|  |  |
| --- | --- |
| Name of the country/entity submitting information | Dr Wayne Holmes  Associate Professor (Artificial Intelligence and Education)  UCL Knowledge Lab  IOE, UCL’s Faculty of Education and Society  University College London |

This response is given in my personal/professional capacity as an academic who has been researching and writing about AI and education for a decade. My research has led to reports on AI and education for organisations such as UNESCO, the Council of Europe, and Education International (the Global Union Federation of teachers); many academic publications and books on AI and education; and invited talks and keynotes on AI and education at universities, ministry conferences, and international NGOs around the world (including Belgium, China, Colombia, Germany, Italy, Mexico, Oman, Portugal, Slovenia, Spain, Uruguay, and the USA). I have also advised the Ministries of Education of the UK and Portugal about AI and education.

Wayne Holmes

# Please provide examples of how AI tools and systems, including generative AI, are used in education process and related decision making in your country, organization or educational institution, with examples of specific software where relevant.

NOT APPLICABLE (in my personal capacity)

# Please provide specific evidence of the known impact of AI tools and systems on learners and teachers and on education systems in general, both positive and negative and explain how the impact is monitored. For example, how does the use of AI affect: (a) persons with special learning needs, learners with different linguistic and cultural backgrounds, women and girls; (b) access to education of populations marginalized or underserved due to ethnicity, socio-economic status, displacement and other factors; (c) human interaction between teachers and students; (d) students’ and teachers’ human rights, privacy, safety, engagement, agency and critical thinking; (e) perpetuation of stereotypes and inequalities; (f) the type of information or disinformation that learners and educators are exposed to; (g) assessment of learning; (h) education management.

A first point to note is that there are more than 20 different types of AI-enabled tools developed for use in education (see Holmes, W., & Tuomi, I. (2022). *State of the art and practice in AI in education.* European Journal of Education: Research, Development and Policies, 57(4), 542–570. https://doi.org/10.1111/ejed.12533), of which generative AI (which many conflate with AI) is only one.

While the AI in Education (AIED) research community have undertaken thousands of studies, the claims made about AIED tools (especially commercial AIED tools) are mostly unevidenced, speculative, aspirational and/or marketing.

The overwhelming majority of the AIED studies that have been conducted are conducted by the developers of the tool being studied (i.e., they are not independent, undermining the integrity of the study), or are small-scale (i.e., undermining their credibility).

In other words, the little specific evidence that is available is not trustworthy: we do not yet know the real impact of these tools on learners, teachers, or educational systems, especially over time (see Holmes, W. (2023). The Unintended Consequences of Artificial Intelligence and Education. Education International Research. https://www.ei-ie.org/en/item/28115:the-unintended-consequences-of-artificial-intelligence-and-education).

In any case, the vast majority of those studies assess only efficacy (e.g., Does this particular tool, in this particular quasi-experimental research context, with this particular group of students, improve those students’ test scores?).

Rarely do the studies consider the broader effectiveness of the tool when used by ordinary teachers in conventional classrooms (e.g., What happens when teachers try to integrate the tool in their real curriculum and classroom context?).

Even more rarely do they consider the safety of the tool (e.g., What is the impact of such tools on the mental health, privacy, or agency of the students and teachers?).

And they almost never consider the impact of the tool on the educational ecosystem (e.g., What is the impact of the tool’s deployment on the classroom, on relationships between teachers and students, and on the professional skills of teachers?).

In short, the studies rarely if ever address any of the eight critical examples given in the prompt question. It is not just a case of ‘more’ research, but more targeted research that is independent of the developers and commercial sector, in order to examine carefully the multiple impacts of AI on each of the eight examples and beyond.

If robust evidence is provided by others in response to this questionnaire, I very much look forward to seeing it.

Meanwhile, there is, however, growing evidence of the harm that AI systems is causing in educational settings. For example, many AIED tools, by default, carry out some degree of social scoring (comparing behaviours, characteristics and outcomes of learners and educators), which might lead to discrimination.

Similarly, AIED tools being implemented in classrooms usually extract data from the students’ interaction with the tool. The developers’ avowed aim is to improve their AI model so that they can support more students; the developers’ real aim is to improve their profits. The point is that this means the students are effectively undertaking unpaid labour whenever they are asked to use one of these systems, contrary to a child’s right to be protected from economic exploitation.

In addition, some AI-enabled online examination proctoring often unfairly prevents students taking their exams. They can discriminate against some students due to the colour of their skin or a disability and can exacerbate mental health problems. In short, the imposed adoption of such products can also interfere with the students’ human dignity.

Equally concerning, some students are being failed in their courses because they have been accused of (but deny) having used generative AI in their assessment tasks. This is problematic given that there is no reliable way of detecting the use of generative AI in writing. It also suggests we need to radically rethink the why, what and how of assessment.

# Please provide examples of legislation, regulations (including codes of conduct or institutional rules) or policies addressing or covering the use of AI in educational context, including ethical or human rights concerns around AI development and use, data privacy, bias mitigation, transparency, academic integrity, plagiarism and proper attribution. Is due diligence mandated for the use of AI in educational context? Do students have clear guidance for citing AI usage?

I am currently working with the Council of Europe which is developing a (i) legal instrument to regulate the use of Artificial Intelligence systems in education to promote and to protect human rights, democracy and the rule of law, together with (ii) a Committee of Ministers recommendation to ensure that teaching and learning about AI incorporates the impact of AI on human rights, democracy and the rule of law and prioritises the active participation and agency of teachers and learners.

I am aware that some countries (e.g., South Korea) have introduced some relevant regulations but I am concerned that these regulations have been overly influenced by BigTech.

# Please provide examples of participation of teachers, parents, students or communities in the development of nationwide or internal regulations addressing the use of AI in education. What has been the feedback from teachers, students and parents? Are there mechanisms in place to solicit such feedback?

I am unaware of any such participation (other than, as I understand it, the CoE intend involving students, parents, teachers, civil society representatives, policymakers, and EdTech industry developers in the development of their legal instrument).

# How does the education system support management staff, teachers and students in understanding how to use AI and how AI works? Please provide examples and /or texts of curricula that address both the technological and human dimensions of AI competency (both how it works (the techniques and the technologies) and what its impact is on people (on human cognition, privacy, agency).

I’m very grateful for the wording of this question.

Given that this call is both about **teaching and learning with AI**-enabled tools (which is known as AIED and is covered by questions 1-4) and **teaching and learning about AI** (which is often known as AI Literacy and is referred to in this question), you might consider using the inclusive term “Artificial Intelligence **and** Education” or AI&ED. AI&ED includes both AIED and AI Literacy.

There are in fact very many online sites that purport to teach people about AI (i.e., to help them achieve a level of AI literacy). There are also a small number of books (including my own “Artificial Intelligence in Education. Promises and Implications for Teaching and Learning”). However, almost all of these (apart from my own) focus on the technological dimension (how it works) almost to the exclusion of the human dimension (e.g., its impact on people, human cognition, privacy, agency, fake news, exploitation, and so on).

It is true that, usually, these resources do mention ethics (usually instantiated only as ‘bias’), but mostly this is almost an add-on to the technological focus which tends to ignore the breadth and depth of ethical issues (see Holmes, W., & Porayska-Pomsta, K. (Eds.). (2023). *The Ethics of AI in Education. Practices, Challenges, and Debates.* Routledge. https://www.routledge.com/The-Ethics-of-Artificial-Intelligence-in-Education-Practices-Challenges/Holmes-Porayska-Pomsta/p/book/9780367349721).

Another problem that the human dimension work needs to address are the various misunderstanding, mythologies and hyperbole spoken around AI. In particular, the language used (which has been led by the AI research community) is often misleading (see Holmes, W. (2023). *AIED—Coming of Age?* International Journal of Artificial Intelligence in Education. https://doi.org/10.1007/s40593-023-00352-3).

For example, it is misleading to suggest that machine-based systems ‘learn.’ Learning is a complex phenomenon that only conscious animals do. What ‘machine learning’ does is only analogous to animal learning – it is not the same thing. Calling ‘machine learning’ learning is misleading (‘adapting’ is a more accurate term). It is also potentially dangerous as it credits the machine systems with more than they can actually do (potentially leading to a dangerous overreliance).

Second, it is misleading to suggest that these machine systems can adapt ‘autonomously’. Humans are involved at every stage of the AI development and implementation cycle; the machines are only autonomous in very limited situations. Again, this overly credits the capability of machine systems.

Third, it is also inaccurate and misleading to state that generative AI ‘generates new content’. The content these systems generate might appear to be new but in fact it is just a rehash of existing content (which is why GenAI system developers are being sued by many for Intellectual Property infringement). Again, it is important to ensure that we do not mislead teachers and learners into misunderstanding the truth of these systems. There are many other misunderstandings about generative AI.

Fourth, with regard to education, as I have stated earlier, there is extremely little robust evidence that AI ‘enhances teaching and research methodologies, student assessment and personalising learning experiences’ (and it isn’t even clear why personalisation is a ‘good’ thing, given that it undermines social interaction and collaborative learning). All such claims for AI in education remain speculative and without robust evidence.

# Please provide examples of existing professional development programmes for teachers to use AI technologies. What training and support are provided to educators to effectively utilize AI tools in their daily work?

I am currently working with the Portuguese Ministry of Eduction DGE to develop a MOOC exploring AI and education. I’m not aware of other such examples (apart from in South Korea).

# Please provide examples of policies addressing gaps and inequalities in access to necessary conditions for the use of AI in teaching and learning, for instance aimed at reducing the digital divide between students with easy access to AI tools at home and those dependent on school resources. What measures are in place to ensure that trustworthy and pedagogically appropriate AI tools and resources are accessible to all students, regardless of their socio-economic background or geographical location?

I am unaware of any such policies and look forward to learning about any that do exist.

# Please provide examples of state-supported collaboration or partnership between public educational institutions and corporations producing AI tools for education. Does the education system enforce contracts with specific software providers or is there a choice, at which level and is it informed by feedback from teachers, parents and students, as appropriate? How are data sovereignty and localization being addressed in the context of using international or foreign-developed AI tools in education?

NOT APPLICABLE (in my personal capacity)

# What are the main challenges encountered during the implementation of AI in education? Have there been any technical, ethical, financial or regulatory hurdles in deploying AI solutions in the educational context?

NOT APPLICABLE (in my personal capacity)

# Are there any specific areas within education where you see significant potential for AI integration in the future?

The vast majority of AI tools developed for education have been designed with the aim of automating and hence replacing teacher functions (trying to realise Bloom’s widely-misunderstood 2-sigma effect). However, this radically misunderstands and underestimates what human teachers do in teaching. In addition, as I have mentioned, there is little robust evidence that such tools are either acceptable or effective in educational settings.

Meanwhile, generative AI is being used mainly for shortcuts, which probably means they are undermining learning and cognitive development, while encouraging teachers and learners to become (overly) reliant on the products of BigTech.

So, for me, the one specific area where I see significant potential for AI is the development of tools that genuinely help teachers do what they – as professionals – want to do. Not tools that pretend to be able to do what the teachers do or that aim to replace teacher functions, and ultimately teachers.

But these more acceptable AI-enabled tools should not be developed based on unpaid labour and misappropriated intellectual property, they should not disempower teachers nor undermine student agency, they should not represent the commercialisation of education by the back door, and they should not be what developers think will help teachers, but things that genuinely help teachers (experienced and novices alike).

This requires a radical change in trajectory, and new aims and ambitions for the AIED research and commercial community; aims that avoids techno-solutionism. This also requires that teachers are involved in the development of the AI tools from their conception, so that the tools are designed to tackle real teacher problems (not problems that the developers assume).