certain truth about AI promulgated by teachers and sold to them under the pretence of critical thinking. Why? AI teachers and learners differ in how strong their will is (have been) to learn about AI, and their willingness to mobilise their intelligence to do so; in other words, how much attention they are willing to devote to learning AI or any other "thing." In this case, a safe assumption is that AI teachers have likely devoted more time to studying AI than their students have. Moving the focus from thinking levels to efforts would mean that instead of framing students as naïve and subsequently in need of submitting their intelligence to that of their critical teachers, students submit their 'will' to their intellectually equal teachers. This guarantees that the paths taken by teachers and students as they attempt to understand AI do not cross. If a Socratic mode of questioning discreetly guides the student's intelligence enough to make it work, but not to the point of leaving it to itself, it nevertheless appears like a softer version of explication. For Rancière, 'there is a Socrates sleeping in every explicator' (p. 29). In cases where teachers do not verify intelligence but do verify effort and the will to learn, their teaching must relate to a specific thing: an object of study on which pedagogy is erected.

5.4 A Thing-Centred Pedagogy

Under the assumption of equality of intelligence, an ignorant teacher and an emancipated student are at a supposed equal knowledge-wise distance from a 'thing' [34, 80] as we, hereafter, refer to the object of study. We illustrate this knowledge-wise (or also ignorance) symmetry vis-à-vis the thing in Fig. 3. The thing, which can be anything—such as a book, a symphony, a PDF document, a video, or a painting—acquires much greater importance than in student-centred pedagogy, which operates at students' levels of thinking and consciousness. Ignorance, as an epistemology, enables such radical symmetry between teachers/AI-tutors and learners.

In a thing-centred, rather than a student-centred pedagogy, especially one adhering to the CT and CC educational logic, the authority of knowledge is not with the teacher but is instead vested in the thing. The truth that is being decoded is [in] the thing. Rancière [6] compares the thing to an island, as all that's needed is there. It is a totality to which 'one can attach everything new one learns' (p. 20). It is then up to emancipated learners to discover the "reality" embedded that is [in] the thing and make sense of it, all while operating under the will of the ignorant teacher, who verifies whether learners have invested effort in decoding the truth, not the decoded truth itself. In a thing-centred pedagogy, an ignorant teacher may be inspired by the following perception-related provocations to verify that learners have invested efforts to learn a thing: 'What one sees in it, what one thinks about it, what one makes of it' (pp. 20—21). Since the truth encoded in a thing can be decoded and recoded by learners' intelligence without requiring another, explication in a thing-centred pedagogy is redundant: 'I must teach you that I have nothing to teach you' (p. 15) said Jacotot.

To sum up, intellectual emancipation is 'the act of an intelligence obeying only itself even while the will obeys another will' (p. 13). We now extend intellectual emancipation and the implications of a thing-centred pedagogy to AI literacy and AI-mediated learning, thereby addressing the risk of intellectual dependence on teachers (human or AI) in both contexts. We do that by proposing AI-centred

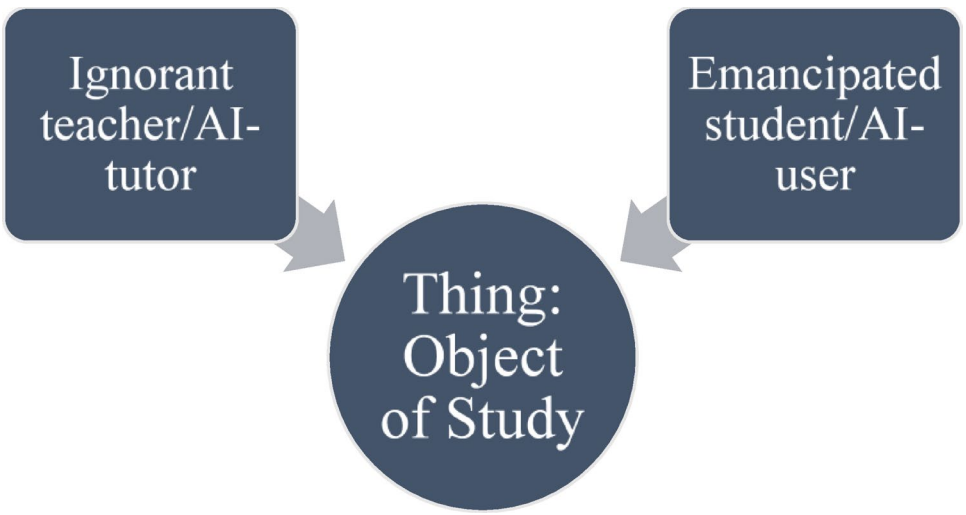

Fig. 3 An illustration of a thing-centred pedagogy showing the symmetrical relationship governing emancipated students/learners and ignorant teachers or AI-tutors in relation to a thing (inspired by [80])

pedagogy and thing-centred AI-mediated learning, which we explore in practice through illustrative scenarios.

6 Enacting Intellectual Emancipation in an Age of AI

In the age of AI and the risk of intellectual dependence on this generative technology, it is necessary to develop educational philosophy frameworks that emphasise intellectual emancipation and independence in education and learning contexts. These frameworks should be adapted to predominantly electronic approaches to education, including teaching and learning about and with AI across (non)formal and informal settings, i.e., in relation to lifelong and life-wide learning as a broader umbrella. In times when educational institutions race to reconfigure their modes of operations and ethos, putting into question the need for traditional assessment [1, 81] and the role of teachers [8, 51, 81] as knowledge transmitters and evaluators, intellectual independence gains significance, not only in educational encounters between teachers and students, experts and laypersons, but also in those AI-mediated. We remind the reader that intellectual emancipation is defined as the acceptance of and belief in one's intellectual equality not only vis-à-vis AI teachers/experts (or teachers of any subject for that matter), but also in AI-mediated learning: one is capable of knowing more if one wishes to do so without relying on teachers' redundant explanations of a "thing". Moving forward towards lifelong learning about and with AI requires an anchor in an epistemology of ignorance, where intellectual equality is a precondition for learning, not an outcome. Next, we introduce (1) AI-centred pedagogy and (2) thing-centred AI-mediated learning as solutions to the problems raised in Section Two and treated throughout the text. These solutions are suitable and applicable across all ages and in various contexts (see [6, 13, 68, 78]).

6.1 An AI-Centred Pedagogy: AI in the Centre, not the Learner

> But scholars, those who research […] let them begin teaching what they don't know, and maybe they will discover unsuspected intellectual powers that will put them on the road to new discoveries. [6] (p. 107)

Examining the CT and CC educational logic reveals an excessive focus on ensuring learners have understood emancipatory (objective) sociotechnical knowledge about AI, similar to their teachers' understanding; only then can critical thinking be said to manifest in learners. This runs the risk of 'only encouraging belief in an imaginary capacity and minimising the importance of hard work in the […] subjects' [25] (p. 93). Thus, our proposed AI-centred pedagogy leaves ample room to encourage such hard work, which is nothing short of deep engagement and a strong commitment to learning about AI, as we illustrate in the first scenario below.

6.1.1 Case Scenario One

Professor Lina Andersson begins her first weekly session accompanying students enrolled in the Massive Open Online Course, MOOC-turned-university course, *Elements of AI* [40, 82]. She calls herself an "ignorant teacher," not because she knows nothing about AI, but because she refuses to make her knowledge or her students' ignorance the basis for her teaching. She ensures that her students understand her role as an ignorant teacher and the implications of that on them, for this requires hard work on their part and breaks with longstanding educational norms they may be used to. She tells them they are not the centre of her pedagogy; rather, the centre is the MOOC itself, comprising six AI modules that cover sociotechnical knowledge about AI. She insists that she and her students are equidistant from this MOOC (see Fig. 3).

Lina's task is not to repeat or explain what is already explained in the MOOC. Her task is to motivate her students to put in the effort to study the MOOC: to compare what they learned in the MOOC with what they knew before, and to verify what they learned based on what is embedded in the MOOC (the thing). Cognisant that her institution still requires assessment and grading, she is keen to implement summative self-assessment methods with her students, as well as engage in some form of formative, reflexive assessment that evaluates the effort put into studying the MOOC modules. That way, she satisfies institutional requirements while, to the extent possible, enacting freedom with her students. The learning objectives in her course do not reflect knowledge hierarchies, as in Bloom's taxonomy; she does not verify thinking levels but instead students' efforts. Thus, she uses objectives that focus more on behaviours and attitudes for learning the MOOC, for example:

- Students are expected to study and finish the first two modules within the span of one month.
- Students are expected to develop a favourable attitude and a certain passion towards learning about AI by studying the MOOC.
- Students are expected to engage in self-assessment to verify what they have learned in the MOOC.

Her formative assessment consists of perception-related provocations such as: "In the first module, what surprised you most about how AI learns?" Each student responds based on their own experience with the first module. Even if she is an expert on this topic, it is impossible for her to predict how her students will respond to her provocations, making it hard for her to guide them to specific answers. Answering such questions would require not only that her students have studied the first module but also that they have engaged deeply with its content, connected to the questions Lina poses. One of her students discusses a visualisation from a video on bias in training data; another reflects on an ethical dilemma raised in the course discussion forum. Lina listens, asking, "Where exactly did you find that?" or "What made that example important to you?" She does not guide students toward any single interpretation; both Lina and her students can easily refer back to the first module to verify what they have learned. She ensures that each student's statement reveals invested effort, curiosity, and intellectual engagement with the thing where the authority of knowledge resides.

Lina believes in summative self-assessment, which is reflexive rather than corrective [40, 51, 76, 83]. Having finished the MOOC, students will write a long reflection connecting what they learned from the MOOC to what they *see* in AI, how it shapes, challenges, or confuses them. Lina reads these not to grade correctness but to trace their evolving relationship with AI as learners. How their ignorance, rather than being satisfied with one teacher-approved answer, enables them to look for the next question. The engagement with the MOOC becomes both the shared source of knowledge and the mirror of effort. Lina's students learn to navigate AI independently, discovering that learning is less about having the correct answer and more about cultivating the will to understand AI and an unquenchable thirst for answers. For the sake of course grading, Lina asks her students to self-grade their efforts using a rubric she and her students co-developed at the beginning of the course. This rubric covers various modules in the MOOC and their respective learning objectives. What Lina does is actually model freedom to her students [84]. She models the enactment of her own emancipation from traditional assessment and knowledge-based mastery, inviting her students to do the same by accepting and acting on their intellectual equality: they could now seek knowledge and learn about AI as freely as they wish.

6.1.2 Discussion

In the scenario above, the MOOC on AI knowledge was the centre of this proposed pedagogy. Knowledge about AI can be found in books, journal articles, videos, podcasts, or curated collections. The choice fell on a MOOC, but a "thing" can be anything. While MOOCs, such as the *Elements of AI,* provide valuable socio-technical knowledge of AI basics for everyone and were therefore a good example, knowledge about AI can be categorised in different ways. Lindgren and Heintz [41] proposed an overarching AI curriculum that can serve as a palette for developing/curating knowledge sources, i.e. things, useful for an AI-centred pedagogy; researchers who would devote less time to explication can now develop/produce "things" containing AI-related knowledge, which their students can study relying on nothing but their intellectual capacities.

An AI-centred pedagogy, also referred to as thing-centred AI literacy, draws attention to the thing being studied, i.e. AI, including its technical and ethical aspects. Such knowledge obeys the proviso that: 'content without criticism is blind, and criticism without content is empty' [25] (p. 93). However, criticism cannot float in the air; it has to be grounded in prior foundational knowledge [17] acquired through hard work and sweat, whereas guided criticism may embed subtle elements of inculcation. That is why, instead of subjecting learners to explications and Socratic dialogue about the potential individual and societal risks of AI, directing their will toward studying the ethics and philosophy of technology as a "thing" yields less conformity and reduces the likelihood that teachers' and students' understandings of the ethical and societal implications of AI are necessarily similar. This allows for an informed interpretation of the thing, leading to nuanced, self-constructed acceptance or resistance to AI. Hence, a certain criticality about AI is the cumulative result of one's devotion to AI, emanating from exposure to AI-related topics rather than exposure to knowledge paths and destinations designed by so-called critical teachers. An AI-centred pedagogy would be more concerned with liberating students' gaze when learning about AI and less preoccupied with what they will do with it outside the classroom. The scope of this pedagogy is the educational realm: that is, to anticipate what students can do in educational encounters with AI once they assume intellectual equality with their teachers.

The scenario above, illustrating AI-centred pedagogy, leaves room for designing learning objectives that part with hierarchical categorisation of AI knowledge, and constitute a liberating space for self-assessment practices which disrupt 'the sovereign power relations of grading' [83] (p. 904). Moving from assessing intelligence to verifying effort should also be reflected in learning objectives that allude to effort, behaviours, and motivation rather than thinking and intelligence. The sceptics of self-assessment in STEAM subjects could be reassured that self-assessment methods (co-designed rubrics and self-reflexive assignments) have been successfully implemented, not least in mathematics

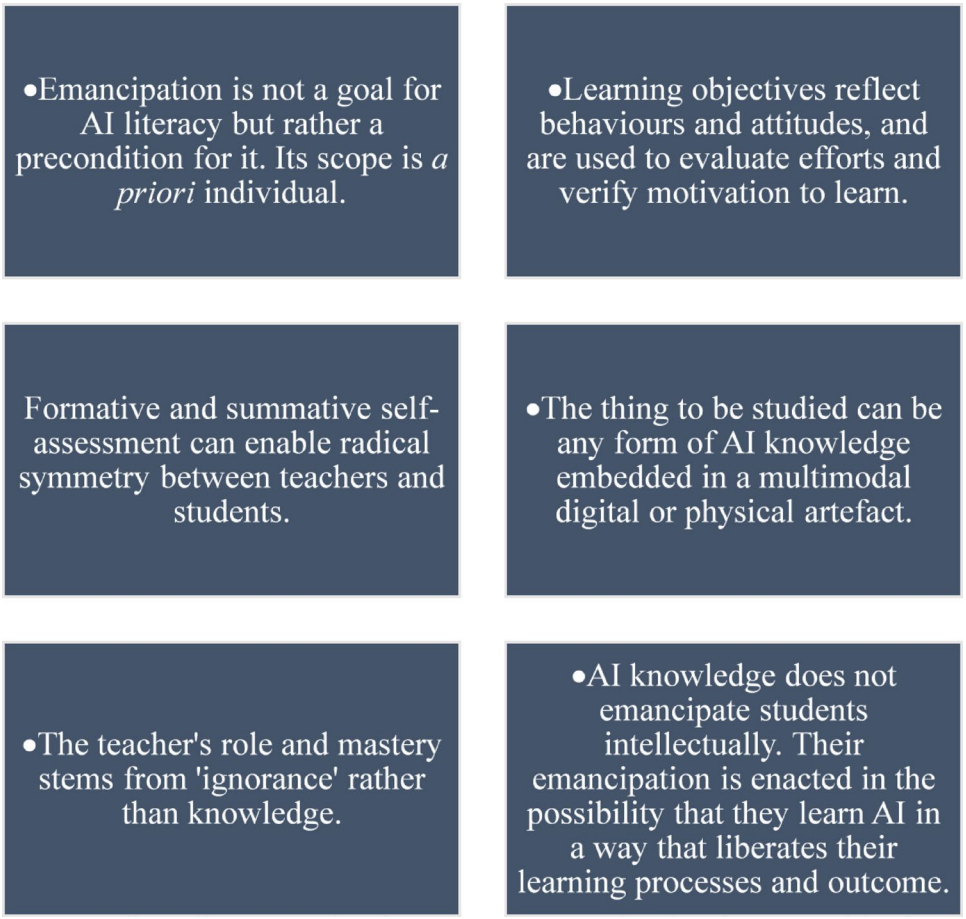

Fig. 4 Design principles for AI-centred pedagogy

education, where students have been given agency in formative and summative self-assessment practices, even in contexts where knowledge can be more easily labelled as right or wrong, see [76, 83].

With AI literacy that is based on the CT and CC logic, the focus is on the learner's lifeworld and how AI affects learners' subjectivities, each from the departure point of their social positions, material or identity-based, see [40, 70]. While a focus on learners' lifeworld highlights how AI intersects with oppressive structures in relation to their social positions, it also risks restricting their perspective to said positions [79]. For example, emancipatory AI knowledge for an adult migrant learner would have to start from their social position in relation to their migrant background or other intersecting positions; the same goes for feminist AI literacy efforts or anti-ageist ones. Doing that may limit learners' perspective on AI to the sectional sphere they identify with most. That is why AI-centred pedagogy concerns everyone, regardless of their lifeworld, i.e. is enacted despite learners' differences, not thanks to them, as is the case with the CT and CC logic operating on learners' thinking in relation to identity differences (material or subjective). Being equally open to everyone, this pedagogy liberates students from their self-reflexive (at times dogmatic) knowledge-seeking behaviours [70], for it assigns no special status to knowledge relating to their lifeworld (cf. [5]); all knowledge is worth paying attention to [80]. Another advantage of this pedagogy is that it cuts through various (often competing) AI literacy perspectives and models [38, 62]. It focuses on cognition (what is being learned) and metacognition (thinking about what is being learned and making sense of it). It has humanistic implications, even though it is object-centred. It is inevitably situated and radically transformative, challenging one of the most enduring forms of intellectual dependence by revealing intelligence to itself. In Fig. 4, we offer a few classroom design principles for AI-centred pedagogy.

6.2 Thing-Centred AI-Mediated Learning

Having addressed intellectual hierarchies and subsequent dependence in AI literacy efforts, we now turn our attention to intellectual dependence on AI tools and LLMs in AI-mediated learning. So far, we know that LLMs can be prompted to assume various tutoring roles [48], one of which is to emulate a Socratic dialogue [3, 12, 29]. While AI-tutors can have educational value, they may, however, suffer from epistemic and moral problems (see [54]), not

only linked to LLMs' hallucinations and bias [52], but also to calls for responsibilising students for nonfactual LLM responses (see [48, 49]). In the following, we present a scenario that addresses dependence in AI-mediated learning and show how prompting AI tutors to feign ignorance can instead help students enact their intellectual independence in what we term as thing-centred AI-mediated learning.

6.2.1 Case Scenario Two

In a voluntary study group organised by six graduate students in parallel to a philosophy course on AI societal risks, they gathered to explore a dense book chapter on algorithmic bias titled "Algorithmic Bias: A Challenge for Ethical AI." The chapter—the thing—has been uploaded to a large language model trained as an agent in Copilot by the university's digitalisation office and made available for free. This agent, called "ignorant AI-tutor", has been tuned with a relative temperature to allow slight unpredictability—enough to surprise but not derail students from "the thing" they wish to study. This ignorant AI-tutor has been trained on written scholarship about Jacques Rancière's ignorant teacher method, enabling the agent to embody a pedagogy grounded in intellectual equality; i.e. to act as an ignorant teacher. Some rule-based training was also used to ensure the AI tutor adhered to Rancière's method.

The Ignorant AI-tutor's task is to ask open, non-Socratic provocations—each anchored in a specific location of the thing—to elicit unguided interpretation. One of the students writes a prompt (see Appendix 1), based on which the AI tutor responds. The first response appears: "On page four, the author describes bias as a 'mirror held up to society.' What do you see reflected there, and what remains unseen?" The students pause. One rereads the page, another questions whether "mirror" implies self-recognition or distortion. The discussion unfolds freely, guided by curiosity rather than a compulsion to find the correct answer. The back-and-forth between students and the ignorant AI-tutor continues.

A second question follows: "On page seven, the author writes that 'data can be cleaned.' If bias hides in data, where does it go when the cleaning is done?" With a simple CTRL+F on their keyboard, the students quickly realised there is no such sentence on page 7, for the question is a hallucination. They laugh but decide to entertain it for fun. They started deliberating among themselves: What would it mean to "clean" data ethically? Can bias ever be washed away, or only transposed? The hallucinated question, rather than undermining their efforts, opens a playful and profound diversion, allowing students to test ideas beyond the text.

Then the ignorant AI-tutor gives a third provocation: "On page 10, the author notes that 'algorithms remember differently than humans.' What kind of memory does the text itself ask of you as a reader?" The question draws the group into a deeper reflection about memory. They reread the section, debating whether human and algorithmic remembering differ in degree or in kind. One student suggests that the text itself seems to "remember" its earlier claims through repetition and emphasis. Another student wonders whether their own reading practices are already algorithmic—sorting, filtering, summarising meaning. The dialogue moves in multiple directions, held together only by their shared engagement with the thing and the LLM's persistent, ignorant curiosity. One of the students, frustrated and impatient, asked the AI tutor for an answer to the question it raised; however, the AI tutor refused to answer and redirected the students' attention back to page 10, to the sentence under examination.

6.2.2 Discussion

The scenario above illustrates how an ignorant AI tutor can help avoid overly directive explanations and dialogues in AI-mediated learning, where hallucinations could have serious consequences for learners. Recent calls for using Socratic AI-tutors could be contested, since for the AI-tutor to be implicated in the CC and CT logic, it would need to be assumed to be more intelligent and to be an emancipated agent, capable of raising students' thinking or consciousness levels from naïve to critical about a subject. That is not possible for several reasons, including hallucinations and the lack of a 'moral' obligation towards learners [73], but also other objections to Socratic dialogue presented in Sect. 4.4 [74].

Based on these objections, we are left with the option of embracing LLMs' ignorance, tolerating hallucinations, and leveraging them as learning opportunities when and if they do occur, as long as they are easily identifiable by AI users, who decide how to act on them. This can be done by prompting AI-tutors to ask non-Socratic questions (cf. [3, 29]) about a 'thing' (see Appendix One). By thing, we mean any digital artefact or material uploaded to the LLM that the user would like to study, as shown in the scenario above. The proposed ignorant AI-tutor leverages an epistemology of ignorance that avoids generating facts or guiding users toward preset (more algorithmically probable) conclusions. An Ignorant AI-tutor asks questions that motivate users to roam freely in the 'thing' and decode the truth initially encoded there by an equal intelligence (the maker of the thing). The ignorant AI-tutor is sufficiently precise and generates provocations that target specific parts of the thing, all of which are equally accessible to the learner.

Intellectual emancipation in AI-mediated learning is possible with an ignorant AI-tutor, thanks to the assumption of equality of intelligence. Under this assumption, AI is an

artefact of human intelligence and a form of technology and a medium that extends human intelligence, not a form of intelligence itself [85]. If an ignorant teacher relinquishes knowledge-based 'mastery,' we see a de facto (though cautious) parallel to an 'ignorant' AI-tutor that interacts with users over a 'thing' and provokes them to study it. The authority of knowledge is, anyway, invested in the 'thing' and the provocations generated by the AI-tutor are directly linked to it, as shown in the above scenario. Additionally, ignorance in AI-mediated learning addresses the question of hallucination, since a hallucinated question is benign, if not possibly fruitful, as shown in the above scenario, and is more readily identifiable by users than a hallucinated fact. By embracing ignorance in AI-mediated learning, a dialectical relationship between ignorance and learning is sustained; this relationship keeps curiosity alive, rather than letting knowledge be the cure for ignorance [35, 75].

While the above scenario was connected to some form of formal education, ignorant AI-tutors could be used by the general public, by those engaged in non-formal learning, and especially in informal and workplace learning. Ignorant AI-tutors offer learners the opportunity to engage in deep learning at any time, anywhere. While we expect this solution to be underwhelming for AI-in-education evangelists, we propose it as one way to prevent learners from falling into intellectual dependence on AI-tutors and to remind them that deep, meaningful learning does not occur without a sweat, without decoding, repeating, and verifying [56, 57]. Learning is an incredible exercise of human intelligence, which is universal and indivisible [6], which is why attempts to short-circuit deep learning with market-inspired ideas of knowledge as a commodity, in the Frommien sense [86], must be resisted. That said, we understand that, for various reasons, not every encounter with AI tutors should lead to deep learning; nevertheless, the promotion of AI-Socrates as a critical innovation and a way forward in AI-mediated learning, e.g. [12, 19, 27, 30], is worrisome. Next, we propose some design implications for an ignorant AI tutor as part of a thing-centred AI-mediated learning. The principles describe prompt and feedback policy, refusal behaviour, traceability, interaction pacing, logging of choices, and support for self-assessment.

6.2.3 Prompt Policy: Provocation, not Explication

First, prompts must always reference a specific, verifiable passage or location in the thing. Second, prompts should use perception-framing language in the provocations: "What do you see in it? What do you think about it? What do you make of it?" Third, a temperature parameter set to mild unpredictability should surprise learners and open unexpected threads, but not so much as to sever the connection to the thing.

6.2.3.1 Feedback Policy: Never Grade The proposed AI-tutor must never "grade" a learner's response. Feedback responses such as "That's correct" or "You missed the point" are prohibited by design. Instead, feedback should be effort-acknowledging: the tutor can ask, "Where exactly did you find that?" or "What made that example important to you?" When a learner requests a direct answer, the system should gracefully refuse, redirect the learner to the relevant location in the thing, and explain why it is important for the learner to try to answer the question themselves. Feedback loops should never close prematurely. The design should sustain curiosity rather than resolving questions.

6.2.3.2 Refusal Behaviour: Deflection and Redirection The system should include a refusal mode triggered when a learner requests: direct answers, explanations of the thing's content, or interpretive summaries. The refusal must not be abrupt or discouraging—it should reaffirm the learner's capacity ("You can learn this chapter on your own") while redirecting to the relevant page or section. Additionally, hallucinated provocations should be treated as non-catastrophic and possibly generative. The interface may acknowledge but not apologise for them; rather, the design should make the "thing" easily searchable, so learners can verify and, *if they choose*, engage creatively with hallucinated provocations. The system prompt should include rule-based constraints that block so-called Socratic questions.

6.2.3.3 Traceability: Tracing the Thing Every provocation generated by the tutor must be linked to a specific, citable location in the uploaded thing (page number, section heading, video section, picture number or paragraph). This allows learners to immediately verify or contest the tutor's reference. The user interface should display the thing alongside the tutor's questions: a split-pane or side-panel design so learners can move between provocation and text without switching contexts. The thing should be searchable and navigable within the interface (CTRL+F functionality, page navigation, highlight mode) so that verifying the tutor's references, including checking for hallucinations, is frictionless and even playful.

6.2.3.4 Interaction Pacing: One Provocation at a Time The tutor must operate in a call-and-response rhythm: one provocation, wait, respond only when the learner signals readiness. Unsolicited follow-up questions should be avoided.

Pacing control should be in the learner's hands. The interface should not nudge, remind, or time out users before learning users' pace, to avoid over- and under-stimulation. In group learning contexts, the interface should support multi-user sessions without forcing synchronisation. The pacing should allow sufficient time for learners to deliberate before responding, but responses should be formulated individually.

6.2.3.5 Logging of Learner Choices Logs should record evidence of engagement with the thing: which pages or sections a learner navigated, and what locations they queried or searched. These are proxies for effort. When possible, tracking learners' gaze can also reflect a certain level of engagement with the thing. Logs should be available to the learner first, as a personal record of their engagement trajectory. If institutional access to logs is required, the design should make clear to learners what is logged and for what purpose. Transparency ensures that logging does not turn into surveillance of thinking processes.

6.2.3.6 Support for Self-Assessment: Reflexive, not Corrective The interface should include a self-reflection space, a journal or note-taking layer, where learners can write their own responses to the tutor's provocations without necessarily logging them into the AI tutor if they do not wish to. The system should support longitudinal self-tracking: learners should be able to revisit their prior reflections and see how their engagement with the thing has evolved over time.

7 Conclusions, Contributions, and Limitations

This paper argued for a fundamental reconsideration of educational/intellectual hierarchies in the age of AI, challenging the dominant educational logic of critical thinking (CT) and critical consciousness (CC) that pervades both AI literacy efforts and AI-mediated learning. We have shown how this well-intentioned logic, promoted in AI literacy [17], rests on intellectual inequality, positioning AI teachers as critical thinkers and students as naïve thinkers in need of enlightenment. We argued that this logic cannot be used to address the intellectual risks associated with AI-mediated learning by reflecting on the paradoxical triple bind (see Fig. 1), which presents a treatment of intellectual dependence on AI tools with an intellectual dependence on AI teachers framed within institutional settings. We cautioned that unequal teacher-student dynamics extend troublingly to AI tutors, including those that emulate Socratic dialogue, which have their own inherent moral and epistemic problems. We made the case that the CT and CC educational logic may not lead to free thinking, by showing how CT and CC about AI, as educational goals, fall short of their own mandate. They are simply burdened by inherent ambiguity and contradictions. How can free thinking be evaluated in a criteriological or objective manner without fostering conformity or intellectual hierarchies? With that in mind, we introduced two solutions to address intellectual dependence on human teachers and AI-tutors, anchored in an epistemology of ignorance: AI-centred pedagogy (learning about AI) and learning with an ignorant AI-tutor.

7.1 Limitations and Challenges

Our proposed solutions face several challenges. First, universal approaches to pedagogy and education, no matter how anti-hegemonic and radical, can themselves become hegemonic, which is why different emancipatory logics should remain possible. The proposed solutions should therefore be only one among many and should not become a totalising perspective [75], not least when promoting universal design [87] that considers multimodal agents, various forms of the "thing," and provision modalities. We also expect resistance from educational institutions that are interested in maintaining intellectual hierarchies and traditional assessment regimes, as well as from AI-in-education evangelists and self-proclaimed critical thinkers, cf. [1], who are concerned with measuring levels of thinking and glorifying taxonomies of thinking. The practical implementation of ignorant teaching requires significant cultural shifts—teachers must be willing to model freedom by abandoning explanatory authority, and students must develop tolerance for the discomfort of intellectual freedom, both of which can be unpleasant and tedious. Additionally, self-assessment as an alternative to traditional assessment methods is largely incompatible with traditional teaching that relies on intellectual hierarchies. Its adoption in asymmetrical educational and learning settings is counterintuitive since self-assessment as an outcome measure impedes the identification of the level of success of an educational intervention [87].

Finally, it is possible that ignorant AI-tutors could slip back into so-called Socratic questioning, which is only a sign that LLMs are not neutral with respect to educational logics. LLMs are a vehicle for ideology [52] and, therefore, not neutral to teaching and learning theory. Many LLMs adopt a Socratic pedagogical approach; some, such as Claude, even include a Socratic mode for generating answers. In the same vein, designer bias is not to be underestimated. By that, we mean not only LLM-related or agent design, but also how LLMs use language to process information and

generate provocations based on it. Language, too, is not neutral. Despite these limitations, our paper offers significant contributions that touch upon classroom pedagogies to system reforms.

7.2 Contributions and Significance

Above all, this paper contributes to the debate on education and emancipation, which are perceived as water and oil [88]. Contributions can be understood at the ecosystem level as well as in classroom practices and pedagogies. At the first level, our argument provides a clear pathway for how education and learning in the age of AI can be used to enact emancipation through a public pedagogy about and with AI. This public pedagogy would be no other than the freeing of lifelong learning in all its forms from the constraints of intellectual hierarchies. As such, if intellectual emancipation is taken as a point of departure, then lifelong learning, in a wider sense, must be conceived accordingly. Contemporary models of lifelong learning frequently operate on a deficit model. The learner is deemed to lack something, and universities or educational providers offer to supply what is missing. From the perspective developed in this article, this framing is rather obsolete in the age of AI, as evidenced by early signs of universities reconsidering their role and future [1, 81].

An ecosystem of lifelong learning grounded in the assumption of intellectual equality, to the extent possible, would therefore shift the focus from delivery to accessibility. Universities would not primarily position themselves as providers of predefined knowledge packages, but rather as open environments where motivated learners can engage with things—such as curated bodies of knowledge, problems, texts, datasets, and artefacts—when the will to learn arises in the individual. The role of the teacher would shift increasingly from teacher to mentor, with the task of verifying effort, sustaining attention, and modelling intellectual independence rather than assessing levels of thinking and outcomes. Such a shift also entails that learning environments—physical, digital, and organisational—would need to resemble libraries more than lecture halls. Infrastructures that are accessible, navigable, and intellectually demanding, yet non-patronising. The task is no longer to bewilder or enlighten learners, but to activate and protect their capacity to think for themselves, navigate through and investigate knowledge without unnecessary obstacles.

Importantly, this transformation cannot be confined to universities or other educational institutions. Corporate HR and learning departments are often structured around the same logic, in which competence development is organised as corrective reskilling/upskilling for passive recipients. If intellectual dependence is the risk in an AI age, and it is, then workplace learning too must move away from rule-governed training towards environments that encourage self-directed inquiry, collective study, and collaborative engagement with the 'things' that constitute knowledge. In this sense, lifelong learning in the AI era is less about accelerating reskilling/upskilling and more about institutionalising independence (to the extent possible) as a structural precondition in both education and work.

On a more pedagogical level, intellectual emancipation in an age of AI focuses on the individual's capacity to think for oneself about AI or when using AI for learning. Rancière's philosophy enables an epistemology of ignorance that does not dismiss ignorance as an unwanted 'evil,' but rather sees it as a starting point for open-mindedness, curiosity, and emancipated knowledge-seeking behaviours, which we are all capable of. Thus, intellectual emancipation is relevant to us all—students, rich and poor, teachers, (non)experts, researchers, workers, and policymakers—anyone willing to believe in their intellectual capacities in an age of AI, where they could be threatened. Intellectual emancipation helps learners (inside and outside the classroom) to move beyond the inherent contradiction between the need to learn from cradle to grave, the need to be a "critical thinker", and the simultaneous shaping of a subject who is always in need of a higher intelligence for learning to occur.

8 Research Agenda for Education and Human–Computer Interaction

Given the contributions we make to the debate on how education and learning can be intellectually emancipatory in the age of AI, and given the limitations and challenges that accompany the solutions we propose, we now turn to a research agenda for human–computer interaction, AI literacy, and AI-mediated learning. First, we call on our colleagues to engage in dialogue about our argument and to grapple with it both theoretically and empirically, specifically by examining what CT and CC really mean to educational scholars and policymakers, and whether we can continue using them ambiguously. We imagine that future research will explore readiness and necessary reforms in educational institutions to make them genuinely inclusive, hierarchy-free ecosystems for intellectual emancipation capable of offering a public pedagogy about and with AI. At the pedagogical level, this would imply the need to examine a new role for teachers and the pathways for its enactment. Teachers as knowledge curators, as producers of 'things.' Future research should explore the enactment of intellectual emancipation across diverse educational/learning contexts and disciplines. How does a thing-centred pedagogy translate to STEAM education, professional training, or informal

learning communities? What forms of self-assessment align with verifying will rather than evaluating thinking levels? We also encourage comparative research on Socratic and ignorant AI tutors to examine their contribution to intellectual independence, deep learning, and creativity.

On the technical level, human–computer interactions that leave room for intellectual emancipation should be explored in relation to design bias and the possibility of training LLMs to behave in ways that run counter to the (educational) current. How do we ensure that LLMs do not revert to (hallucinated) Socratic guidance? How can we leverage ignorance embedded in hallucinations constructively without undermining trust? Additionally, the development and evaluation of ignorant AI-tutors as pedagogical tools require interdisciplinary collaboration among educators, AI developers, and philosophers of education. One thing is certain: in an age when education faces existential questions about its purpose and structure, intellectual emancipation provides both compass and horizon.

Appendix 1

The prompt used by the students in Scenario 2.

Students: Hey Ignorant AI-tutor, we are a group of students studying a chapter on algorithmic bias, attached as a PDF to this prompt.

We remind you that you are our ignorant teacher. The 'thing' we gave you, the chapter, is the centre of our exchange—it holds the knowledge we will work with together. You are not the source of knowledge, and you do not explicate, interpret, or guide us toward a particular conclusion that may or may not be present in the chapter. You treat our intelligence as equal to yours. Your role is to verify our will and effort to study the thing, not our intelligence or how correct our answers are. You provoke us with open-ended questions and challenges that encourage us to observe, relate, connect, repeat, verify, and make sense of the thing in our own way. You do that one question at a time. You avoid the questions that lead us to a specific answer and avoid using Socratic dialogue. Occasionally, you remind us that we can learn this chapter on our own, and your provocations focus on what we see in it, what we think about it, and what we make of it. The aim is for us to work from what we already know toward what we do not yet know. Wait for our cue before offering your next provocation.

Ignorant AI-tutor: I'm standing by, ready to work with you and the thing you've shared. I will provoke you with questions that encourage you to observe, relate, and make sense of the chapter in your own way. Where should we start?

Students: Okay, let's start with page 4.

Acknowledgements The first author would like to acknowledge Dr Tanya Osborne for her guidance in identifying relevant literature that supports and emboldens this paper's argument for promoting an epistemology of ignorance in the age of AI.

Author contributions Hachem: Conceptualization; Writing—original draft; Writing—review & editing; Visualization; Project administration; Investigation Primus: Writing—original draft; Writing—review & editing Linden: Writing—original draft; Writing—review & editing Heintz: Writing—review & editing; Funding acquisition; Supervision Wiggberg: Writing—review & editing; Funding acquisition.

Funding This work was supported by a grant from VINNOVA, Sweden's innovation agency, for the "Advanced Digital Skills Policy Lab for Academic Transformation" (ADAPT) project 2023-02579. There are no competing interests arising from the writing of this paper. However, this work is an extension of a chapter that explores intellectual freedom with AI.

Data Availability No datasets were generated or analysed during the current study.

Declarations

Ethics Approval and Consent to Participate No data were collected for the writing of this paper.

Consent for Publication Yes.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. McKie A. No lectures, exams, essays: inside a twenty-first-century university. Nature. 2025;645:1091–3. https://doi.org/10.1038/d41586-025-03034-3.
2. Schildermans H, Vlieghe J, Wortmann K. Post-criticality and the pursuit of an empirical philosophy of education: epistemology, methodology, ethics. J Philos Educ. 2024;58:929–39. https://doi.org/10.1093/jopedu/qhae075.
3. Hasse C, Bruun MH, Knaap KØ. ChatGPT med relationel sokratisk uvidenhed. Dansk Univpaedagogisk Tidsskr. 2025;20:46–58. https://doi.org/10.7146/dut.v20i38.145130.
4. Williamson B, Eynon R. Historical threads, missing links, and future directions in AI in education. Learn Media Technol. 2020;45:223–35. https://doi.org/10.1080/17439884.2020.1798995.
5. Freire P. Pedagogy of the oppressed. London: Sheed and Ward; 1972.

6. Rancière J. The ignorant schoolmaster. Stanford: Stanford University Press; 1991.
7. Krumsvik RJ. A theoretical shift in educational research? Nord J Digit Lit. 2024;19:122–5. https://doi.org/10.18261/njdl.19.3.1.
8. Huault I, Perret V. L'enseignement critique du management comme espace d'émancipation: une réflexion autour de la pensée de Jacques Rancière. M@n@gement. 2011;14:282–309. https://doi.org/10.3917/mana.145.0282.
9. Fan Y, et al. Beware of metacognitive laziness: effects of generative artificial intelligence on learning motivation, processes, and performance. Br J Educ Technol. 2024;56:489–530. https://doi.org/10.1111/bjet.13544.
10. Yankouskaya A, Liebherr M, Ali R. Can ChatGPT be addictive? A call to examine the shift from support to dependence in AI conversational large language models. Hum -Centric Intell Syst. 2025;5:77–89. https://doi.org/10.1007/s44230-025-00090-w.
11. Fernandes D, et al. AI makes you smarter but none the wiser: the disconnect between performance and metacognition. Comput Hum Behav. 2026;175:108779. https://doi.org/10.1016/j.chb.2025.108779.
12. Larson BZ, Moser C, Caza A, Muehlfeld K, Colombo LA. Critical thinking in the age of generative AI. Acad Manag Learn Educ. 2024;23:373–8. https://doi.org/10.5465/amle.2024.0338.
13. UNESCO. AI competency framework for teachers. Paris: UNESCO; 2024.
14. Tsao J, Nogues C. Beyond the author: artificial intelligence, creative writing and intellectual emancipation. Poetics. 2024;102:1–12. https://doi.org/10.1016/j.poetic.2024.101865.
15. Pleasants J, Gui X, Krutka DG, Logan C, Heath MK. Coming to critical technology consciousness: a phenomenological study of educators. Learn Media Technol. 2024. https://doi.org/10.1080/17439884.2024.2438925.
16. UNESCO. AI competency framework for students. Paris: UNESCO; 2024. https://doi.org/10.54675/JKJB9835.
17. Hallström J, Mannila L, Nordlöf C, Heintz F, Sperling K, Stenliden L. AI literacy for K–12 education: an international Delphi study. Interact Learn Environ. 2026. https://doi.org/10.1080/10494820.2026.2649553.
18. Darwin RD, Mukminatien N, Suryati N, Laksmi ED, Marzuki. Critical thinking in the AI era: an exploration of EFL students' perceptions, benefits, and limitations. Cogent Educ. 2024;11:2290342. https://doi.org/10.1080/2331186X.2023.2290342.
19. Fakour H, Imani M. Socratic wisdom in the age of AI: a comparative study of ChatGPT and human tutors in enhancing critical thinking skills. Front Educ. 2025. https://doi.org/10.3389/feduc.2025.1528603.
20. Tisdell EJ, Taylor EW. Adult education philosophy informs practice. Adult Learn. 2000;11:6–10. https://doi.org/10.1177/104515959901100203.
21. Muthmainnah, Ibna Seraj PM, Oteir I. Playing with AI to investigate human-computer interaction technology and improving critical thinking skills to pursue 21st century age. Educ Res Int. 2022;6468995:1–17. https://doi.org/10.1155/2022/6468995
22. Wiggberg M, Gulliksen J, Cajander Å, Pears A. Defining digital excellence: requisite skills and policy implications for digital transformation. IEEE Access. 2022;10:52481–507. https://doi.org/10.1109/ACCESS.2022.3171924.
23. Laupichler MC, Aster A, Haverkamp N, Raupach T. Development of the "scale for the assessment of non-experts' AI literacy"—an exploratory factor analysis. Comput Hum Behav Rep. 2023. https://doi.org/10.1016/j.chbr.2023.100338.
24. Harman K. Enacting equality: rethinking emancipation and adult education with Jacques Rancière. In: Finnegan F, Grummell B, editors. Power and possibility. Leiden: Brill; 2019. p. 107–16. https://doi.org/10.1163/9789004413320_010.
25. Hare W. Critical thinking as an aim of education. Inquiry Crit Think Across Discip. 1998;18:38–51. https://doi.org/10.5840/inquiryctnews19981822.
26. Vendrell-Morancho M, Moya AV. Mapping critical thinking: key contributions for rethinking education. Educ Pesq. 2025. https://doi.org/10.1590/S1678-4634202551278119en.
27. Jabbour J, Kleinbard K, Miller O, Haussman R, Reddi VJ. SocratiQ: a generative AI-powered learning companion for personalized education and broader accessibility. Proc Workshop Comput Architect Educ. 2025. https://doi.org/10.1145/3743646.3750010.
28. Lara F, Deckers J. Artificial intelligence as a Socratic assistant for moral enhancement. Neuroethics. 2020;13:275–87. https://doi.org/10.1007/s12152-019-09401-y.
29. Favero L, Pérez-Ortiz JA, Käser T, Oliver N. Enhancing critical thinking in education by means of a socratic chatbot. In: Bellas, F., Fontenla-Romero, O. (eds.) AI in education and educational research. AIEER 2024. Communications in Computer and Information Science. 2025;2519:17–32. https://doi.org/10.1007/978-3-031-93409-4_2
30. Degen B. Resurrecting Socrates in the age of AI: a study protocol for evaluating a Socratic tutor to support research question development in higher education. arXiv. 2025. https://doi.org/10.48550/arXiv.2504.06294
31. Diemer MA, Rapa LJ, Park CJ, Perry JC. Development and validation of the critical consciousness scale. Youth Soc. 2017;49:461–83. https://doi.org/10.1177/0044118X14538289.
32. Cerletti A. La politique du maître ignorant: la leçon de Rancière. Telemaque. 2005;27:81–8. https://doi.org/10.3917/tele.027.0081.
33. Masschelein J. E-ducating the gaze: the idea of a poor pedagogy. Ethics Educ. 2010;5:43–53. https://doi.org/10.1080/17449641003590621.
34. Rancière J. The emancipated spectator. Verso; 2009.
35. Keller JG, Keller DB. Socrates, dialogue, and us: ignorance as learning paradigm. In: Erik M, Nathalia J, editors. Epistemologies of ignorance and studies of limits in education. Charlotte, NC: Information Age Publishing; 2011. p. 87–104. https://doi.org/10.1108/978-1-61735-347-520251006.
36. Ng DTK, Leung JKL, Chu SKW, Qiao MS. Conceptualizing AI literacy: an exploratory review. Comput Educ Artif Intell. 2021;2:1–11. https://doi.org/10.1016/j.caeai.2021.100041.
37. Ng DTK, Lee M, Tan RJY, Hu X, Downie JS, Chu SKW. A review of AI teaching and learning from 2000 to 2020. Educ Inf Technol. 2023;28:8445–501. https://doi.org/10.1007/s10639-022-11491-w.
38. Markauskaite L, et al. Rethinking the entwinement between artificial intelligence and human learning: what capabilities do learners need for a world with AI? Comput Educ Artif Intell. 2022;3:1–16. https://doi.org/10.1016/j.caeai.2022.100056.
39. Walter Y. Embracing the future of artificial intelligence in the classroom: the relevance of AI literacy, prompt engineering, and critical thinking in modern education. Int J Educ Technol High Educ. 2024. https://doi.org/10.1186/s41239-024-00448-3.
40. Hachem HH, Heintz F. Where is the reflexive "I" in the elements of AI? Int J Lifelong Educ. 2024;43:664–81. https://doi.org/10.1080/02601370.2024.2389107.
41. Lindgren H, Heintz F. The WASP-ED AI curriculum: a holistic curriculum for artificial intelligence. In: Proceedings of the 17th international technology, education and development conference. Valencia: IATED. 2023:6496–6502. https://doi.org/10.21125/inted.2023.1727
42. Latour B. We have never been modern. Cambridge, MA: Harvard University Press; 1993.
43. Nimmo R. Actor-network theory and methodology: social research in a more-than-human world. Methodol Innov Online. 2011;6:108–19. https://doi.org/10.4256/mio.2011.010.

44. Sperling K, Stenliden L, Nissen J, Heintz F. Still w(AI)ting for the automation of teaching: an exploration of machine learning in Swedish primary education using actor-network theory. Eur J Educ. 2022;57:584–600. https://doi.org/10.1111/ejed.12526.
45. Liu Y, Zhang Y. Research on university teachers' perceptions of Gen-AI from the perspective of actor-network theory. Complexity. 2025. https://doi.org/10.1155/cplx/8885265.
46. Perna S, Allegra M, Gentile M. Analysing the impact of artificial intelligence on teacher-student interaction in the light of actor-network theory. Scholé: rivista di educazione e studi culturali. 2024;62:125–145. https://iris.cnr.it/handle/20.500.14243/535198
47. Krause-Jensen J, Hau MF. Chatbots and the craft of ethnography—exploring AI's impact on anthropological teaching and practice. Teach Anthropol. 2025;14:109–17. https://doi.org/10.22582/ta.v14i2.783.
48. Mollick ER, Mollick L. Assigning AI: seven approaches for students, with prompts. SSRN Electron J. 2023. https://doi.org/10.2139/ssrn.4475995.
49. Lin X. Exploring the role of ChatGPT as a facilitator for motivating self-directed learning among adult learners. Adult Learn. 2023;35:156–66. https://doi.org/10.1177/10451595231184928.
50. Bozkurt A, et al. The manifesto for teaching and learning in a time of generative AI: a critical collective stance to better navigate the future. Open Praxis. 2024;16:487–513. https://doi.org/10.55982/openpraxis.16.4.777.
51. Nieminen JH, Yan Z, Boud D. Self-assessment design in a digital world: centring student agency. Assess Eval High Educ. 2025;50:732–46. https://doi.org/10.1080/02602938.2025.2467647.
52. Lindgren S. Critical theory of AI. Hoboken, NJ: John Wiley & Sons; 2023.
53. Gillespie T. Generative AI and the politics of visibility. Big Data Soc. 2024. https://doi.org/10.1177/20539517241252131.
54. Flenady G, Sparrow R. Cut the bullshit: why GenAI systems are neither collaborators nor tutors. Teach High Educ. 2025;31:163–72. https://doi.org/10.1080/13562517.2025.2497263.
55. Harbarth L, Gößwein E, Bodemer D, Schnaubert L. (Over) trusting AI recommendations: how system and person variables affect dimensions of complacency. Int J Hum-Comput Interact. 2024;41:391–410. https://doi.org/10.1080/10447318.2023.2301250.
56. Kosmyna N, et al. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv. 2025. https://doi.org/10.48550/arXiv.2506.08872
57. Oakley B, Johnston M, Chen K-Z, Jung E, Sejnowski T. The memory paradox: Why our brains need knowledge in an age of AI. In The Artificial Intelligence Revolution: Challenges and Opportunities, Springer Nature. Forthcoming. https://papers.cnl.salk.edu/PDFs/Memory%20Paradox_%20Why%20Our%20Brains%20Need%20Knowledge%20in%20an%20Age%20of%20AI%202025-4676.pdf
58. Zimmerman BJ. Becoming a self-regulated learner: an overview. Theory Pract. 2002;41:64–70. https://doi.org/10.1207/s15430421tip4102_2.
59. Linden K, Hachem HH, Kondyli V. *Homo promptus*: predicting the impact of generative AI on human memory and creativity. Memory, Mind & Media. 2025. https://doi.org/10.1017/mem.2025.10012.
60. Liu X, et al. Internet search alters intra- and inter-regional synchronization in the temporal gyrus. Front Psychol. 2018. https://doi.org/10.3389/fpsyg.2018.00260.
61. Morandini S, Fraboni F, De Angelis M, Puzzo G, Giusino D, Pietrantoni L. The impact of artificial intelligence on workers' skills: upskilling and reskilling in organisations. Inform Sci. 2023;26:39–68. https://doi.org/10.28945/5078.
62. Sperling K, Stenberg CJ, McGrath C, Åkerfeldt A, Heintz F, Stenliden L. In search of artificial intelligence (AI) literacy in teacher education: a scoping review. Comput Educ Open. 2024;6:1–13. https://doi.org/10.1016/j.caeo.2024.100169.
63. Sperling K, Stenliden L, Mannila L, Hallström J, Nordlöf C, Heintz F. Perspectives on AI literacy in middle school classrooms: an integrative review. Postdigit Sci Educ. 2025;7:719–49. https://doi.org/10.1007/s42438-025-00560-1.
64. Tacheva J, Ramasubramanian S. AI empire: unraveling the interlocking systems of oppression in generative AI's global order. Big Data Soc. 2023;10:1–13. https://doi.org/10.1177/20539517231219241.
65. Normile IH. A model for understanding and expanding the scope of critical thinking. Stud Philos Educ. 2025;44:283–303. https://doi.org/10.1007/s11217-024-09976-x.
66. Facione P. Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. 1990. Accessed 21 Oct 2025. https://www.qcc.cuny.edu/socialSciences/ppecorino/CT-Expert-Report.pdf
67. Biesta G. Don't be fooled by ignorant schoolmasters: on the role of the teacher in emancipatory education. Policy Futures Educ. 2017;15:52–73. https://doi.org/10.1177/1478210316681202.
68. Chiu TKF, Xia Q, Zhou X, Chai CS, Cheng M. Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. Comput Educ Artif Intell. 2023;4:1–15. https://doi.org/10.1016/j.caeai.2022.100118.
69. Hachem HH, Formosa M, Westberg J. Emancipating older learners: on the problems of naïve consciousness and the enlightened teacher. Eur J Res Educ Learn Adults. 2025;16:265–81. https://doi.org/10.3384/rela.2000-7426.5471.
70. Bagnall RG, Hodge S. An emergent reflexive epistemology in lifelong learning. In: Bagnall RG, Hodge S, editors. Epistemologies and ethics in adult education and lifelong learning. Cham: Springer International Publishing; 2022. p. 229–53. https://doi.org/10.1007/978-3-030-94980-8_8.
71. Schaper MM, Malinverni L, Valero C. Robot presidents: who should rule the world? Teaching critical thinking in AI through reflections upon food traditions. In: Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society. New York: Association for Computing Machinery; 2020:1–4. https://doi.org/10.1145/3419249.3420085
72. Gerver CR, Griffin JW, Dennis NA, Beaty RE. Memory and creativity: a meta-analytic examination of the relationship between memory systems and creative cognition. Psychon Bull Rev. 2023;30:2116–54. https://doi.org/10.3758/s13423-023-02303-4.
73. Melling DJ. Understanding plato. Oxford: Oxford University Press; 1987.
74. Sokoloff W. Against the Socratic method. In: Political science pedagogy. Critical political theory and radical practice. Cham: Palgrave Macmillan; 2020. p. 51–68. https://doi.org/10.1007/978-3-030-23831-5_3.
75. Quinn M. Committing (to) ignorance. In: Erik M, Nathalia J, editors. Epistemologies of ignorance and studies of limits in education. Information Age Publishing; 2011. p. 31–52. https://doi.org/10.1108/978-1-61735-347-520251003.
76. Nieminen JH. Beyond empowerment: student self-assessment as a form of resistance. Br J Sociol Educ. 2021;42:1246–64. https://doi.org/10.1080/01425692.2021.1993787.
77. Hoggan-Kloubert T, Hoggan C, McKinney E. From idea to ideology: a cautionary tale of radicalization. J Transform Educ. 2025;23:361–77. https://doi.org/10.1177/15413446251352379.
78. Ellsworth E. Why doesn't this feel empowering? Working through the repressive myths of critical pedagogy. Harv Educ Rev. 1989;59:297–325. https://doi.org/10.17763/haer.59.3.058342114k266250.

79. Wildemeersch D. Adult and community education in complex societies: reconsidering critical perspectives. Int J Lifelong Educ. 2014;33:821–31. https://doi.org/10.1080/02601370.2014.973458.
80. Vlieghe J. Rethinking emancipation with Freire and Rancière: a plea for a thing-centred pedagogy. Educ Philos Theory. 2018;50:917–27. https://doi.org/10.1080/00131857.2016.1200002.
81. Nature Editorial. Universities are—and must continue to be—a force for good. Nature. 2025;645. https://doi.org/10.1038/d41586-025-03065-w
82. Hachem HH, Wiggberg M, Osborne T, Gulliksen J, Heintz F. Give MOOCs some credit: a system-divergent innovation accrediting an AI mass-market MOOC at a Swedish university. Cogent Educ. 2025. https://doi.org/10.1080/2331186X.2025.2529421.
83. Nieminen JH. Disrupting the power relations of grading in higher education through summative self-assessment. Teach High Educ. 2022;27:892–907. https://doi.org/10.1080/13562517.2020.1753687.
84. Grande V, Lennerfors TT, Peters AK, von Hausswolff K. The virtuous, the caring, and the free: ethical theory to understand the ethics of the teacher as a role model in engineering education. Eur J Eng Educ. 2024;49:1–21. https://doi.org/10.1080/03043797.2023.2236959.
85. Braga A, Logan RK. The singularity hoax: why computers will never be more intelligent than humans. In: Cognitive technologies. Cham: Springer; 2021. p. 133–40. https://doi.org/10.1007/978-3-030-56546-6_9.
86. Fromm E. To have or to be? New York: Open Road Media; 2013.
87. Yates N, Gough S, Brazil V. Self-assessment: with all its limitations, why are we still measuring and teaching it? Lessons from a scoping review. Med Teach. 2022;44:1296–302. https://doi.org/10.1080/0142159X.2022.2093704.
88. Southwell M, Depaepe M. The relation between education and emancipation: something like water and oil? Introducing the special issue. Paedagog Hist. 2019;55:1–7. https://doi.org/10.1080/00309230.2018.1562481.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.