DOI: 10.1111/ejed.12532

ORIGINAL ARTICLE

WILEY

The future of AI and education: Some cautionary notes

Neil Selwyn

School of Education, Culture & Society, Faculty of Education, Monash University, Melbourne, Victoria, Australia

Correspondence
Neil Selwyn, School of Education, Culture & Society, Faculty of Education, Monash University, 19 Ancora Imparo Way, Melbourne, VIC 3800, Australia.
Email: neil.selwyn@monash.edu

Abstract

In light of fast-growing popular, political and professional discourses around AI in education, this article outlines five broad areas of contention that merit closer attention in future discussion and decision-making. These include: (1) taking care to focus on issues relating to 'actually existing' AI rather than the overselling of speculative AI technologies; (2) clearly foregrounding the limitations of AI in terms of modelling social contexts, and simulating human intelligence, reckoning, autonomy and emotions; (3) foregrounding the social harms associated with AI use; (4) acknowledging the value-driven nature of claims around AI; and (5) paying closer attention to the environmental and ecological sustainability of continued AI development and implementation. Thus, in contrast to popular notions of AI as a neutral tool, the argument is made for engaging with the ongoing use of AI in education as a political action that has varying impacts on different groups of people in various educational contexts.

1 | INTRODUCTION

The future of artificial intelligence (AI) in education—as with any aspect of the future—is uncertain, unpredictable and essentially unknowable. As such, this is not an article that adds to over-confident claims around educational AI being a likely "game changer" (Richardson & Clesham, 2021, p.1) with "the potential to address some of the biggest challenges in education today" (UNESCO, 2019). Similarly, there is little sense in speculating about the

dehumanisation of classrooms, rise of robot teachers, and similar dystopian possibilities (Edwards & Cheok, 2018). Instead, this article looks to make a more modest contribution to the theme of this special issue—reflecting on how discussions of AI and education might best progress over the next decade or so. While AI currently remains a peripheral feature of most schools and universities, the ways in which 'early adopted' AI-driven tools and technologies have already come to bear on educational processes and practices raise a number of contentions that deserve to be taken seriously over the next years. In light of fast-growing popular, political and professional discourses around AI and education, the following elaborates on five broad areas of contention that might foster more rigorous discussion and decision-making; specifically, addressing AI and associated (1) hyperbole, (2) limitations, (3) social harms, (4) ideology, and (5) environmental sustainability.

2 | RECOGNISING AI AS AN AREA OF HYPERBOLE AND OVER-SELLING

While certainly not the first instance of digital technology being over-sold to education audiences (see Cuban, 2001), the past ten years have seen artificial intelligence quickly become a subject of heightened rhetoric and extravagant promotion, replicating previous periods of intensive government and industry interest in AI (e.g., the so-called 'AI Springs' during the mid-1960s and 1980s). While some well-grounded conversations continue to take place, this most recent period of AI resurgence has prompted intense hyperbole and exaggerated claims—particularly as prevailing discourses become co-opted by IT vendors and marketeers looking to profit from AI-related products. In short, there are now plenty of people wanting to talk about (and talk up) the idea of AI in education primarily for profit-motivated reasons.

At a base level, then, any discussion of AI and education is perhaps best started by asking the basic question, *is this actually AI?* This is not a flippant line of inquiry, as illustrated by a 2019 audit of over 2800 purported AI start-ups in Europe that judged two-fifths of these firms to be making no meaningful use of artificial intelligence in their products (Schulze, 2019). Subsequent investigations have similarly found other IT firms claiming to use sophisticated AI capabilities while actually relying on low-tech manual procedures (Johnson, 2021; Morris, 2022). All told, we need to be mindful that education remains vulnerable to what can be termed AI theatre.

Even when an education technology product can be classed as constituting some form of AI, another basic step is ascertaining its actual capacities. Exactly what educational AI products can do (as distinct from what vendors claim they are capable of) therefore needs to underpin any discussions around the tasks that AI might (and might not) be applied to in education settings. As Meredith Broussard has reasoned, it is all too easy to fixate on the ex-ante claims of software developers and IT industry actors, and only much later turn our attention to the actual capabilities of their technology products:

> A lot of the problem we run into with AI is that people make dramatic claims about what the software can do (ex-ante claims) and then the analysis afterward (ex-post) reveals that the claims are false. (Broussard, @merbroussard, 22 April, 2021)

This directs attention to a number of different approaches to how discussions around AI and education might be conducted. For example, some prominent voices in the computer science community—such as Michael I. Jordon—now advocate that we refrain from using the term 'AI' altogether, and instead take care to be more specific about the type (and extent) of machine learning and algorithmic training processes that underpin a particular technology (Pretz, 2021). Alternately, some legal and media scholars now strive to discuss these technologies in more specific terms of 'automated decision making' or 'algorithmic forecasting'. Such shifts in terminology might go some way to acknowledge, as Emily Tucker (2022) puts it, that "whatever the merit of the scientific aspirations originally encompassed by the term 'artificial intelligence', it [has become] a phrase that now functions in the vernacular primarily to obfuscate, alienate and glamorize".

If not deciding to reject the term outright, then it certainly seems reasonable to suggest that those working in education push back against the tendency to presume forms of narrow AI (i.e., AI systems designed to handle one specific limited task and operating within pre-defined boundaries) as a step toward notions of general AI, the singularity, and other niche dreams of technology capable of reaching sentient levels of human intelligence. Instead, as Siddarth et al. (2021) contend, it seems far more constructive to remain focused on matters relating to *actually existing AI*—i.e., the actual computational, material and meta-physical limits of what this technology is capable of doing.

All told, education audiences are best advised to be more circumspect (if not suspicious) of the hype that surrounds the development of emerging AI technologies. This involves being less willing to give serious consideration to what Dwayne Monroe (2021) frames as improbable industry claims around the possible outcomes of speculative future uses of AI, in lieu of paying full attention to the real dangers of already existing implementations of technology. For Lee Vinsel (2021), even setting out to critique and debunk unsubstantiated and speculative marketing is of little use, other than to lend credibility to industry hyperbole. Of course, there can be value in paying some attention to futures thinking and sociotechnical imaginaries that drive the design and development of emerging technologies. Yet it seems unwise to get distracted by the speculative qualities that continue to be attached to AI technologies in education at the expense of engaging with their actual substance. As Deb Raji observed:

> [...] Even AI critics will fall for the PR hype, discussing ethics in the context of some supposedly functional technology. But, often, there is no moral dilemma beyond the fact that something consequential was deployed and it doesn't work. (Raji, @rajiinio, 24 April, 2021)

3 | DISCUSSING AI IN TERMS OF ITS ABSOLUTE LIMITATIONS

These latter points call for closer attention to the specific limitations of AI technologies in education. In this sense, future conversations around AI in education need to move well beyond any sense of what Campolo and Crawford (2020) describe as *enchanted determinism*, i.e., "*the belief that AI systems are both magical and superhuman—beyond what we can understand or regulate, yet deterministic enough to be relied upon to make predictions about life-changing decisions*" (Campolo & Crawford, 2020, p. 1). Such a reframing of AI in education might take place along a number of different lines.

First, is the need to more explicitly promote understandings of AI products as bounded mathematical systems. This raises the fundamental question of which aspects of education can be satisfactorily represented through data, and which aspects of education are decidedly non-quantifiable. One of the inherent limitations in any educational application of AI is the working assumption that all significant facets of student activity and the learning process can be captured in data form. Moving forward, professionals in education need to challenge the capacity of AI to model real world issues that are embedded in social contexts such as classrooms and the attendant problems of representativeness, reductiveness, and explainability of data-driven interventions. As I have argued previously:

> Echoing Murray Goulden's distinction between 'technologically smart' but 'socially stupid' systems, the concern persists that there are not enough data points in the world to adequately capture the complexities and nuances of who a student is, or how a school functions. (Selwyn, 2019, p. 12)

Regardless of the future development of AI techniques, it is likely that attempting to account statistically for the contextual layers implicit in any educational situation will inherently be compromised by the breadth of the social components which these calculations attempt to capture. The statistical limits to what Thea Snow (2021, n.p) describes as "*trying to make the illegible, legible*" were illustrated in a Princeton University study which challenged teams of statisticians, data scientists, AI and machine learning researchers to predict

various life outcomes for children, such as eventual grade point averages and their perseverance with regards to schoolwork. Even when provided with nearly 13,000 data points on over 4000 families stretching back over fifteen years, all teams failed to develop even moderately successful statistical models (Salganik et al., 2020). As news media reporting of this trial bluntly concluded: "*AI can't predict how a child's life will turn out even with a ton of data*" (Hao, 2020).

Second, is encouraging more realistic understandings around the capability of AI products to approximate human traits. Perhaps the most obvious limitation is the incapacity of AI to simulate a wide range of intelligences. As the US computer scientist Melanie Mitchell (2021) contends, these limitations are often glossed over due to the AI community's "*limited understanding of the nature and complexity of intelligence itself*" (Mitchell, 2021, p.1). Here, Mitchell argues that AI development tends to be predicated around a narrow information-processing model of mind that sees intelligence as something that is brain-based, disembodied, and therefore readily relocated to software and hardware. However, this discounts a wide range of human intelligences—such as embodied cognition, common-sense, emotions, and irrational thoughts that are entwined with the complexities of people's social lives. As highlighted in recent debates over the falsely presumed sentience of language generation systems, we need to remain mindful of risks associated with designing AI systems in ways that make users believe they are encountering real, independent intelligence in a program; when in reality AI systems can do no more than fabricate displays of narrow facets of human behaviour (Bender, 2022a).

Another key conflation relates to the ways in which AI technologies are associated with simplistic promises of human autonomy—not least the hope that AI is somehow capable of freeing people from the mental burden of labour and the general daily demands of being human. As Siddarth and Nabben (2021) point out, such expectations gloss over the fact that genuine human autonomy is not an individual quality or personal trait, but something that is circumscribed collectively by the communities and societies in which we live. As such, someone cannot simply decide to be autonomous—instead, individual and community self-sovereignty is borne from social relations with others.

There are many other conflations that might be pointed to. For example, Moser et al. (2022) point to the conflation of so-called AI-driven reckoning (decision-making based on summing up of various kinds of data and technical images) with more nuanced forms of judgement that remain beyond the scope of mechanised processing (i.e., decision-making based on reasoning, reflection, imagination and empathy). Similarly disingenuous is the recent push to promote AI as capable of inferring the presence of emotion and mood (Crawford, 2021), or discerning complex issues of morality or aesthetics. In all these cases, as Dwayne Monroe (2021) concludes, it makes little sense to expect these complex contextually-bounded acts of social comprehension to arise from computational models that can only apply statistical methods to large amounts of already collected data.

Instead, such examples highlight the need to develop more nuanced understandings within education of the absolute limitations of AI technology, and to reset prevailing ambitions and actions accordingly. Of course, pointing out the frailties of AI in education is not to imply that human reckoning and decision-making is somehow substantially better and always preferable. Teachers and education administrators have always made biased, illogical and out-right bad decisions. Education is full of occasions when those in positions of authority might benefit from additional advice or an automated nudge in the right direction. Yet, even reaching the level of being no worse than a human does not justify the adoption of flawed AI technology in an educational setting. As Frank Pasquale puts it, "*We don't have to choose between biased AI & biased humans. We can regulate to improve the AI that is complementing humans.*[...] *not to stop AI entirely*" (Pasquale, @FrankPasquale, 31 July, 2021).

4 | ACKNOWLEDGING THE SOCIAL HARMS ASSOCIATED WITH AI IN EDUCATION

In addition to these previous points, is the need to be more confident in calling out instances when AI products result in social harms in educational contexts. The educational use of AI has already manifested in examples of

social harm that systematically disadvantage and oppress minoritized groups. For example, we have seen various reports of algorithmic discrimination—where AI models amplify discriminations baked into their training data and subsequently erroneously judge students with non-native accents of cheating on tests, or compute higher automated grades for students who fit the profile of those who have historically been more likely to be awarded high grades (e.g., NAO, 2019). We have seen facial recognition software used in schools that regularly fails to recognise students of colour (Feathers, 2020). We have seen the implementation of AI systems designed around processes that predominantly advantage those who are able-bodied and neuro-typical—for example, using eye-tracking data that presumes a steady gaze to denote engagement (Shew, 2020).

Alongside these notable breakdowns, it is also important to acknowledge many other more banal glitches in AI technologies that nevertheless also constitute instances of social harm. Even the most well intended uses of AI perpetuate broader logics of monitoring, categorising, standardisation, synchronisation and social sorting that tend to enable oppression and exacerbate power imbalances. For example, Costanza-Chock (2020) points to various 'dysaffordances' of AI systems—such as instances of students with non-binary gender identities having to mis-identify their gender as either 'M' or 'F' in order to register a profile that is recognised as valid by the system. Such exclusionary features might not be designed deliberately into technologies, but neverthheless perpetuate the regular micro-aggressions that minoritized students encounter throughout their engagement with education systems. Also relevant here are concerns over the capacity of AI systems to support sustained conditions of 'soft surveillance' in classrooms, exacerbating the hypervisibility of minoritized students and teachers (Aradau & Bunz, 2022). All told, as Birhane and Van Dijk (2020) contend, the most pressing matter in the roll-out of AI technology across societal settings is the oppressive use of AI technology against vulnerable groups in society.

Crucially, it is important for those working with education to not presume that these harms can somehow be fixed and made more fair by more inclusive training data-sets, or more mindful approaches to progressive software design. Instead, attention should be paid to an increasing number of arguments—led by scholars such as Joy Buolamwini, Safia Noble, Simone Browne, Ruha Benjamin and Timnit Gebru—that frame AI products as advancing forms of 'engineered inequality' (Benjamin, 2019) in already inequitable social contexts. This points to the ways in which AI technologies result in oppressive and disadvantaging outcomes given their origins in societies that are tightly structured by interlocking forms of domination (Benjamin, 2019). For example, given the inherent racializing and racist logics baked into the structure of facial processing technology, obvious implicit harms are likely to arise from any conceivable application of facial recognition products (however benign the use might seem) in education settings that themselves are historically disposed toward racializing and racist practices.

As such arguments imply, it is important to acknowledge the harms associated with AI in education as relational in nature, and therefore likely to be experienced differently depending on individuals' different backgrounds and circumstances. Any instance of *some* people being disempowered and disadvantaged by the implementation of AI technologies in education is accompanied by *others* being empowered and advantaged. As such, any particular AI technology might appear to work perfectly well, and be of great advantage, for many teachers and students. Nevertheless, for many others, the same technology can simultaneously be experienced in harmful ways. This raises the need for discussions to progress beyond broad-brush concerns over presumed forms of universal harm arising from AI technologies; for example, what Viljoen (2021) describes as 'dignitarian' and humanist fears over a universal loss of human dignity, selfhood and general 'dehumanisation' of social life. Instead, more attention needs to be paid to localised harms being experienced by specific individuals and groups—especially those from minoritized backgrounds.

Conversely, this also foregrounds the need to reconsider (and perhaps reject outright) any talk of *AI for good* in education—as evident in ongoing debates over fairness, accountability and transparency; trustworthy AI; humane AI; and so on. As David Golumbia (2021) reasons, promoting vague notions of AI for good often acts as an inadvertent (if not outright) dishonest way of silencing more complex discussions around racism, ableism and other forms of social discrimination. Instead, future discussions of AI in education need to explicitly address these concerns around educational harms in a more direct, forthright and honest manner. As Abeba Birhane (2022) contests:

> Let's ditch the common narrative that AI is a tool that promotes and enhances human 'prosperity' (whatever that means) & start with the assumption that AI is a tool that exacerbates inequality & injustice & harms the most marginalized unless people actively make sure it doesn't. (Birhane, @Abebab, 17 May, 2022)

5 | ACKNOWLEDGING THE IDEOLOGICAL NATURE OF DEBATES ABOUT AI IN EDUCATION

A further point of contention is the need to engage directly with the value-driven nature of claims and counterclaims around AI in education. In contrast to notions of AI being essentially neutral, even the most basic AI research and development is animated by narrow sets of motivations and concerns that shape what projects are chosen, what problems are addressed and how outcomes are conceived. Then, as AI technologies are subsequently marketed and implemented in societal domains such as education, other values and agendas come to the fore. In short, any talk around the future of AI in education needs to be seen as a site of competing values, interests, agendas, and ideologies.

One dominant set of values that continues to shape debate around AI derives from what might be termed technicist perspectives—i.e., values that tend to underpin the work of software developers, AI researchers and others aligned with computer science. As Birhane et al. (2021) argue, within these professional technicist circles, AI projects tend to be driven primarily by a narrow set of values and concerns over improving technical performance, efficiency and/or generalisability of systems—often shaped by researchers' previous work and understandings, or else the perceived novelty of the application. The emphasis here is very much on what computer scientist Bettina Berendt (2019) describes as a *problem-solving mindset* and ambitions to push boundaries of what is technically possible (rather than what is socially desirable). As such, societal implications and possible ethical consequences tend to be broadly specified, and often framed in terms of technical challenges that can be addressed through better design and development of AI. This mindset is evident in current enthusiasms for notions such as *privacy by design*, or addressing complex issues of social bias and discrimination through correcting statistical bias, underrepresentation and variance in datasets.

This technicist approach to AI is often justified as enabling sound scientific practice alongside the pragmatic desire to develop useable and useful technology. Yet growing numbers of people are pointing to the dangers of seeing AI in purely computational terms. In contrast, then, is a growing critically-minded counter-commentary around AI concerned primarily with questions of social impact and social justice. Such commentators see themselves as raising long-standing concerns that are grounded in historical precedents—not least what Birhane and Guest (2020) identify as the stagnant, sexist, and racist shared past of fields such as behavioural psychology and neuroscience that are now converging upon the development of AI products intended to infer a person's gender and age, or purportedly to detect students' emotions, motivations and intentions. From this point of view, then, it can be argued that "*the AI community suffers from not seeing how its work fits into a long history of science being used to legitimize violence against marginalized people, and to stratify and separate people*" (Van Noorden, 2020). Framing AI along these lines therefore leads to growing calls to imbue AI development with a heightened sense of the social, political and cultural dimensions of this work.

Another competing set of values inherent in the framing of AI and education is what might be termed the corporate perspective of large multi-national Big Tech actors. Here, we can point to a number of notable IT industry logics that increasingly shape popular and political understandings of AI and education. First is the framing of all people (regardless of their circumstances or context) as individual users that operate wholly within a bounded digital ecosystem. Second is what can be described as a politics of scaling. In other words, the idea that all undertakings need to be able to move beyond localised actions, and instead engage in scalable, universalised actions that can dominate markets on a society-wide basis (Pfotenhauer et al., 2022). Third is what can be described as a

'politics of modularity' (Birch & Bronson, 2022). This refers to how Big Tech firms strive to create conditions where they can easily 'plug in' smaller social actors into digital ecosystems that they ultimately retain control over. Such modularity creates a logic of just-in-time, unbundled provision—what might be termed *education as a service* where specialised services and resources are rented out on an on-demand basis (Komljenovic, 2022).

In contrast, is what might be seen a competing set of values reflecting traditional education concerns and interests. Here, we are beginning to see a burgeoning push-back against the prevailing AI logics of capture, control and prediction, and what is perceived to be a general shift toward the mechanisation of human capability (e.g., Felix, 2020). Picking up on decades-old debates over the nature of teaching as an art and craft (rather than a science), such arguments highlight the deeply relational nature of teaching and learning, and the fact that "*even human motivations and basic reasoning capabilities fundamentally arise out of social interactions rather than as individual decision-making capabilities*" (Siddarth et al., 2021, n.p). As Chris Gilliard puts it: "*Good teaching is a combination of art and skill and experience. I'm of the firm belief that no amount of data capture is going to be able to reproduce that*. [However] *this is apparently a fringe belief in educational technology circles*" (Gilliard, 2021, p. 267). Also of relevance here is the question of how current forms of AI in education bump up against understandings around the 'publicness' of education—for example, raising questions of how current popular forms of AI in education reinforce (or undermine) values of democratic education, plurality, inclusiveness, diversity and any form of difference (Saltman, 2020).

All these different perspectives are being put to work in framing and reframing current debates around AI in education. Any set of arguments, assertions or even basic descriptions of AI in education (this article included) need to be seen as positioned within these broader ideological tensions and conflicts. As such, professionals in education need to engage with such discussions and positions in an appropriately knowing manner. Discussions around AI in education are not straightforward matters of technical efficiency or bringing corporate know-how to bear on classrooms. Instead, the topic of AI in education needs to be approached in contestable terms—as a site of struggle and politics, rather than a neutral benign addition to classrooms.

6 | ACKNOWLEDGING THE ENVIRONMENTAL AND ECOLOGICAL COSTS OF AI IN EDUCATION

Finally, there is an underpinning caveat to any speculations around the future of AI in education. This pertains to the ecological and environmental sustainability of continued AI development. Of course, this is not unique to AI and the application of AI in education. For example, the excessive environmental burden of current data-driven technology development has recently come to the fore in the case of cryptocurrency—with a product such as bitcoin estimated to incur an annual energy consumption equivalent to that of Thailand or Norway (de Vries et al., 2022). Yet, such excessive energy demands are also evident in less contentious forms of AI. For example, cutting-edge deep learning techniques now require computational models estimated to each incur resourcing costs of up to US$ 100 billion and resulting in levels of carbon emissions equivalent to those emitted in a month by New York City. In terms of energy drain alone, the cost of continued AI development and innovation is becoming unsustainable (Thompson et al., 2021).

This specific issue of energy consumption relates to a broader point of contention that urgently needs to feature in future conversations around AI in education—i.e., the ways in which this field of work is entwined with growing ecological and environmental harms associated with the production, consumption and disposal of digital technologies. In this sense, any enthusiasms for the increased use of AI in education have to reckon with the materiality of this technology, and its deleterious consequences for the planet. Striving to refashion education around AI and other emerging digital technologies feeds directly into the depletion of scarce resources in manufacturing, usage and disposal processes, alongside the excessive amounts of energy used to support data processing and storage, *and* the exacerbation of waste and pollution issues (see Brevini, 2021). A strong argument can therefore

be made that calling for the continued excessive application of AI technology in any context—education included—makes little sense in term of environmental sustainability.

At present, such issues are rarely (if ever) acknowledged, even within critically minded conversations around education and technology. Thus, while there is growing acceptance of human-initiated climate change and the role of global capitalism in initiating catastrophic devastation to the planet and its people, there is curious unwillingness to implicate digital technology in this process. As Jonathan Crary observes, even the most climate-concerned commentators will "*carelessly presume* [that] *the internet and its current applications and services will somehow persist and function as usual in the future, alongside efforts for a habitable planet and for more egalitarian social arrangements*" (Crary, 2022, p. 4).

Nevertheless, future discussions of AI and education need to find ways of engaging seriously with the environmental fragility of any future dependence on AI technology, and ground any imaginings of possible futures within genuinely connected understandings of the precarious and precious nature of life at a planetary level. This raises the fundamental question of whether it is desirable (or even possible) to continue to sanction the development and use of AI technologies over the medium to long term. Optimistically, there is perhaps merit in working to explore possible alignments of educational AI with 'green-tech' principles—exploring the extent to which increased use of AI technologies in education might actually contribute to the pursuit of forms of ecologically sustainable growth. This might pick up on emerging hopes around green forms of machine learning and carbon-responsive computing and its emphasis on small datasets and refined processing techniques that do not rely on brute force computational approaches (see Nafus et al., 2021). Other proposed alternatives include the deliberate use of AI technology to support the reduction of carbon emissions associated with campus-based travel and education—not least by lowering emissions of students and teachers otherwise commuting to-and-from classes (Versteijlen et al., 2017) alongside the reduction of on-campus power consumption (Caird et al., 2015).

Less optimistically, however, there is also merit in education actors engaging seriously with the contention that the whole concept of AI technology is irredeemable, and the need to learn to live without such technologies. This implies a radical reassessment of the entire education technology project. It might be that any ambitions to instigate what are perceived as cleaner and greener forms of carbon-neutral digital technology and renewable energy prove to be futile attempts to perpetuate what are fundamentally devastating products and practices. In short, it might be that we need to lose any illusions we might have about different forms of AI somehow addressing the environmental challenges that future forms of education will face. Instead, we need to accept that all forms of AI technology *"are intrinsically incompatible with a habitable earth, or with the human interdependence needed to build egalitarian post-capitalist forms of life"* (Crary, 2022, p. i). Seen in this light, discussions around the future of AI in education will clearly have to take on a decidedly different tone.

7 | CONCLUSIONS

All these issues and contentions merit far more consideration than is possible in a short overview such as this, but hopefully even this brief discussion has begun to point toward various ways in which ongoing conversations around AI in education might progress. Despite what we might be told (by those who have a vested interest in telling us), AI is not a straightforwardly good thing for education. It is not a neutral tool that we can look forward to transforming our classrooms, schools and universities over the next few years. Instead, the very idea of AI is something that needs to be extensively scrutinised, challenged and questioned by those who make decisions that affect education, and those who work in the field of education. The future of AI in education is perhaps best approached as a struggle—as something to be contested rather than a *fait accompli*, something to be taken as given. As implied throughout this article, the education community needs to start asking difficult questions of AI, refining expectations of AI in terms of a problematic to be investigated rather than a problem to be solved. Critical discussions of this sort are now beginning to thrive within various academic, industry and civil society forums—led by the likes of

the *Distributed AI Research Institute*, the *AI Now* institute and the *Data & Society* research institute. There is no reason that education concerns cannot take a leading role in these more critically minded conversations, and shift the framing of AI in education toward fundamental questions along the lines of "*what is being done in the name of 'AI', to whom, who benefits, and how can democratic oversight be exerted?*" (Bender, 2022b, @emilymbender, 10 June).

Crucially, any such conversations around AI and education need to be seen as profoundly political in nature, and entangled with broader issues of power, disadvantage and marginalisation (see Verdegem, 2021). As has been argued elsewhere, the dominant industrialised forms of AI technology and infrastructures that are beginning to pervade education "*skew strongly toward the centralisation of power*" (Crawford, 2021, p. 223). Moreover, AI-driven centralisations of power in education already seem to be working to advantage the otherwise advantaged and, conversely, disadvantage the otherwise disadvantaged. The question that now confronts the AI Education (AIED) community (and others that might see themselves as personally benefitting from the continued application of AI in education) is straightforward enough: are you content with this unequal predicament and, if not, how might things be otherwise?

As this latter challenge implies, the topic of AI in education needs to be approached as a political project. Reframed in this light, there are a number of ways that future discussions around AI in education might progress. For example, everyone working around AI and education needs to be ready to examine (and to make explicit) the underpinning values and ideologies that are driving debates around particular issues. This involves reflecting on one's own positionality, as well as pushing back against any claims for AI to be non-political and neutral. Indeed, as Ben Green (2021) reminds us, attempting to claim neutrality is a fundamentally conservative position that constitutes tacit support for maintaining the status quo and, therefore, the interests of dominant social groups and hegemonic political values. People working in the area of AI and education should also not shy away from occasions when their contributions to discussions need to be overtly political—for example, offering advice to policymakers, pushing back against corporate hyperbole and profiteering, as well as standing up for marginalised and disadvantaged groups. All the points of argument and contention outlined in this article require us to take value-driven positions on issues that are inherently political. For example, deciding to either point out or deny how AI might be implicated in climate change is a political act. Similarly, pointing out or denying how AI might be implicated in the perpetuation of institutional racism is a political act.

Finally, is the need to ensure that future discussions and decision-making around AI and education is diversified and enriched by the voices and actions of many others who are not currently included. In short, we need to call out the tendency for discussions of AI and education to be driven by already privileged and dominant voices. Talk of AI and education rarely originates from the people and groups who are *most* disadvantaged by the implementation of AI in education. Instead, discussion of AI in education to date has been something of a closed shop—dominated by those already invested in (and advantaged by) AI. This homogeneity of interests leads to narrow and unimaginative discussions about what AI ought to be, and the education issues that are deemed worthy of being addressed, and what specific social responsibilities are chosen above others. In particular, ensuring a greater diversity of participation is of clear importance if we are to meaningfully pursue discussions around the futures of AI in education. As Laura Forlano reasons:

> 'Who gets to Future?' is an essential question […] So, how might we move beyond status quo futures and toward more pluriversal futures? Black futures? Feminist futures? Queer futures? Trans futures? Crip futures? Working-class futures? Asian futures? Indigenous futures? And multispecies futures? (Forlano, 18 October, 2021)

All told, this article hopefully raises some ideas for further reflection on how the ongoing use of AI in education is a political action that has varying impacts on different groups of people in various educational contexts. This implies pursuing educational AI along more considered lines—in a manner that involves time and effort, and considerable amounts of deliberation, debate, dialogue and consensus-building. Of course, this is

not to argue for the complete rejection of AI in education—as Frank Pasquale (2020) acknowledges, a wholesale dismissal of AI in education would be foolish, regardless of our concerns over the troubling outcomes of particular technologies and trends. Instead, we are now faced with the challenge of engaging with the complex problem of how to make *different* use of AI technologies for *more just* education outcomes. This requires us all in beginning to engage with the topic of AI in education in more nuanced, dis-interested and politically-minded ways. The future starts here!

ACKNOWLEDGMENT

Open access publishing facilitated by Monash University, as part of the Wiley - Monash University agreement via the Council of Australian University Librarians.

REFERENCES

Aradau, C., & Bunz, M. (2022). Dismantling the apparatus of domination? *Radical Philosophy*, *21*(2), 10–18.

Bender, E. (2022a). Human-like programs abuse our empathy – Even Google engineers aren't immune. *The Guardian*, 14 June. https://www.theguardian.com/commentisfree/2022/jun/14/human-like-programs-abuse-our-empathy-even-google-engineers-arent-immune

Bender, E. [@emilymbender] (2022b, June 10). *I'm not interested in how impressed the journalist was. That's not news. What I […]*. Tweet. https://twitter.com/emilymbender/status/1534986608591532032

Benjamin, R. (2019). *Race after technology*. Polity.

Berendt, B. (2019). AI for the common good? *Paladyn, Journal of Behavioral Robotics*, *10*(1), 44–65. https://doi.org/10.1515/pjbr-2019-0004

Birch, K., & Bronson, K. (2022). Big tech. *Science as Culture*, *31*(1), 1–14. https://doi.org/10.1080/09505431.2022.2036118

Birhane, A. [@Abebab] (2022, May 17). *Let's ditch the common narrative that "AI is a tool that promotes and enhances human […]"*. Tweet https://twitter.com/Abebab/status/1526530264717004802

Birhane, A. & Guest, O. (2020). *Towards decolonising computational sciences*. Online paper. Cornell University. https://doi.org/10.48550/arXiv.2009.14258

Birhane, A., Kalluri, P, Card, P., Agnew, W., Dotan, R. & Bao, M. (2021) *The values encoded in machine learning research*. Online paper. Cornell University. https://doi.org/10.48550/arXiv.2106.15590

Birhane, A., & van Dijk, J. (2020, February). Robot rights? Let's talk about human welfare instead. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (pp. 207–213). Association for Computing Machinery.

Brevini, B. (2021). *Is AI good for the planet?* Polity.

Broussard, M. [@merbroussard] (2021, April 22). *A lot of the problem we run into with AI is that people make dramatic […]*. Tweet. https://twitter.com/merbroussard/status/1384934004030418945

Caird, S., Lane, A., Swithenby, E., Roy, R., & Potter, S. (2015). Design of higher education teaching models and carbon impacts. *International Journal of Sustainability in Higher Education.*, *16*(1), 96–111. https://doi.org/10.1108/IJSHE-06-2013-0065

Campolo, A., & Crawford, K. (2020). Enchanted determinism. *Engaging Science, Technology and Society*, *6*(1), 1–19. https://doi.org/10.17351/ests2020.277

Costanza-Chock, S. (2020). *Design justice*. MIT Press.

Crary, J. (2022). *Scorched earth*. Verso.

Crawford, K. (2021). *Atlas of AI*. Yale University Press.

Cuban, L. (2001). *Oversold and underused*. Harvard University Press.

de Vries, A., Gallersdörfer, U., Klaaßen, L., & Stoll, C. (2022). Revisiting Bitcoin's carbon footprint. *Joule*, *6*(3), 498–502. https://doi.org/10.1016/j.joule.2022.02.005

Edwards, B., & Cheok, A. (2018). Why not robot teachers? *Applied Artificial Intelligence*, *32*(4), 345–360. https://doi.org/10.1080/08839514.2018.1464286

Feathers, T. (2020). Facial recognition company lied to school district about its racist tech. *Vice Motherboard*, 2 December. https://www.vice.com/en/article/qjpkmx/fac-recognition-company-lied-to-school-district-about-its-racist-tech

Felix, C. (2020). The role of the teacher and AI in education. In E. Sengupta, P. Blessinger, & M. Makhanya, (Eds.) *International perspectives on the role of technology in humanizing higher education* (pp. 33–48). Emerald Publishing Limited.

Forlano, L. (2021). *The future is not a solution*. Blogtext. Public Books website 18 October, https://www.publicbooks.org/the-future-is-not-a-solution/

Gilliard, C. (2021). 'Smart' educational technology. *Surveillance & Society*, *19*(2), 262–271.
Golumbia, D. [@dgolumbia] (2021). *axiom: The "good" uses to which a technology might be put are at best parts […]*. Tweet. https://twitter.com/dgolumbia/status/1420384545866661889
Green, B. (2021). Data science as political action. *Journal of Social Computing*, *2*(3), 249–265. https://doi.org/10.23919/JSC.2021.0029
Hao, K. (2020 April 2). AI can't predict how a child's life will turn out even with a ton of data. *MIT Technology Review*. Online article.
Johnson, K. (2021 April 5). Government audit of AI with ties to white supremacy finds no AI. *Venture Beat*. Online article.https://venturebeat.com/2021/04/05/government-audit-of-ai-with-ties-to-white-supremacy-finds-no-ai/
Komljenovic, J. (2022). The future of value in digitalised higher education. *Higher Education*, *83*(1), 119–135. https://doi.org/10.1007/s10734-020-00639-7
Mitchell, M. (2021). *Why AI is harder than we think*. Online paper. Cornell University. https://doi.org/10.48550/arXiv.2104.12871
Monroe, D. [@cloudquistador] (2021, October 25). *Here's the key insight which describes all machine learning systems - These models […]*. Tweet. https://twitter.com/cloudquistador/status/1452298573874356227
Morris, M. (2022). 'AI' shopping startup exaggerated tech capabilities to potential investors. *The Information* . Online article.www.theinformation.com/articles/shaky-tech-and-cash-burning-giveaways-ai-shopping-startup-shows-excesses-of-funding-boom
Moser, C., den Hond, F., & Lindebaum, D. (2022). Morality in the age of artificially intelligent algorithms. *Academy of Management Learning & Education*, *21*(1), 139–155. https://doi.org/10.5465/amle.2020.0287
Nafus, D., Schooler, E., & Burch, K. (2021). Carbon-responsive computing. *Energies*, *14*(21), 6917. https://doi.org/10.3390/en14216917
NAO. (2019). *Investigation into the response to cheating in English language tests*. National Audit The Office https://www.nao.org.uk/wp-content/uploads/2019/05/Investigation-into-the-response-to-cheating-in-English-language-tests.pdf
Pasquale, F. (2020). *New laws of robotics*. Harvard University Press.
Pasquale, F. [@FrankPasquale] (2021, July 31). *We don't have to choose between biased AI & biased humans. We can regulate to […]*. Tweet. https://twitter.com/FrankPasquale/status/1421211750691651588
Pfotenhauer, S., Laurent, B., Papageorgiou, K., & Stilgoe, A. (2022). The politics of scaling. *Social Studies of Science*, *52*(1), 3–34. https://doi.org/10.1177/03063127211048945
Pretz, K. (2021, March 31). Stop calling everything AI, machine-learning pioneer says. *IEEE Spectrum*. Online article.https://spectrum.ieee.org/stop-calling-everything-ai-machinelearning-pioneer-says
Raji, D. [@rajiinio] (2021, April 24). *I can't reiterate how much this is true. Even AI critics will fall for the PR […]*. Tweet. https://twitter.com/rajiinio/status/1385935151981420557
Richardson, M., & Clesham, R. (2021). Rise of the machines? *London Review of Education*, *19*(1), 1–13. https://doi.org/10.14324/LRE.19.1.10
Salganik, M., Lundberg, I., Kindel, A., Ahearn, C., Al-Ghoneim, K., Almaatouq, A., & Altschul, D. (2020). Measuring the predictability of life outcomes with a scientific mass collaboration. *Proceedings of the National Academy of Sciences*. www.pnas.org/content/117/15/8398, *117*, 8398–8403.
Saltman, K. (2020). Artificial intelligence and the technological turn of public education privatization. *London Review of Education*, *18*(2), 196–208. https://doi.org/10.14324/LRE.18.2.04
Schulze, E. (2019, March 6). 40% of A.I. start-ups in Europe have almost nothing to do with A.I., research finds. *CNBC*. Online article. www.cnbc.com/2019/03/06/40-percent-of-ai-start-ups-in-europe-not-related-to-ai-mmc-report.html
Selwyn, N. (2019). What's the problem with learning analytics? *Journal of Learning Analytics*, *6*(3), 11–19. https://doi.org/10.18608/jla.2019.63.3
Shew, A. (2020). Ableism, technoableism, and future AI. *IEEE Technology and Society Magazine*, *39*(1), 40–85. https://doi.org/10.1109/MTS.2020.2967492
Siddarth, D., Acemoglu, D., Allen, D., Crawford K., Evans, J., Jordan, M. & Weyl, G. (2021) *How AI fails us*. Technology & Democracy Discussion Paper. Online publication. Harvard University. https://ethics.harvard.edu/files/center-for-ethics/files/howai_fails_us_2.pdf?m=1638369605
Siddarth, D. & Nabben, K. (2021). What tech futurists get wrong about human autonomy. *Noema*. Online article. https://www.noemamag.com/ai-blockchain-human-autonomy-future/
Snow, T. (2021, February 12). The (il)logic of legibility–why governments should stop simplifying complex systems. *Impact of Social Sciences Blog. Blogtext*. Online article. https://blogs.lse.ac.uk/impactofsocialsciences/2021/02/12/the-illogic-of-legibility-why-governments-should-stop-simplifying-complex-systems/
Thompson, N. Greenewald, K., Lee, K. & Manso, G. (2021September 24). Deep learning's diminishing returns. *IEEE Spectrum*. Online article. https://spectrum.ieee.org/deep-learning-computational-cost

Tucker, E. (2022, March 17). *Artifice and intelligence*. Tech Policy Press. Online article. https://techpolicy.press/artifice-and-intelligence/

UNESCO. (2019). *Beijing consensus on AI and education*. Author https://unesdoc.unesco.org/ark:/48223/pf0000368303

Van Noorden, R. (2020, November 18). The ethical questions that haunt facial-recognition research. *Nature*. Online article. https://www.nature.com/articles/d41586-020-03187-3

Verdegem, P. (2021). *AI for Everyone?* University of Westminster Press.

Versteijlen, M., Salgado, F., Groesbeek, M., & Counotte, A. (2017). Pros and cons of online education as a measure to reduce carbon emissions in higher education in The Netherlands. *Current Opinion in Environmental Sustainability*, *28*, 80–89. https://doi.org/10.1016/j.cosust.2017.09.004

Viljoen, S. (2021). A relational theory of data governance. *Yale Law Journal*, *131*(1), 573–654.

Vinsel, L. (2021, February 2). You're doing it wrong: notes on criticism and technology hype. *STS News*. Online article. https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5

How to cite this article: Selwyn, N. (2022). The future of AI and education: Some cautionary notes. *European Journal of Education*, *57*, 620–631. https://doi.org/10.1111/ejed.12532