



# How should we change teaching and assessment in response to increasingly powerful generative Artificial Intelligence? Outcomes of the ChatGPT teacher survey

Matt Bower<sup>1</sup> · Jodie Torrington<sup>1</sup> · Jennifer W. M. Lai<sup>1</sup> · Peter Petocz<sup>2</sup> · Mark Alfano<sup>3</sup>

Received: 18 May 2023 / Accepted: 7 December 2023 / Published online: 26 January 2024  
© The Author(s) 2024

## Abstract

There has been widespread media commentary about the potential impact of generative Artificial Intelligence (AI) such as ChatGPT on the Education field, but little examination at scale of how educators believe teaching and assessment should change as a result of generative AI. This mixed methods study examines the views of educators ( $n = 318$ ) from a diverse range of teaching levels, experience levels, discipline areas, and regions about the impact of AI on teaching and assessment, the ways that they believe teaching and assessment should change, and the key motivations for changing their practices. The majority of teachers felt that generative AI would have a major or profound impact on teaching and assessment, though a sizeable minority felt it would have a little or no impact. Teaching level, experience, discipline area, region, and gender all significantly influenced perceived impact of generative AI on teaching and assessment. Higher levels of awareness of generative AI predicted higher perceived impact, pointing to the possibility of an ‘ignorance effect’. Thematic analysis revealed the specific curriculum, pedagogy, and assessment changes that teachers feel are needed as a result of generative AI, which centre around learning with AI, higher-order thinking, ethical values, a focus on learning processes and face-to-face relational learning. Teachers were most motivated to change their teaching and assessment practices to increase the performance expectancy of their students and themselves. We conclude by discussing the implications of these findings in a world with increasingly prevalent AI.

**Keywords** Artificial Intelligence · Generative AI · ChatGPT · Teacher beliefs · Motivation

## 1 Generative Artificial Intelligence as an educational disruption

The release of GPT3 in November of 2022 by OpenAI constituted a bellwether for humanity, heralding a new age of powerful and easily accessible generative Artificial Intelligence (AI). Capturing widespread public and media attention (e.g. Mollman, 2023; Roose, 2023; Wingard, 2023), ChatGPT (OpenAI, 2023) could now provide extended text responses to a wide variety of natural language prompts that often passed for intelligent and informed human prose. Amongst those media reports, there were various opinions about the possible implications of generative AI for education, and what teachers might think, but none of these were based on research evidence. Understanding educator beliefs about how teaching and assessment will need to change in response to generative AI is important, because it underpins the sorts of changes that we can expect to see in educational institutions around the world. Understanding the teacher motivations for change in response to generative AI will enable educational leaders to provide impactful professional learning that builds on teacher belief systems. Hence, the purpose of this study is to examine educator perceptions of how we should change teaching and assessment in response to generative AI, and their motivations for changing. Establishing teacher beliefs and motivations will also act as a baseline for further research as teachers and generative AI technologies change over time.

AI has increasingly become an embedded part of contemporary life through its accepted use in technologies that provided personal assistance (e.g., Siri, Alexa), weather forecasting, facial recognition, medical diagnoses, legal support, and beyond (Holmes et al., 2019). Amongst this milieu, a range of ways that AI could potentially benefit education have been identified, for instance through personalised learning platforms, adaptive assessment systems, intelligent predictive analytics, and provision of conversational agents (Akgun & Greenhow, 2021). However, the ability of ChatGPT and similar generative AI tools to provide students and teachers with often high-quality responses to a wide range of common educational tasks raised fundamental questions about what educators worldwide should be teaching and how students should be assessed.

Deep epistemological and pedagogical questions about the use of AI in Education have existed for some time, with Holmes et al. (2019) provocatively asking “If you can search, or have an intelligent agent find, anything, why learn anything? What is truly worth learning?” (p. 3). Touretzky et al. (2019) contend that it is not enough to know how to use AI tools effectively, but that AI needs to become a compulsory topic area woven throughout the curriculum so that all students develop a requisite understanding of how AI works. Whether and how AI should be used by students in the classroom has been compared to whether use of spell-checkers and calculators by students should be allowed (Popenici & Kerr, 2017). However, sophisticated new *generative* AI such as ChatGPT constitutes a wholesale leap in what cognitive and learning tasks can be supplanted by technology, with the potential for students to simply submit copy-pasted ChatGPT output in response to quite elaborate assessment specifications.

Our study addresses an important gap in the research literature. There have been a number of reviews of AI in Education that summarise broad categories of AI uses in the classroom (for instance, Celik et al., 2022; Chen et al., 2020; Chiu et al., 2023b; Xu & Ouyang, 2022; Zawacki-Richter et al., 2019; Zhai et al., 2021). There have also been a range of more philosophically-oriented commentaries by educational experts about how teaching and assessment needs to change in a world of increasingly powerful AI (e.g. Cope et al., 2021; Markauskaite et al., 2022; Schiff, 2021). However, there is a paucity of studies that examine, at scale, teacher perceptions of AI in education. One notable exception was the study by Chounta et al. (2022) that explored perceptions of K-12 Estonian teachers. However, by examining only one country and one level of education, it was not possible to detect whether different regional, level, or other demographic factors influenced teacher perceptions. More importantly, there has not been a large-scale study that examined teacher perceptions in light of the new and more powerful *generative* AI tools such as ChatGPT, so there is no indication of how teachers will or should respond. There is also a pressing need to understand what motivates teachers to change their teaching and assessment in response to generative AI, so that we can provide the appropriate support and policy settings for effective integration of generative AI into education systems.

Consequently, this study addresses the following research questions:

RQ1 To what extent do teachers believe generative AI tools such as ChatGPT will have an impact on their teaching and assessment practices, and how does this vary by demographic and contextual factors?

RQ2 How should teaching and assessment change as a result of generative Artificial Intelligence such as ChatGPT?

RQ3 What motivates educators to change their teaching and assessment as a result of generative Artificial Intelligence?

Addressing these research questions soon after the release of generative AI tools such as ChatGPT also offers a baseline for comparison as teacher beliefs evolve over time.

## 2 Literature review

### 2.1 Defining artificial intelligence in education

Defining Artificial Intelligence is challenging, not least because of the different ways that intelligence can be conceived (Wang, 2019). Baker et al. (2019) define Artificial Intelligence as “computers which perform cognitive tasks, usually associated with human minds, particularly learning and problem-solving” (p. 10). However, for the purposes of this paper we define Artificial intelligence (AI) as “computing systems that are able to engage in human-like processes such as learning, adapting, synthesizing, self-correction and use of data for complex processing tasks” (Popenici & Kerr, 2017, p. 2), because it more completely describes the relevant cognitive processes and highlights the importance of underlying data.

AI in Education (AIED) systems can take many forms, and can be categorised into learner-facing AIED systems such as Intelligent Tutoring Systems (ITS), educator-facing AIED systems such as teacher dashboards and automatic grading support, and AIED systems for institutional support that can identify students at risk of attrition (Luckin et al., 2022). More recently, *generative* AI systems have emerged, based on Large Language Models (LLMs), that are able to answer a broader range of questions and more intelligently than previous AI platforms based on dialogical or dialectical interactions with users (Ouyang et al., 2022).

## 2.2 How text-based generative AI works

Generative AI systems such as GPT-3, ChatGPT, GPT-4 and Bing Chat are based on transformer models that learn “context and thus meaning by tracking relationships in sequential data” (Merritt, 2022; see also Vaswani et al., 2017). These systems are trained on a large collection of textual corpora scraped from the Internet. For instance, for GPT-3, a 45 TB dataset of text (approximately 409 billion ‘tokens’) from multiple sources was passed through an encoder. Text was drawn from sources such as Wikipedia and other websites indexed by Microsoft’s Bing search engine.

To train the model, these corpora are passed through BERT (Bidirectional Encoder Representations for Transformers), which enables the model to treat text not just as an unordered ‘bag of words’ but rather a sequence of words (actually tokens, which can also include punctuation, emoji, and other characters), where what comes before or after matters (Brown et al., 2020; Devlin et al., 2018). Tokens are then processed by the BERT decoder, an auto-regressive model that is designed to predict the next token. In short, this final step is used to predict, iteratively, the next word or other token given the sequence of tokens up until that point. The model initially assigns random probabilities, but it is trained up over many iterations of feedback to be more and more accurate in its predictions. When the user presents the trained model with a prompt, the BERT encoder works on the text as above. Mirroring the training of the model, the text features (tokens) are used to predict what textual features are likely to be associated with them in the next token. These predictions are then decoded into tokens that can be output back to the user. Through this process, generative AI systems trained on large corpora of data can often provide human-like ‘intelligent’ responses to a large variety of prompts.

## 2.3 Use of AI in classes

While the popular emergence of *generative* AI is relatively recent, AI has been used in Education in a wide variety of other ways. Research into AI use in classes has focused on how AI can be used for student feedback, to support reasoning, to enable adaptive learning, to facilitate interactive role-play and to support gamification (Zhai et al., 2021). A recent review of AI in Education found that the application of AI in the classroom can be divided into four main roles: (i) assigning tasks based on individual competence, (ii) providing human–machine conversations, (iii) analysing student work for feedback, and (iv) increasing adaptability and interactivity in digital

environments (Chiu et al., 2023b). Intelligent Tutoring Systems are a particularly common application of AI, supporting learning by teaching course content, diagnosing strengths or gaps in student knowledge, providing automated feedback, curating learning materials based on individual student needs, and even facilitating collaboration between learners (Zawacki-Richter et al., 2019).

The use of AI offers a number of benefits in classes, including selecting the optimum learning activity based on AI feedback, facilitating timely intervention, tracking student progress, making the teaching process more interesting, and increasing interactivity for students (Celik et al., 2022). Teachers have proposed a variety of possible learning designs that integrate AI, which centre around interdisciplinary learning, authentic problem solving, and creativity tasks (Kim et al., 2022). Recently, novel learning designs have emerged based on more powerful AI systems, such as using AI in junior secondary school classes to encourage students to explore different poetic features (Kangasharju et al., 2022), or using AI-induced guidance to optimise learning outcomes by keeping students in the zone of proximal development during discovery learning (Ferguson et al., 2022). Such uses portend great promise for a future with increasingly powerful generative AI. However, there are concerns that AI may only adjust access to content without substantially impacting on core educational practices (Zhai et al., 2021). Of critical pedagogical importance is the relationship between the student and the AI platforms being used, which can vary from AI-directed (learner-as-recipient) to AI-supported (learner-as-collaborator), to AI-empowered (learner-as-leader) (Ouyang & Jiao, 2021, in Xu & Ouyang, 2022).

## 2.4 AI as a topic area in the curriculum

Learning about AI has started to become a part of school and university curriculum (Dai et al., 2020; Touretzky et al., 2019; Xu & Ouyang, 2022). The UNESCO Beijing Declaration on Artificial Intelligence in Education identifies the development of AI literacy skills required for effective human–machine collaboration as a fundamental priority across all levels of society (UNESCO, 2019). Holmes et al. (2019) make the important distinction between ‘learning about AI’ and ‘learning with AI’, with the former being a prerequisite for the latter. According to Touretzky et al., learning about AI means developing a fundamental understanding of how computers work, including processes such as model creation, machine learning, and human–computer interaction.

Some scholars argue that alongside an explicit curriculum about AI, it is critical to overlay general capabilities and dispositions in order for students (and society) to thrive in a world of increasingly powerful AI. Markauskaite et al. (2022) highlight the need for well-developed critical thinking, evaluative skills, creativity, self-regulation, empathy, and ethics when using AI. Carvalho et al. (2022) similarly argue that pedagogical practices that emphasize human skills (creativity, complex problem solving, critical thinking, and collaboration) are needed for supporting one’s ability to communicate and collaborate with AI tools in life, learning, and work. Holmes et al. (2019) argue for the need for deeper learning goals that emphasize versatility,

relevance, interdisciplinarity, transfer, and the embedding of skills, character, and meta-learning into the teaching of traditional knowledge domains. It is unclear more broadly what teachers believe to be important in terms of integrating AI into the curriculum.

## 2.5 Use of AI for assessment

AI may assist teachers in assessment processes, for instance through construction of assessment questions, providing writing analytics, automated use of learning process data, and creating more adaptive and personalized assessment (Swiecki et al., 2022). Another important use case for AI in schools is through automated grading and evaluation of papers and exams, for example, through text interpretation, image recognition, and so on (Chen et al., 2020). AI is sometimes used to evaluate student engagement and academic integrity, and has the potential to generate personalized assessment tasks based on the specific needs of individual students (Zawacki-Richter et al., 2019). AI also provides increased opportunities for focusing on process-oriented assessment and evaluation of collaborative performance (Kim et al., 2022).

However, there are also a range of challenges that need to be addressed when using AI in assessment, including the sidelining of expertise, deferral of accountability, adoption of surveillance pedagogy, and a potentially unproductive separation of humans and machines in the assessment process (Swiecki et al., 2022). In the case of increased access to generative AI tools, there are also concerns about identity, plagiarism, and assurance of learning (e.g. Hisan & Amri, 2023). If students can pass assessment tasks by submitting work that is not their own, then the purpose and integrity of education may be undermined (Cotton et al., 2023). While AI detection tools have emerged almost as rapidly as generative AI itself (for instance, *AI Text Classifier*, *GPTZero*, *AI Cheat Check*, *AI Content Detector*), these systems are probabilistic rather than exact, and there is a risk that teachers may not be able distinguish whether a student's writing is their own work (Cotton et al., 2023).

## 2.6 AI and teacher practice

In terms of assisting teacher work, AI can be used to help support educational decision making with evidence, provide adaptive teaching strategies, and support teacher professional learning (Chiu et al., 2023b). The role of the teacher and how they position AI use in the class is emerging as a critical influence on learning. One study examining the use of AI chatbots with Year 10 students found that teacher support significantly influenced motivation and competence to learn with the AI platform (Chiu et al., 2023a). Alternately, other research has found that not all students benefit equally from the use of AI in education, and passive/mechanical approaches to using AI may actually lead to reduced performance (Wang et al., 2023). Thus, there is a need to understand the teacher's role when helping students to learn with AI, in terms of the learning approach that they take.

The use of AI by teachers involves challenges, including the limited reliability, capacity, and applicability of AI (Celik, 2023). Scholars have highlighted the

importance of Fairness, Accountability, Transparency, and Ethics (FATE) of AI in education, encouraging the use of eXplainable AI (XAI) whereby the reasons for decisions made by AI are transparent and available (Khosravi et al., 2022). The range of ethical issues that needs to be addressed is extensive and nuanced, for instance, perpetuation of existing systemic bias and discrimination, privacy and inappropriate data use as well as amplifying inequity for students from disadvantaged and marginalized groups (Akgun & Greenhow, 2021; Miller et al., 2018). Differences in access to AI platforms have the potential to expand inequality gaps for certain sub-populations, for instance, low-socio economic, female, and Indigenous students (Celik, 2023). Teachers may play a critical role in addressing these issues.

While teachers have a general sense that AI may also present a range of opportunities for education, many have limited knowledge about AI and how to effectively integrate it into their teaching (Chounta et al., 2022). Some teachers have limited interest or motivation to integrate AI into their classes (Celik, 2023; Chiu et al., 2023b). What is important, however, is that teachers have the requisite *AI Readiness*, whereby they understand, at least in non-technical terms, how AI works and what it is capable of achieving, so that they can integrate it effectively into their classes if they so choose (Luckin et al., 2022). Recent reviews have established that there is often a lack of connection between the AI technologies and their use in teaching (Chiu et al., 2023b) and a need for further exploration of educational approaches to applying AI in Education (Zawacki-Richter et al., 2019).

## 2.7 Theorising AI in education research

Educational research relating to AI has tended to be under-theorised (Zawacki-Richter et al., 2019). Several theoretical positions for conceptualising AI have been suggested, including: social realism, AI mediated dialogue, networked learning, knowledge artistry, human-centred AI, Sen's Capability Approach, 4Cs of Human Creativity, Self-Regulated Learning, Networked Learning, and hybrid cognitive systems (Markauskaite et al., 2022). For this study, which examines beliefs and motivations to change technology use, we refer to the most recent version of the Universal Theory of the Acceptance and Use of Technology framework, the UTAUT2 model (Venkatesh et al., 2012).

The UTAUT2 framework (Venkatesh et al., 2012) consists of seven factors that influence the behavioral intentions to use technology: performance expectancy, effort expectancy, social influence, facilitating conditions, hedonic motivation, price value, and habit. The model has been used to study teacher perceptions of a range of technologies in education, including PowerPoint presentation software (Chávez Herting et al., 2020), Massive Open Online Courses (Tseng et al., 2022) and Immersive Virtual Reality (Bower et al., 2020). When used quantitatively, the model has been found to describe 74% of variance in behavioural intention to use technology (Venkatesh et al., 2016). The high explanatory power of the model and its capacity to provide a qualitative model for examining teacher motivations (e.g. Bower et al., 2020) has led to it being used as an a-priori framework for coding teacher motivations in this study.



### 3 Methodology

The objective of this study was to understand whether teachers believe that generative AI tools such as ChatGPT will impact on teaching and assessment, and if so, how. As well, the study sought to determine teacher motivations for changing their teaching and assessment practices, and the demographic and contextual factors might influence their beliefs. A general trend in AI educational research is a preference towards quantitative (80%) and theoretical or descriptive papers (17%) over qualitative (2%) or mixed methods (3%) research (Tang et al., 2021). By adopting a mixed methods approach in this study of generative AI, qualitative results were able to offer explanatory insight into the quantitative data that were collected (Almalki, 2016). Providing insights into motivations for change also enables educational leaders to better understand how they may affect change in the use of generative AI by teachers.

A purposefully designed survey was used to collect data, because of the known capacity of survey methods to efficiently harvest standardised responses from a large number of geographically dispersed participants (Nardi, 2018). Survey methods are often used to conduct high quality research. One recent systematic review paper found that conducting surveys was the third most popular research method in the educational technology field, with 57 out of 365 papers (16%) using survey approaches (Lai & Bower, 2019). The frequent use of survey methods to conduct high-quality research in the educational technology field also accords with other reviews (Baig et al., 2020; Escueta et al., 2017).

#### 3.1 Instrument design

The survey was constructed to address the specific research questions, with Likert scale items relating to participants' perceived impact of generative AI tools such as ChatGPT on what they teach and how they assess, as well as open-ended items relating to the sorts of changes that they believed they should make to their teaching and assessment. The instrument also included open-ended questions asking participants about what motivated them to change their teaching and assessment. Demographic items relating to age, gender, country, teaching experience, teaching levels (e.g., elementary, secondary, university), teaching discipline (e.g., physics, philosophy) were also included to shed light on the impact of individual circumstances. Participants were also asked to rate their prior awareness of generative AI tools such as ChatGPT, as a potential moderating variable. The survey also included other questions relating to professional learning and quality of ChatGPT responses that were not included in this study. In order to increase the content and construct validity of the instrument, a prototype of the survey was sent to seven colleagues with extensive survey research experience and expertise in the Artificial Intelligence field. These experts were selected based on their experience publishing research relating to AI as well as the expertise with developing research questionnaire instruments. Feedback from these research experts was provided about the clarity of wording with respect to intended focus of questions, which was



subsequently incorporated into the final survey design. The final survey instrument is included in Appendix 1.

In order to avoid people not knowing what generative AI tools were or how they could be used in learning and teaching, a link to a basic video overview was provided that included an introduction to ChatGPT and an explanation of how it could be used by students to answer a range of different assessment tasks. The video avoided making any specific judgments about the impact of AI on teaching and assessment, or how people might need to change their teaching, so as not to unduly influence participant responses. Participants were also asked to indicate whether they watched the video, so that the impact of watching the video on participant perceptions could be measured. The video can be found at [https://youtu.be/92y\\_oOXvj6c](https://youtu.be/92y_oOXvj6c).

### 3.2 Distribution and participants

All research protocols adopted in this study were approved by the Macquarie University Human Ethics Committee, reference number 520231285244798. The survey was distributed via social media (e.g., Twitter, Facebook, LinkedIn) and professional teaching listservs. Examples include the LinkedIn High School Teachers group, the Society for Teaching and Learning in Higher Education, and the Australian Council for Computers in Education. Participants provided full and informed consent before completing the survey. Participants were offered the opportunity to enter a prize draw for an iPad or a \$US50 Amazon shopping voucher as an incentive to respond. A snowballing approach was used, where recipients were encouraged to forward the survey to colleagues who may be interested in responding. The data collection period extended from 24th of January to 9th of March 2023, during which time GPT3.5 was the prevailing model. A total of 763 survey responses were received.

After excluding incomplete, non-English, and disingenuous responses, there were 318 participant responses remaining in the sample that were used as the basis of the analysis. According to research by Pérez-Sanagustín et al. (2017), our sample of 318 participants places the size of our sample well within the top third of studies published in the world's highest-ranked educational technology journal, *Computers & Education* (top 9% of qualitative studies, top 43% of quantitative studies, and top 25% of mixed methods studies). Participants included 129 females and 183 males (0 non-binary, 6 preferred not to indicate gender). The average age was  $M=47$  years old,  $SD=10.6$  years. The average teaching experience was  $M=17$  yrs,  $SD=9.8$  yrs. There were 14 elementary/primary school teachers, 82 high school teachers and 222 university teachers in the sample. The Web of Science Research Domain schema (Clarivate Analytics, 2023) was used to categorize teaching areas. This revealed a broad range of teaching areas: 71 Arts & Humanities, 14 Life Sciences & Biomedicine, 40 Physical Sciences, 129 Social Sciences, and 59 Technology.

### 3.3 Quantitative analysis

The two relevant outcome variables from the Likert scale responses were *impact on teaching* and *impact on assessment*, which measured respondents' views of the

impact of AI tools such as ChatGPT on teaching content and assessment, using a four-point Likert scale (1 = none, 2 = minor, 3 = major, 4 = profound). The responses for each variable were modelled using a general linear model with possible explanatory variables of *age group* (<35, 35–55, >55 years), *gender* (male or female), *teaching experience* (<=10, 11–20, >20 years), *geographic region* (Americas, Australasia, Europe, Other), *teaching level* (Primary, Secondary, University), *teaching area* (arts & humanities, life & biomedical sciences, physical sciences, social sciences, technology) and *experience with generative AI* (from 1 least to 4 most). A small number of respondents, those who preferred not to state their gender and those who were in ‘other’ teaching areas, were excluded from the modelling, resulting in 308 responses included in the statistical analysis.

The initial general linear model included all the main effects of the explanatory variables as well as two-factor interactions, with the latter removed stepwise until only significant interactions remained. Residuals passed checks for normality and distributional assumptions (though the effect of the discrete nature of the outcome variable was evident). All analyses were carried out using the *SPSS* package (*IBM*, Version 28).

### 3.4 Qualitative analysis

The qualitative data for this study included the open-ended responses relating to changes that teachers believed they should make to teaching and assessment as a result of generative AI such as ChatGPT, as well as their responses to what motivated them to make those changes. All questions were imported into *NVivo Release 1.7* software and coded separately using thematic analysis (Clarke & Braun, 2013). Due to the newness of generative AI, the questions relating to changes to teaching and changes to assessment were coded entirely inductively. As previously identified, the questions relating to motivations to change assessment and teaching adopted the Unified Theory of Acceptance and Use of Technology (UTAUT2) framework (Venkatesh et al., 2012). This framework includes seven themes: Performance Expectancy (improvement of outcomes); Effort Expectancy (anticipated work); Social Influence (perceptions of others); Facilitating Conditions (environmental supports); Hedonic Motivation (enjoyment); Price Value (benefit for the cost); and Habit (regular behaviour). Subthemes for each UTAUT2 category were coded inductively as they emerged from the data.

For each question, a sample of 20% of responses was coded by two raters, to establish the reliability of the coding schemes that had emerged. The frequency weighted Cohen’s (1960) Kappa for the four qualitative dimensions analysed were themes that emerged were: ‘changes to teaching’  $K=0.92$ , ‘changes to assessment’  $K=0.90$ , ‘motivations to change teaching’  $K=0.96$ , ‘motivations to change assessment’  $K=0.87$ . This indicates a ‘high’ level of Inter-Rater Reliability (Landis & Koch, 1977). Due to the high levels of Inter-Rater Reliability, the remaining 80% of responses were coded by one of the two raters. The coding scheme for each topic area, example responses, and details of specific Kappa scores for the themes and subthemes from each topic are provided in Appendix Tables 1, 2, 3, and 4.

Qualitative analysis is reported after the respective quantitative analysis (e.g., what should be taught and corresponding motivations) to promote more direct interrelation of qualitative and quantitative results. Note that participant responses regarding how teaching and assessment should change are their perceptions only, and do not constitute evidence that these changes should be made. All data has been deidentified to preserve anonymity of participants, in accordance with the ethical protocols approved for this study. The approach to reporting qualitative findings uses participant quotes to preserve fidelity to the underlying data as well as frequency counts to indicate the prevalence of the various themes. In this way the reporting aims to characterize the nature of participant responses in the most reliable and unbiased way possible.

## 4 Results

### 4.1 Changes to teaching as a result of generative AI

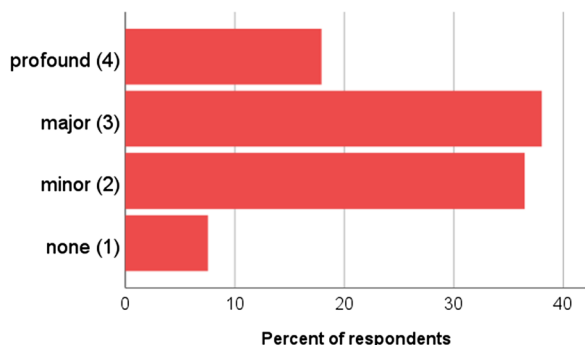
#### 4.1.1 Perceived impact of generative AI on what should be taught

Figure 1 shows the percentage of respondents in each category. The modal and median category is ‘major’, and with the categories numbered 1 = none to 4 = profound, the mean is 2.7, *SD* 0.9. More than one-third of respondents are in each of the ‘minor’ and ‘major’ categories (37% and 38%) while most of the others are in the ‘profound’ category (18%); only 7% selected ‘none’.

The final general linear model included all main effects and a single interaction effect, between *Experience Group* and *Region*, resulting in an R-squared of 20% (15% adjusted). *Region* and *Awareness* were statistically significant ( $p=0.007$  and  $0.003$  respectively), *Gender* and *Level* were marginally significant ( $p=0.045$  and  $0.046$ ), while *Age Group*, *Teaching Area* and *Experience Group* were non-significant ( $p=0.63$ ,  $1.00$  and  $0.11$ , respectively). *Experience Group* by *Region* showed a significant interaction ( $p<0.001$ ).

**Fig. 1** Impact of AI tools such as ChatGPT on Teaching content

**Impact of AI tools such as ChatGPT on Teaching content**



Using the numerical values of the categories, respondents with high awareness had higher averages than those with medium or low awareness (2.7 vs. 2.4 and 2.3, representing a medium effect of 0.45 on the Cohen's-*d* scale). Females had a higher average than males (2.6 vs. 2.4, a small effect of 0.25). Primary teachers had lower average than secondary and university teachers (2.1 vs. 2.6 and 2.7, a large effect of 0.7). In the Americas, all three experience groups had averages of 2.5 or 2.6 while Australasian teachers with high experience and European teachers with low or medium experience had lower averages of 1.8 to 2.1 (the difference representing a large effect of 0.75). A small number of respondents from 'other' regions with high experience had a much higher average of 3.6 (again, the difference representing a large effect of 1.2).

#### 4.1.2 Changes to what should be taught as a result of generative AI

A total of 546 references were coded, with results relating to *what* to teach (**curriculum**), *how* to teach (**pedagogy**), and **other** comments. In terms of **curriculum**, the most frequent response from participants about the changes to *what* to teach was teaching students **how to use AI** ( $n=53$ ) as an integrated part of learning activities in the classroom. Responses ranged from broad ideas, such as using ChatGPT to “*support learning*”, “*ask specific questions*” and “*gather information*”, to more nuanced application for specific learning areas, for instance “*to translate Latin texts*”, to use ChatGPT “*as an instructor*” in Spanish classes, and to harness its capability to provide individual learning support to students, particularly with writing. Many respondents indicated that they were already confident with the integration of digital tools in their teaching, revealing that AI provided them the opportunity to extend their usual tasks, for instance, using it “*for enhancing our multimedia research-creation digital workflow*”. Many teachers felt that it was important to teach specifically about **how AI tools work** ( $n=40$ ) including “*what they are, as well as the capabilities and limitations*” and their “*usefulness*”. Similarly, the importance of teaching students **critical thinking**, especially relating to evaluating AI responses, was strongly emphasized by teachers ( $n=38$ ). Many detailed how it was imperative to move students beyond being passive recipients of AI, for example, by being “*alert and critique what is offered as 'knowledge'*”. Teachers acknowledged the importance of teaching about **ethical and responsible use of AI** ( $n=26$ ), commenting both generally (for example, “*how to ethically use them*”) and specifically. Detailed responses included teaching about “*ethical dilemmas*”, “*why citing matters*”, and an increased focus on teaching about plagiarism and academic integrity. Lastly, 17 respondents described how they would use ChatGPT in their teaching to **provide exemplars or models** for their students. From “*explaining code related to statistics*” to “*analyzing and editing*” a ChatGPT response, teachers described how leveraging AI in this way helped provide worked examples for students to critique, elaborate on, or use as scaffolding prompts to produce more in-depth learning.

As per the quantitative results, there were also a number of participants who indicated that they would **not change their teaching** ( $n=31$ ). Interestingly, there were only three people who indicated that they would **discourage the use of AI**, which is a small number across the sample of 318 respondents. There were also several

teacher educators who expressed that they would change the curriculum by **teaching trainee teachers** how to teach using AI ( $n=8$ ). They stated that new teachers needed to be equipped to teach about AI in schools to “*develop younger students’ digital literacy*”.

In conjunction with identifying changes to *what* would be taught, teachers also outlined *how* their teaching (i.e., **pedagogy**) would change in response to AI tools. The most frequent subcategory was specifying their **changed pedagogical approach** ( $n=60$ ). Many responses identified how current teaching styles needed to be adapted to foster active involvement and engagement of students through increased groupwork, higher-order thinking, creative processes, and attention from the teacher within the classroom. For instance, teachers named “*more face-to-face and group projects*”, increased monitoring of the progress of tasks by “*drafting processes*”, “*in-class observations and interviewing*”, and “*utilising creative media*”. Similarly, teachers described how the **design of learning tasks** would need to be modified ( $n=28$ ). Here, too, teachers identified the need to design activities with creative, higher-order modifications and use fewer essays, recall or simple explanations that could be generated by AI. Many responses indicated a greater focus on “*process writing*”, that is, completing writing in stages, for greater accountability. Interestingly, one teacher detailed how AI would be incorporated to “*create an avatar of a historical figure*”, providing an exciting and engaging new approach for students to conduct interviews. Teachers also acknowledged how AI could be utilized to **support their teaching practice** ( $n=23$ ), reporting it as a convenient and efficient tool to “*improve workflow*”. Teachers used AI to create resource materials, collect ideas for lessons, and compare their own resources with those created by AI (“*compare my material and evaluate against core material utilized by AI tools*”). Notably, it was acknowledged that AI could quickly and efficiently create differentiated resources for students “*personalized to their interests and also calibrated to their ability*”. Finally, sixteen teachers provided general statements reflecting that their pedagogy would change, although they believed the overall content would remain unchanged.

The **other** category included general reflections or opinions about ChatGPT/AI unrelated to changes to their own teaching ( $n=13$ ), responses about changes needed at a policy level rather than identifying changes to their own teaching ( $n=11$ ), or unclear responses ( $n=16$ ).

### 4.1.3 Motivations to change teaching

A total of 370 references were coded, averaging 1.2 per respondent (Appendix Table 2). Most of the responses ( $n=233$ ) were classified as Performance Expectancy, that is, respondents were motivated to change how and what they will teach due to the belief that AI will assist them, and their students, in their future performance. Interestingly, the most frequent responses were about their students’ future performance rather than their own. Over one-third of participants were motivated by the **performance expectancy of their students** ( $n=136$ ), commenting that AI would “*enhance student learning*”, give them “*an authentic learning experience*”, and provide them “*knowledge and skills*”. A sense of urgency in the lexicon was apparent: educators concerned that their students were “*not missing out*”, “*need to be made aware*” or “*have to be ready*” for an AI-integrated future. The second-highest

response category was **performance expectancy for teachers** themselves ( $n=97$ ). For example, teachers saw benefit in utilizing AI to keep current with their knowledge “*in order to enhance my ability to share state-of-the-art knowledge with my class*” and to “*facilitate better learning and teaching*”. Being “*relevant*” or “*authentic*” as an educator were terms commonly used to justify their change motivation.

Three categories had noticeably fewer coded responses: facilitating conditions ( $n=36$ ), effort expectancy ( $n=26$ ), and social influence ( $n=24$ ). Respondents acknowledged that they were motivated to change due to infrastructure for AI already existing (i.e., **facilitating conditions**). This is shown with comments such as “*Chatbots are here*”, we are “*...surrounded by technology*”, and that AI will continue to morph with already available technology to become “*ubiquitous in the future*”. Responses coded to **effort expectancy** (the anticipated effort to learn and teach using AI in the classroom) commonly used words such as “*efficient*”, “*convenient*”, “*fast*” and “*timesaving*” to describe using AI as a resource for their teaching practice. Only 7% of participants were motivated to change due to persuasion from others, or societal pressure (**social influence**). While some were motivated to “*follow the trend*”, others were pragmatic, almost defeatist, in their reasoning: “*I don’t have a choice... it would be irresponsible of me not to prepare my students to the world that’s to come*” and “*...there’s not really a choice. Evolve or ‘die’*”. The language used in these responses was primarily negatively phrased.

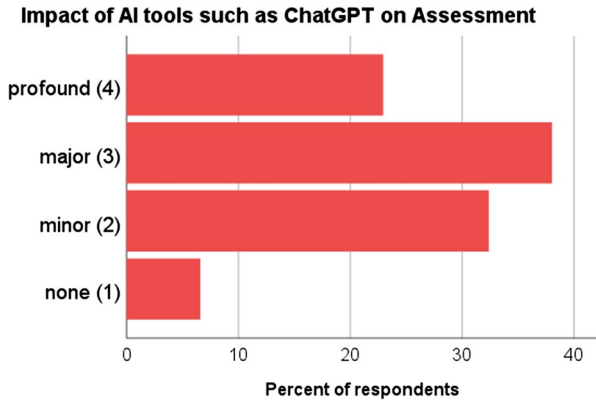
Only 15 educators (5%) expressed their motivation for change due to the fun or pleasure derived from AI (**hedonic motivation**). They noted that AI was “*extraordinary*”, and that the technology was frequently referred to as “*fascinating*” and “*exciting*”. The passionate language was encouraging, with one respondent declaring “*I love disruptive tools!*”. In comparison to the Social Influence category, these 15 responses were phrased positively and seemed uplifted in their enthusiasm for AI. There were no responses coded for the two categories of Price Value or Habit. This suggests that the participants were not motivated to use AI because of its cost, or lack thereof (Price Value), which is an interesting finding given that entry-tier versions of the technology are free to use. Habit (automatic performing of a behaviour) was not recognized by participants as a motivation for change, which was unsurprising given the newness of ChatGPT. Finally, 26 respondents were unsure of what motivated them, or reiterated that they were not making any changes to their teaching. There were 10 ambiguous responses.

## 4.2 Changes to assessment as a result of generative AI

### 4.2.1 Perceived impact of generative AI on how we should assess

Figure 2 shows the percentage of respondents in each category. The modal and median category is ‘major’ (and if the categories are numbered 1 = none to 4 = profound, the mean is 2.8,  $SD$  0.9). More than one-third of respondents are in the ‘major’ category (38%), almost one-third in the ‘minor’ category (32%), and almost one-quarter are in the ‘profound’ category (23%). Only 7% selected ‘none’.

**Fig. 2** Impact of AI tools such as ChatGPT on assessment



The final general linear model included all main effects and the same single interaction effect, between *Experience Group* and *Region*, and resulted in an R-squared of 20% (14% adjusted). *Gender*, *Region* and *Awareness* were statistically significant ( $p=0.003$ ,  $0.001$  and  $0.005$  respectively), *Level* and *Teaching Area* were marginally significant ( $p=0.039$  and  $0.011$ ), while *Age Group* and *Experience Group* were non-significant ( $p=0.46$  and  $0.74$ , respectively), and *Experience Group* by *Region* showed a marginally significant interaction ( $p=0.028$ ).

Using the numerical values of the categories, respondents with high awareness had higher averages than those with medium or low awareness (3.0 vs. 2.6 and 2.8, representing a small to medium effect size of 0.35); males had lower average than females (2.7 vs. 3.0, again a small to medium effect of 0.35); primary teachers had lower average than secondary and university teachers (2.4 vs. 3.0, a large effect of 0.75); teachers in the life and biomedical sciences had higher averages than those in other teaching areas (3.5 vs. 2.5 to 2.7, representing a large effect of 1.1). In the Americas all three experience groups had averages of 2.5 or 2.6 while Australasian teachers with high experience and European teachers with low or medium experience had lower averages of 1.8 to 2.1 (the difference representing a large effect of 0.75); a small number of respondents from ‘other’ regions with high experience had a much higher average of 3.6 (again, the difference representing a large effect size of 1.3). In terms of regions by experience groups, the lowest averages were seen in the Americas among those with medium and high experience (2.3) and the largest in other regions (3.1), with again an unusually high average among a small number of those with high experience (3.6); these differences represent a large effect of 1.0 on the Cohen’s-*d* scale.

#### 4.2.2 Changes to how we should assess as a result of generative AI

A total of 567 references relating to assessment were coded, representing 1.8 codes per person (see Appendix Table 3 for details of the coding scheme). The major themes related to **assessment task types**, the **content of assessments**, means of **evaluation**, approach to **supervision**, and **AI assistance** for assessment processes.



Respondents discussed a range of different **assessment task types** ( $n = 67$ ) that were suitable for a world with increasingly powerful generative AI. The most frequently mentioned included **spoken tasks** ( $n = 29$ ) such as “*real time discussions*” and “*verbal assessment such as mini-viva*”. **Examinations** ( $n = 21$ ) were the second most frequently mentioned assessment type, followed by **paper and pen approaches** ( $n = 7$ ). Sometimes these sub-themes were raised together, for instance “*students need to complete all assessments in an examination situation and in pen to paper*”. Other suggested assessment task types included **video production or visual essay** ( $n = 7$ ) and **portfolio** ( $n = 3$ ).

Respondents identified that changes to the **types of content** included in assessment were warranted to address AI issues. A large number of respondents identified that assessment tasks needed to be inclined towards **higher-order thinking** ( $n = 71$ ), including “*critical thinking and application*” and “*evaluation-type questions*”. Other suggestions were shifts towards more **authentic tasks** ( $n = 65$ ) and focus on the **process** of assessment ( $n = 50$ ) like “*incorporating a lot more process-driven methodologies (drafting, research booklets etc.)*”. Additional recommendations for more **personalized** tasks ( $n = 35$ ) such as “*reflective commentary about prior learning and other lived experiences*”, and tasks that require the provision of accurate **references and citations** ( $n = 12$ ) were proposed.

There was widespread agreement that more **supervision** of assessment tasks was needed ( $n = 69$ ), including the suggestion that invigilated, face-to-face assessment would be necessary to help address the possibility of students plagiarising from AI tools such as ChatGPT. There were also some people who suggested the benefits of having more **collaborative and peer-based** assessment tasks ( $n = 10$ ). For instance, one respondent proposed to make “*greater use of .... group work rather than relying on an essay*”. Teachers also suggested **changing evaluation** processes ( $n = 11$ ) to address AI created work, for example by “*rewording assessment tasks and rubrics to create tasks that are less able to be generated with ChatGPT*”.

Surprisingly, many respondents proposed using **AI assistance** to “*help with assessment tasks*” ( $n = 54$ ). Some participants suggested embedding AI into assessment tasks, for instance, “*ask students to create first draft in ChatGPT or equivalent and then show how they enhanced it, as well as a reflection on the use of AI in the field*” and “*use ChatGPT in the information gathering phase prior to assessment to build a platform of understanding*”.

Some teachers commented that AI would have **minimal or no impact** on their assessment tasks ( $n = 36$ ), either because the nature of the “*assessments can't really be completed with AI*”, or they believe that AI would not do a better job than human - “*I am not impressed by the code ChatGPT currently produces*”. There were a range of **other** more general comments and reflections ( $n = 49$ ), for instance, “*Lots of great resources out there and we're working our way through discovering and coming up with ideas too. Still early days and how we assess will continually evolve as these tools evolve*”.

#### 4.2.3 Motivations to change assessment

A total of 360 motivations for changing assessment were coded, with over 70% (264 responses) relating to **Performance Expectancy**. Similar to the results of the motivations

to change their teaching, the majority of teachers expressed the intention to change assessment tasks to influence the *performance expectancy of their students* ( $n=160$ ). Participants believed that the change would enable “students to develop their own skills and abilities”, “prepare students for life and workplace with AI”, and they wanted to “de-motivate students from using AI as a way to cheat on their coursework”. Some of the *performance expectancy responses related to teachers* ( $n=104$ ). For instance, teachers acknowledged that the use of AI could improve their performance by assisting them with “plagiarism checks” and allow them to “ensure assessment is fair and a true representation of what student is capable of”. In addition, they hoped that they could use “ChatGPT to custom design some of the assessments’ to meet the needs of their students”.

Regarding *effort expectancy* ( $n=15$ ), teachers were motivated to use AI tools as “it makes assessment much easier”. They perceived applying AI tools “could make me more efficient at providing feedback”. Some motivations for changing assessment were due to the *facilitating conditions* ( $n=12$ ), with the widespread emergence of AI tools in teaching and learning being an environmental catalyst (“for me it is part of digital transformation of education”). Other individuals/stakeholders played some roles to motivate teachers to apply AI tools in their assessment processes (*social influence*,  $n=10$ ). They explicitly stated that “there were a lot of people use it” and “there’s nowhere to hide” from AI tools. They agreed “to adapt” to the new changes in order to “not be left behind”. A few teachers indicated *hedonic motivation* ( $n=7$ ) was one of the reasons they changed how to assess students (“I’m excited about AI”). None of the responses indicated a motivation to change their assessment practices because of Price Value or Habit. There were 19 respondents who commented that they were not motivated to make any changes to the current assessment tasks.

## 5 Discussion

### 5.1 Teacher perceptions of how generative AI impacts on teaching and assessment

There was a wide variety of educator perceptions about the impact of generative AI tools such as ChatGPT on teaching and assessment. While the majority of teachers thought that generative AI tools would have a major or profound impact on teaching and assessment, there was also sizeable minority who believed it would have no or minor impact. Awareness of AI tools such as ChatGPT had a significant influence upon the impact educators felt generative AI would have on teaching and assessment, with greater awareness and experience with AI tools leading to beliefs that such tools would have a greater impact. This result points to an ‘ignorance effect’, where people with low experience with AI tools may not believe that AI will impact on teaching assessment because they don’t fully comprehend the capabilities of such tools.

There were other demographic and contextual factors that influenced educator beliefs about the impact of AI on teaching and assessment. Primary school teachers felt that generative AI would have less of an impact on teaching and assessment than high school or university teachers, which is unsurprising given the less sophisticated

nature of tasks in younger years as well as the heavier reliance on face-to-face teaching and assessment. The interaction between teaching experience and global region was at first surprising to our research team. Upon reflection, it did make sense that different regions may have different perceptions about the impact of generative AI due to differences in media coverage or the nature of their education systems, and that there could be differences in perceptions according to teaching experience. This result provides an important reminder of how situational factors mean that the use, acceptance, and perceptions of technology can vary by context.

## 5.2 Teacher beliefs about how teaching should change as a result of generative AI

Participants made a number of suggestions about how to change teaching in response to generative AI. Teachers felt that the curriculum needed to change to teach students how AI works, how to use AI, as well as the critical thinking skills and the ethical values needed for working in an AI-saturated world. Teaching was seen as needing to shift to emphasize learning processes that included creativity, collaboration, and multimedia, and to achieve efficiencies by using AI to assist pedagogical design and administration. Creative ideas included using generative AI for exemplars, simulations, practicing interview techniques, and examples to critique. It is important to note that the suggestions do not constitute empirical evidence of what should change, but rather teacher perceptions. These changes to teaching are quite different from themes identified in previous reviews of AI in Education (e.g. Chiu et al., 2023b; Zhai et al., 2021). These educational possibilities do accord with some of the more innovative approaches suggested in emerging work from teachers (e.g. Kim et al., 2022), and highlight the range of new opportunities for teaching that generative AI avails. Interestingly, the majority of changes suggested by teachers accorded with what is already considered good teaching practice in terms of the underlying pedagogies (Churchill et al., 2016; Hattie, 2023).

Responses from teachers revealed relatively little *future focus* of the impact of AI in relation to its capabilities and possibilities in Education, as compared to the research literature. For example, of the four roles outlined by Chiu et al. (2023b) that AI can be assigned in the classroom, teacher responses in this study mainly focused on the role of “*increasing adaptability and interactivity in digital environments*” (p. 3). Surprisingly, there was relatively limited acknowledgement of AI being used to provide individual differentiation, human-machine conversations, or providing feedback to students. Similarly, recent research outlining how AI can be utilized by teachers, for example for individual intervention or differentiation, monitoring student progress (Celik et al., 2022; Chiu et al., 2023a), providing feedback (Zhai et al., 2021), and utilising AI to maintain a zone of proximal development for students (Ferguson et al., 2022). Importantly, these possibilities for how AI can be used by teachers were not reflected by the majority of participants in this study, possibly indicating that many teachers may have limited knowledge about the full potential of AI. This observation accords with the quantitative findings regarding an association between AI awareness and perceived future impact.

### 5.3 Teacher beliefs about how assessment should change as a result of generative AI

Participants also identified a number of ways that assessment should change as a result of generative AI. Most teachers appeared to realize the challenges that AI propose to assessment, and made explicit suggestions about different ways to better assess students (e.g., spoken tasks, invigilated face-to-face assessments, examinations, video production, etc.). They often recommended that the nature of the assessment tasks change to include ‘higher-order and critical thinking’ as well as ‘authentic and creative tasks’, which aligns with suggestions throughout the literature (e.g. Kim et al., 2022). Proposed changes were due to teachers’ concerns about the academic integrity and plagiarism issues relating to unauthorized use of AI, concerns that have been raised by other researchers (Hisan & Amri, 2023). These rationales accord with the identified importance of ‘*Fairness, Accountability, Transparency and Ethics (FATE)*’ for AI in education (Khosravi et al., 2022), addressing AI-created work to make assessment as fair as possible, from both teachers’ and students’ perspectives.

The teachers in this study proposed various ways to amend assessments to promote fairness and improve learning outcomes (for instance, by adjusting marking rubrics to account for AI). While participant suggestions may provide useful ideas for educators to consider in response to generative AI, once again, it is important to remember that they do not constitute evidence that these changes are efficacious. As well, relatively few teachers in this study focused on more divergent and advanced AI-assisted assessment processes, for instance, automated and adaptive assessment systems, personalized learning systems, and intelligent predictive analytics (Akgun & Greenhow, 2021). Only a minority of teachers indicated knowledge about how AI may help in assessment processes through ‘construction of assessment questions’ and ‘provision of writing analytics’ (Swiecki et al., 2022). Hence, it appears crucial to provide professional development to teachers about how AI can assist them to make their assessment processes more effective and efficient.

### 5.4 Motivations to change teaching and assessment as a result of generative AI

Almost three-quarters of teachers in this study were motivated to change their teaching due to their belief that AI would assist them and their students in their future performance. This very strong performance motivation indicated that most teachers were intent on keeping current with innovative technological developments and helping their students be ready for the AI integrated workplace. The implication is that teachers are willing to learn and integrate AI in the classroom, viewing it as a technological development that needs to be addressed rather than ignored. Therefore, professional learning to develop AI knowledge and understanding appears to be the primary concern to ensure that the changes made to their teaching are effective and encompass the diversity of pedagogical possibilities suggested by the AI literature.

The majority of teachers (70%) were motivated to change assessment tasks because it would help students and teachers to improve different aspects of their performance. This very high performance motivation is in line with suggestions by Swiecki et al. (2022) that a key driver for using AI in assessment is to improve the quality of

assessment processes. Although related to performance, relatively fewer respondents indicated increasing student engagement and detecting academic integrity were key motivators (as outlined by Zawacki-Richter et al., 2019). Teachers less frequently indicated that efficiency was a motivation for changing assessment processes, for instance through feedback and grading, as proposed by Chen et al. (2020). These omissions may be due to teachers' lack of awareness of the numerous ways that AI can be used to support assessment, which points to the relationship between motivation and understanding, and consequently the need for teacher professional development.

## 5.5 Limitations & future research

This study was the first study we could find relating to teacher perceptions of generative AI tools such as ChatGPT, and the influence of such tools on what should be taught and how assessment should take place. However, it should be noted that respondents self-selected via invitation from professional social media channels, so that the responses may not accurately represent the distribution of perceptions amongst all teachers. Although the sample was the largest we could find on the topic of teacher perceptions of AI, the number of participants in each of the demographic categories (for instance, region, age, teaching experience, etc.) could in no way accurately represent the respective populations. Rather, the results provide exploratory indications of possible trends and factors that could lead to differences. It should also be acknowledged that the sample size adopted in this study did not allow results relating to coding schemes to be definitively determined. We provide the Inter-Rater Reliability Kappa values to enable readers to interpret the reliability of findings for their contexts and uses. All results should be interpreted with due caution.

The emergence of powerful generative AI tools provides a rich and important basis for future research. The results of our study highlight the relevance of investigating specific contexts, that may vary by teaching level, discipline, teacher experience, and so on. In addition, we believe there is urgent need to investigate how educators can effectively develop the AI understanding and literacies that they and their students will need to thrive in a world of increasingly powerful generative AI. Teachers made a number of valuable recommendations about how to improve classroom teaching and assessment, including using generative AI as an instructor, a translator, a source of feedback, an exemplar for critique, a datasource, a dialogic partner, and a tool to expedite their digital workflows, for instance, through the creation of assessment tasks and rubrics. Researching the efficacy of such design patterns is an imperative, so that we can provide teachers with evidence-based recommendations about how to best use generative AI in their teaching practices. Additionally, we recommend research into professional learning needs of teachers as an important area to understand if we are to equip the education field to respond swiftly and competently to the rapid emergence of increasingly powerful generative AI. Finally, just as the findings from this study can inform policymaking in the area of generative AI use in Education, further research is needed to understand the pedagogical and social implications of different policy settings.

## 6 Conclusion

The wider media has made a number of conjectures about the impact of generative AI on education, but without any evidence-base to underpin suggestions about how it might change teaching and assessment. This study responds by providing an international and multi-level sample that reveals the ways educators believe teaching and assessment should change as a result of generative AI. These results in turn help the education sector more broadly to acquire a sense of the changes that may need to be made and are poised to come. Moreover, understanding teacher motivations for changing in response to generative AI provides powerful insights that can be used to design professional learning based around those aspects of teaching and technology use that are key drivers for educators.

There was a wide variety in how teacher participants from this study felt that generative AI tools such as ChatGPT will influence education, meaning that there is no one-size-fits-all when it comes to teacher perceptions. Some teachers felt that generative AI will profoundly influence teaching and assessment, and others believed that it will not impact their practices at all. In this study we found that perceptions can vary by awareness of generative AI, teaching experience, teaching level, discipline, region, and gender. These results are a timely reminder of how the impact of technology can differ greatly between contexts, and also people's perceptions, which in turn can be influenced by exposure and familiarity. We should not search for a singular answer as to the impact of generative AI on Education, but rather endeavour to understand its nuanced impact on specific contexts and teachers.

However, there was general consensus from teachers about the ways that both teaching and assessment need to change as a result of generative AI. In general, teachers saw the need to teach students how AI works, how to use AI effectively, critical thinking, ethical values, creativity, collaboration, with greater pedagogical focus on learning processes rather than learning products. Assessment was seen as needing to shift towards more in-person and invigilated tasks that involved greater authenticity, personalisation, higher order thinking, and disclosure of sources. These are all changes that could be seen as pedagogically ideal irrespective of generative AI, and as such, technology may be acting as a valuable catalyst for positive educational change, as it has done on many occasions throughout recent history (Bower, 2017; Matzen & Edmunds, 2007).

The results of this study have a number of implications for a wide range of stakeholders. Teachers can use the results relating to how teaching and assessment needs to change can to inform their practices, for instance, towards critical, ethical and process-focused uses of generative AI in their classes. Designers of professional learning can use the motivational findings to underpin the creation of generative AI workshops and resources, which could be, for example, centred primarily around benefits to student and teacher performance outcomes. Educational leaders can use the results relating to perceived impact to better respond to the variety of perceptions that may be held amongst their staff, from people who may be 'ignorant' to the potential impact of generative AI to experienced users who may be may perceive profound implications. Policymakers can use the findings of this study to inform the design of frameworks for the use of generative AI in education, both in terms of practices that should be applied (e.g. modelling, invigilation) as well as where teachers may need further guidance (e.g. teacher administration and student feedback).

Many aspects of the world around us are set to change in profound ways due to generative AI, meaning that the role of teachers will grow in complexity and importance as they attempt to prepare students for a future that is increasingly difficult to predict. Generative AI will operate on growing array of media, with production of not only high quality text and images, but also video, multimedia, virtual and augmented reality, as well as physical operation through application in robotics. We can expect that generative AI will increase in quality so that it can outperform humans in almost any intellectual or creative task. However, humans will still want and need to be ultimate arbiters for the world in which we live in, and to this extent, it is critical that teachers are well equipped to help students develop the fundamental thinking skills and dispositions that they will need, such as critical thinking, creativity, advanced problem solving, as well as ethical values, volition, and empathy, as a matter of paramount importance. To this extent, an increased focus on how to most effectively develop these underlying capabilities, based on the professional learning of teachers, is a critical area for future research and development.

## Appendix 1 - Survey Questions

What is your gender?

- Male (1)
- Female (2)
- Non-binary / third gender (3)
- Prefer not to say (4)

In what country do you teach?

Name of institution? (optional)

How many years of teaching experience do you have (to the nearest year)?

What level/s do you teach (e.g. Junior High School, University)

What is/are your main teaching/discipline areas? (e.g. physics, philosophy)

Did you watch the presentation/video that accompanies this survey? (Yes / No)

- Yes (1)
- No (2)

Before watching the video, how aware were you of Open Artificial Intelligence (AI) tools such as ChatGPT that could be used by students and teachers to provide answers to general questions?

- Never heard of or seen them (1)
- Knew of them but had not seen or used any first-hand (2)
- Had briefly tried them (3)
- Extensive experience with them (4)



Prior to completing this survey, how aware were you of Open Artificial Intelligence (AI) tools such as ChatGPT that could be used by students and teachers to provide answers to general questions?

- Never heard of or seen them (1)
- Knew of them but had not seen or used any first-hand (2)
- Had briefly tried them (3)
- Extensive experience with them (4)

I believe that open Artificial Intelligence (AI) tools such as ChatGPT:

- Do **not impact** on what I should be **teaching** my students (1)
- Have a **minor impact** on what I should be **teaching** my students (2)
- Have a **major impact** on what I should be **teaching** my students (3)
- Have a **profound impact** on what I should be **teaching** my students (4)

What **changes** to **what you teach** do you believe you should make in response to AI tools such as ChatGPT? (Please elaborate)

What **motivates** you to make changes to **what you will teach** in response to AI tools such as ChatGPT? (Please elaborate)

I believe that open Artificial Intelligence (AI) tools such as ChatGPT:

- Do **not impact** on how I should be **assessing** my students (1)
- Have a **minor impact** on how I should be **assessing** my students (2)
- Have a **major impact** on how I should be **assessing** my students (3)
- Have a **profound impact** on how I should be **assessing** my students (4)

What **changes** to **how you assess** do you believe you should make?  
(Please elaborate)

What **motivates** you to make changes to **how you will assess** in response to AI tools such as ChatGPT? (Please elaborate)

## Appendix 2

**Table 1** Coding Scheme for question: “What changes to what you teach do you believe you should make in response to AI tools such as ChatGPT?”

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Curriculum - Students	Teach students how to use AI	Specific information about how their students will use it for learning (outlining tasks students could integrate AI into in the classroom)	53	0.85	<p>“Allow students with literature challenges to try them out to see if Chat GPT can help them express themselves.” ID22</p> <p>“To learn how to use it to help gather information and to get help on topics. How to devise prompts to get reliable information. Ways to incorporate it for study and writing practices.” ID80</p> <p>“Teaching students how to use it as a tool for learning in appropriate ways” ID105</p>
	Teaching about AI – understanding the tool	Teaching specifically about the tool of AI and understanding what AI (in general) can do	40	1.00	<p>“Undergraduate students need an understanding of AI tools - including what they are, as well as the capabilities and limitations of AI. Undergrads require a critical understanding of the development and use of these tools currently and in the future.” ID02</p>
	Teaching about critical thinking and AI	Comments about specific teaching of critical thinking skills to be applied when using AI	38	0.91	<p>“Students will need to be taught critical thinking skills and how to evaluate AI generated responses for veracity and modification to make them fit for purpose.” ID24</p>
	No change	Commenting that nothing will change with their pedagogy or content	31	0.77	<p>“None” ID74</p> <p>“No changes to teaching” ID206</p>
	Teaching about ethical and responsible use of AI	Specifically teaching about the ethical and responsible use of AI	26	0.95	<p>“How to ethically use them and build on them” ID93</p> <p>“I believe I should continue to prompt my students to explore the ethical dilemmas that are caused by AI. They should explore potential solutions, intended and unintended consequences of those solutions.” ID124</p>
	Using AI responses as a model	Using AI responses in the classroom for modelling purposes	17	1.00	<p>“Provide a sample from ChatGPT and ask students about the strengths and weaknesses of the writing; to rewrite it for improvement.” ID151</p> <p>“I would ask my students to assess the quality of response from ChatGPT and how well or how inaccurate the response is.” ID249</p>
	Tasks that can't be completed by AI	Comments about including tasks that are challenging to complete using AI	11	0.53	<p>“Make sure to focus on problem solving and elaboration (skills that go beyond what bots like ChatGPT can do). I also think we should reconsider the overall goals of education, focusing less on content, and more of developing skills.” ID111</p>
	Teach how to teach using AI	Comments about teaching pre-service education students and university staff how to incorporate AI into their teaching	8	0.98	<p>“...education undergrads also need to be equipped to teach about AI tools in schools in order to develop younger students' digital literacy.” ID02</p> <p>“Support academics to use AI effectively to build skills of learners that is relevant to the courses they teach” ID12</p>
	Discourage use of AI	Commenting on discouraging AI in the classroom	3	1.00	<p>“Reduce use” ID43</p> <p>“Majority of work should be performed in class without internet access” ID193</p>

**Table 1** (continued)

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Teaching -Pedagogy	Changed pedagogical approach	Specific acknowledgement of how their processes/ pedagogical approach may be changed	60	1.00	"More face-to-face and group projects." ID03 "Monitoring progress of tasks through drafting processes/in class observations and interviewing" ID28
	Changed type of activity	Specific acknowledgement of how current activities will be modified (outlining what will change with their usual task)	28	0.85	"Greater focus on the concrete interpretation of the concepts used in class, that is, on the specific concrete referents of these concepts, and the relationships between them." ID13 "Ask for student explanations e.g., video of a code explanation" ID20 "A shift to asking students to justify responses in greater detail or long-form answers and avoiding factual recall or simple explanation questions." ID30
	Supporting own teaching practice	Includes general support of own teaching practice and administration, including using AI for lesson ideas.	23	0.98	"Familiarise personally with ChatGPT prospects, though highly dynamic and updating they may be" ID238 "Make me more convenient to assist me" ID35 "My students are only 9 years old. Rather than make them aware of the capabilities of ChatGPT, I would be interested to see how I could utilise ChatGPT to help me design regular and engaging retrieval practice activities for my class." ID91
Other	Change pedagogy not content	General comments acknowledging their pedagogical approach will need to be modified, however their content will not change.	16	1.00	"I think the content that I teach students will remain the same, but what I teach in terms of how they should consider interacting with the content would change (e.g., how they study etc)." ID01
	Unclear or ambiguous response		16	0.92	"The class is more lively and interesting." ID39 "You can walk into thousands of homes" ID59
	Philosophical response or reflection	Opinions (not discussing how their teaching or content may change)	13	1.00	"In terms of broader teaching strategy, I think the qualitative content ChatGPT produces is poor compared with the depth of analysis and argument students should be producing at a university level. ChatGPT is good at summarising things, good at generating what looks like insight, but is rarely insight. So, ultimately it can be a tool for students but not a replacement for their studies." ID67
	Policy level changes	Comments about policy and procedure at a broad level	11	0.95	"Using AI to provide space to review concepts, including a clear policy/ approach to how it should be used." ID77

### Appendix 3

**Table 2** Coding Scheme for question: “Motivations for changes to what should be taught”

Category	Sub-category	Description	Frequency	Cohen’s Kappa	Examples from participants
Performance Expectancy	Students	Belief that AI will assist in future performance	136	0.98	<p>“To enhance student learning.” ID09</p> <p>“Ensuring students are not missing out on developing important cognition skills.” ID28</p> <p>“Students are going to use them so let’s teach them how to use them in an ethical and efficient way that can be carried them into their careers” ID120</p> <p>“Students need to be made aware of the appropriate integration of how these tools should be used in academia and how they may use these tools in their various professional industries.” ID154</p> <p>“The fact that my students have to be ready for future demands.” ID242</p>
	Teachers		97	0.92	<p>“We need to teach critical thinking, how to get information and knowledge, and how a tool like this can assist them in their learning process.” ID73</p> <p>“The need to benchmark my knowledge again core scientific know-how utilized by AI tools, in order to enhance my ability to share state-of-the-art knowledge with my class.” ID207</p> <p>“They are out there for people to use so freely, we should embrace that” ID84</p> <p>“Chatbots are here” ID133</p> <p>“Response to innovation.” ID100</p>
Facilitating Conditions		Belief that the required technology and organizational infrastructure is available	36	0.90	
Effort Expectancy		Anticipated work to learn and use AI in the classroom	26	1.00	<p>“It can greatly save people’s physical and mental energy” ID58</p> <p>“ChatGPT actually is useful in designing task and helps save me administrative time, this also translates into more time for higher quality activities that sometimes I get ChatGPT to help and inspire me with.” ID136</p> <p>“How these tools can be a time-saver for tedious jobs.” ID166</p> <p>“I can finish my work more efficiently” ID182</p>

**Table 2** (continued)

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Social Influence	Degree to which individuals perceive that others would want them to use AI		24	0.96	<p>“We must be innovative and embrace this new technology tool...” ID81</p> <p>“Technology will be used by students so incorporate what they use or what is common into learning.” ID113</p> <p>“I am preparing teachers for the next several generations of students. The students will have access to tools we do not know of as yet. They need to be able to adapt to the world their students will be inhabiting and their teaching must change with the societal ask for them to be prepared to think critically and contribute to the society in which they live.” ID122</p> <p>“I don’t have a choice. Assignments have to change, and it would be irresponsible of me not to prepare my students to the world that’s to come.” ID127</p> <p>“It’s important to adapt to the changing normal and learn / be prepared for the needs of future learners.” ID160</p> <p>“The need to adapt.” ID225</p> <p>“I’m motivated by the feeling that there’s not really a choice. Evolve or “die.” This genie, Chat GPT, is not going back in the proverbial bottle, and educators need to teach learners to capitalize on the skills humans are best at to prepare them for a future where AI is built into life.” ID248</p> <p>“To follow the trend” ID316</p>
Hedonic Motivation	The fun or pleasure derived from using AI		15	1.00	<p>“Fascinating technology! And it actually has been a veecery long time since I was fascinated by technology...so this was even a bit fun! I have already had lots of interesting interactions with the ChatGPT - seen a lot of the downsides, but still fascinated by the upsides and what it can do.” ID73</p> <p>“A desire to lean into new and innovative ideas rather than shy away from them.” ID75</p> <p>“I love disruptive tools” ID123</p>
Price Value	Perceived benefits and quality of AI compared to price		0	1.00	-

Table 2 (continued)

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Habit		The extent one performs a behaviour automatically	0	1.00	-
Unsure/No change			26	0.97	"Nothing yet as I'm happy with the way I teach and the range of teaching strategies I use." ID318
Unclear			10	1.00	"Guarantee" ID47

**Appendix 4**

**Table 3** Coding framework question: “What changes to how you assess do you believe you should make? (please elaborate)”

Category	Sub-category	Description	Frequency	Cohen’s Kappa	Examples from participants
Assessment task types	Spoken tasks	Presentations	29	0.81	“I will challenge students to share the outputs of their outputs from such tools and to lead real-time discussion on an analysis of strengths and weaknesses.” ID71
	Exam	All forms of exams	21	0.89	“Also require students to write exams in-class” ID170
	Paper and pen approaches	Paper and pencil, paper and pen	7	0.83	“Students need to complete all assessments in an examination situation and in pen to paper.” ID107
	Video production or visual essay	Video making, visual essay	7	1.00	“Ask for student explanations e.g., video of a code explanation.” ID20
	Portfolio	A set of pictures, or photographs of examples of students’ works	3	1.00	“Personal response portfolios such as of artworks.” ID304
Types of content of assessment	Higher-order thinking	Evaluative questions, Critical, logical, analytical thinking assessments	71	0.92	“Students must practice higher-order thinking skills in assessments; assessments should be design to help students practice and demonstrate higher-order thinking skills. (As they always should have been at the university level.)” ID266
	Authentic tasks	Application, scenario-based, Inquiry-based, project-based assessment, Creative type of assessment	69	0.82	“More authentic assessments, more inclusion and student experiences need to be brought into focus.” ID81
	Focus on the process of assessment	Multiple-staged assessments, formative, scaffolding, evidence-based Fewer summative assessments	50	0.94	“A good strategy, I think, would be to make use of more assessments that occur in multiple stages and builds on itself, e.g., have students write outline for an essay, have that outline assessed, and then have an essay that is produced from that outline which is then assessed.” ID13



Table 3 (continued)

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Changing evaluation processes	Personalized tasks	Personal reflections, reflections Personal connections, personal beliefs	35	1.00	"The element of personal reflection and the opportunity to incorporate reflective commentary about prior learning and other lived experiences will be important too, I think." ID10
	Other content-related	Harder contents or Harder exams Minor adjustments, e.g., reworded assessment tasks, more graphs Open-ended question type Tasks related to soft skills	17	0.84	"We will make changes to check for more knowledge - exams will get harder." ID176
	References and citations	Adding references	12	1.00	"Ask students to use & cite sources meticulously, combining different areas of expertise in unique ways while seeking answers to open-ended or reflective questions." ID224
Supervision	Invigilated, face-to-face assessment	More specific rubrics	11	0.86	"I will change the rubric to reflect higher-order thinking and synthesis of knowledge. Students will be asked to compare and provide a scientifically justified choice of the model selected for a specific biological problem." ID272
		Working in class under supervision Invigilated exam	64	0.91	"Most assessments are live and the mc assessments are without internet." ID61
	Detection	Academic integrity, plagiarism software	5	0.88	"All tasks should go through plagiarism software in Junior secondary not just senior secondary." ID171
AI assistance		Using ChatGPT to create their own assessment tasks	54	0.86	"I believe, at this point, we cannot ban students from using AI tools like ChatGPT, so in this case, for any kind of assessment, the quality and innovation to solve a problem can be more important, students can use such tools to provide a correct answer, but not the best one, accordingly, the "SO WHAT" questions can be replaced to traditional assessment methods." ID08

**Table 3** (continued)

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Collaborative & peer-based tasks	Peer assessment tasks	Peer assessment tasks	10	1.00	"Peer assessment." ID09
	Group assessments	Group assessments			"More group assignments." ID03
Minimal or no impact	It does not affect how we assess	It does not affect how we assess	36	1.00	"I don't anticipate making changes to how I assess students. Regardless of whether they use ChatGPT, students will be responsible for meeting the assignment criteria." ID92
	None	None			
Other	Change to assessment strategies (general)	Change to assessment strategies (general)	49	0.90	"Changes in the type of Assessments" ID301
Unsure	I'm still skeptical	I'm still skeptical	17	0.89	"I haven't figured it out yet." ID127 "I'm not really sure. At first blush, I don't see significant impact on how I'll assess my students, but I really need to educate myself further and to think about this some more" ID130

## Appendix 5

**Table 4** Coding framework for question: “What motivates you to make changes to how you will assess in response to AI tools such as ChatGPT? (please elaborate)”

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Performance Expectancy	Students	The belief of the change will improve the performance of teachers or students.	160	0.87	“I am looking for students to develop their own skills and abilities, rather than them leaning on technology to do it for them. I am also mindful of where certain skills will no longer be so valued because ‘a machine can do it’ in the same way we can do mathematic problems ourselves or use a calculator. For some people the calculator really is the more effective option and for others not so much. I want to assess what students can do themselves and within the scope of curriculum what they actually need to be able to demonstrate.” ID173
	Teachers		104	0.82	“Technology can support a higher level of teaching standard and a higher level of researching.” ID230 “As teachers it is important to be able to keep track with our students’ learning and misunderstandings. I’m hoping that by using ChatGPT to custom design some of my assessments that it might help with that.” ID91
Effort Expectancy		The degree to which it’s easy to use of AI for assessment	15	1.00	“It makes assessment much easier.” ID15 “It’s undeniably convenient” ID53
Facilitating Conditions		The degree to which individuals believe that the existing organisational and technical infrastructure can support the use of AI in assessment	12	0.97	“Keeping up with the latest developments in the field/industry.” ID79
Social Influence		The degree to which individuals perceive that others suppose them to use AI in assessment.	10	0.92	“The rapid development and availability of AI tools... there’s nowhere to hide.” ID130
Hedonic Motivation		The perceived pleasure/enjoyment that is derived from the use of AI in assessment	7	1.00	“I’m excited about AI, so the changes I make are motivated by workload.” ID258
Price Value		The perceived benefits of using a particular technology and the cost involved in the use of AI in assessment	0	1.00	-

**Table 4** (continued)

Category	Sub-category	Description	Frequency	Cohen's Kappa	Examples from participants
Habit		Individuals believe the use of AI in assessment to be automatic	0	1.00	-
Not relevant		Answers are not relevant to the question.	27	0.86	"I just wrote a long answer and the website deleted it when I tried to move on" ID 240
Others		Motivations are not classified into UTAUT2 factors.	6	0.84	"Diversification" ID 55
None		No need to change the assessment.	19	1.00	"None" ID74

**Acknowledgements** ChatGPT and generative AI were not used in any way to create this paper.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions.

**Data availability** The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

#### Declarations

This research was approved by the Macquarie University Human Research Ethics Committee (approval ref.: 520231285244798).

**Conflicts of interest** There are no conflicts of interest to declare with this research.

**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

- Akgun, S., & Greenhow, C. (2021). Artificial intelligence in education: Addressing ethical challenges in K-12 settings. *AI and Ethics*, 2, 431–440. <https://doi.org/10.1007/s43681-021-00096-7>
- Almalki, S. (2016). Integrating quantitative and qualitative data in mixed methods research - challenges and benefits. *Journal of Education and Learning*, 5(3), 288–296.
- Baig, M. I., Shuib, L., & Yadegaridehkordi, E. (2020). Big data in education: A state of the art, limitations, and future research directions. *International Journal of Educational Technology in Higher Education*, 17(1), 1–23. <https://doi.org/10.1186/s41239-020-00223-0>
- Baker, T., Smith, L., & Anissa, N. (2019). Educ-AI-tion rebooted? Exploring the future of artificial intelligence in schools and colleges. *Nesta Foundation*. [https://media.nesta.org.uk/documents/Future\\_of\\_AI\\_and\\_education\\_v5\\_WEB.pdf](https://media.nesta.org.uk/documents/Future_of_AI_and_education_v5_WEB.pdf). Accessed 6 Dec 2023
- Bower, M. (2017). *Design of technology-enhanced learning: Integrating research and practice*. Emerald Publishing Limited.
- Bower, M., DeWitt, D., & Lai, J. W. M. (2020). Reasons associated with preservice teachers' intention to use immersive virtual reality in education. *British Journal of Educational Technology*, 51(6), 2215–2233. <https://doi.org/10.1111/bjet.13009>
- Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., & Askell, A. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901.
- Carvalho, L., Martinez-Maldonado, R., Tsai, Y. S., Markauskaite, L., & De Laat, M. (2022). How can we design for learning in an AI world? *Computers and Education: Artificial Intelligence*, 3, 100053. <https://doi.org/10.1016/j.caeai.2022.100053>
- Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, 107468. <https://doi.org/10.1016/j.chb.2022.107468>
- Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, 66(4), 616–630. <https://doi.org/10.1007/s11528-022-00715-y>

- Chávez Herting, D., Pros, C., & Castelló Tarrida, A. (2020). Patterns of PowerPoint use in higher education: A comparison between the natural, medical, and social sciences. *Innovative Higher Education*, 45, 65–80. <https://doi.org/10.1007/s10755-019-09488-4>
- Chen, X., Xie, H., Zou, D., & Hwang, G. J. (2020). Application and theory gaps during the rise of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2023a). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*, 1–17. <https://doi.org/10.1080/10494820.2023.2172044>
- Chiu, T. K. F., Xia, Q., Zhou, X., Chai, C. S., & Cheng, M. (2023b). Systematic literature review on opportunities, challenges, and future research recommendations of artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 4, 100118. <https://doi.org/10.1016/j.caeai.2022.100118>
- Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, 32(3), 725–755. <https://doi.org/10.1007/s40593-021-00243-5>
- Churchill, R., Ferguson, P., Godinho, S., Johnson, N., Keddie, A., Letts, W., McGill, M., MacKay, J., Moss, J., & Nagel, M. (2016). *Teaching: Making a difference*. John Wiley and Sons Ltd.
- Clarke, V., & Braun, V. (2013). *Successful qualitative research: A practical guide for beginners*. Sage publications.
- Clarivate Analytics (2023). *Web of Science Research Domains*[https://images.webofknowledge.com/images/help/WOK/hs\\_research\\_domains.html](https://images.webofknowledge.com/images/help/WOK/hs_research_domains.html). Accessed 6 Dec 2023
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Cope, B., Kalantzis, M., & Searnsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory*, 53(12), 1229–1245. <https://doi.org/10.1080/00131857.2020.1728732>
- Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. <https://doi.org/10.1080/14703297.2023.2190148>
- Dai, Y., Chai, C. S., Lin, P. Y., Jong, M. S. Y., Guo, Y., & Qin, J. (2020). Promoting students' well-being by developing their readiness for the artificial intelligence age. *Sustainability*, 12(16), 6597. <https://doi.org/10.3390/su12166597>
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*. <https://arxiv.org/pdf/1810.04805.pdf&usq=ALkJrhxzlC6yTht2BRmH9atgvKFxHsxQ>. Accessed 6 Dec 2023
- Escueta, M., Quan, V., Nickow, A. J., & Oreopoulos, P. (2017). Education technology: An evidence-based review. *National Bureau of Economic Research Working Paper No. 23744*. <https://www.nber.org/papers/w23744>. Accessed 6 Dec 2023
- Ferguson, C., van den Broek, E. L., & van Oostendorp, H. (2022). AI-Induced Guidance: Preserving the Optimal Zone of Proximal Development. *Computers and Education: Artificial Intelligence*, 3, 100089. <https://doi.org/10.1016/j.caeai.2022.100089>
- Hattie, J. (2023). *Visible learning: The sequel: A synthesis of over 2,100 meta-analyses relating to achievement*. Taylor and Francis.
- Hisan, U. K., & Amri, M. M. (2023). ChatGPT and medical education: A double-edged sword. *Journal of Pedagogy and Education Science*, 2(01), 71–89. <https://doi.org/10.56741/jpes.v2i01.302>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. The Center for Curriculum Redesign.
- Kangasharju, A., Ilomäki, L., Lakkala, M., & Toom, A. (2022). Lower secondary students' poetry writing with the AI-based poetry machine. *Computers and Education: Artificial Intelligence*, 3, 100048. <https://doi.org/10.1016/j.caeai.2022.100048>
- Khosravi, H., Shum, S. B., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Kim, J., Lee, H., & Cho, Y. H. (2022). Learning design to support student-AI collaboration: Perspectives of leading teachers for AI in education. *Education and Information Technologies*, 27(5), 6069–6104. <https://doi.org/10.1007/s10639-021-10831-6>

- Lai, J. W., & Bower, M. (2019). How is the use of technology in education evaluated? A systematic review. *Computers & Education*, *133*, 27–42. <https://doi.org/10.1016/j.compedu.2019.01.010>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.
- Luckin, R., Cukurova, M., Kent, C., & du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, *3*, 100076. <https://doi.org/10.1016/j.caeai.2022.100076>
- Markauskaite, L., Marrone, R., Poquet, O., Knight, S., Martinez-Maldonado, R., Howard, S., Tondeur, J., De Laat, M., Shum, S. B., Gašević, D., & Siemens, G. (2022). Rethinking the entwinement between artificial intelligence and human learning: What capabilities do learners need for a world with AI? *Computers and Education: Artificial Intelligence*, *3*, 100056. <https://doi.org/10.1016/j.caeai.2022.100056>
- Matzen, N. J., & Edmunds, J. A. (2007). Technology as a catalyst for change: The role of professional development. *Journal of Research on Technology in Education*, *39*(4), 417–430.
- Merritt, R. (2022). What is a transformer model. *Nvidia Corporation*. <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/>. Accessed 6 Dec 2023
- Miller, F. A., Katz, J. H., & Gans, R. (2018). The OD imperative to add inclusion to the algorithms of artificial intelligence. *OD Practitioner*, *50*(1), 8.
- Mollman, S. (2023). ChatGPT passed a Wharton MBA exam and it's still in its infancy. One professor is sounding the alarm. *Fortune*. Retrieved January 22 2023, from <https://fortune.com/2023/01/21/chatgpt-passed-wharton-mba-exam-one-professor-is-sounding-alarm-artificial-intelligence/>
- Nardi, P. M. (2018). *Doing survey research: A guide to quantitative methods*. Routledge.
- OpenAI (2023). *Introducing ChatGPT*. <https://openai.com/blog/chatgpt>. Accessed 6 Dec 2023
- Ouyang, F., & Jiao, P. (2021). Artificial intelligence in education: The three paradigms. *Computers and Education: Artificial Intelligence*, *2*, 100020. <https://doi.org/10.1016/j.caeai.2021.100020>
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., & Ray, A. (2022). Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, *35*, 27730–27744.
- Pérez-Sanagustín, M., Nussbaum, M., Hilliger, I., Alario-Hoyos, C., Heller, R. S., Twining, P., & Tsai, C. C. (2017). Research on ICT in K-12 schools—A review of experimental and survey-based studies in computers & education 2011 to 2015. *Computers & Education*, *104*, A1–A15. <https://doi.org/10.1016/j.compedu.2016.09.006>
- Popenici, S. A. D., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, *12*(1), 1–13. <https://doi.org/10.1186/s41039-017-0062-8>
- Roose, K. (2023). Don't ban ChatGPT in schools - Teach with it. *The New York Times*. Retrieved 12th Jan 2023, from <https://www.nytimes.com/2023/01/12/technology/chatgpt-schools-teachers.html>
- Schiff, D. (2021). Out of the laboratory and into the classroom: The future of artificial intelligence in education. *AI & Society*, *36*(1), 331–348. <https://doi.org/10.1007/s00146-020-01033-8>
- Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J. M., Milligan, S., Selwyn, N., & Gašević, D. (2022). Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, *3*, 100075. <https://doi.org/10.1016/j.caeai.2022.100075>
- Tang, K. Y., Chang, C. Y., & Hwang, G. J. (2021). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, 1–19. <https://doi.org/10.1080/10494820.2021.1875001>
- Touretzky, D., Gardner-McCune, C., Martin, F., & Seehorn, D. (2019). Envisioning AI for K-12: What should every child know about AI? *Proceedings of the AAAI Conference on Artificial Intelligence*. <https://doi.org/10.1609/aaai.v33i01.33019795>
- Tseng, T. H., Lin, S., Wang, Y. S., & Liu, H. X. (2022). Investigating teachers' adoption of MOOCs: The perspective of UTAUT2. *Interactive Learning Environments*, *30*(4), 635–650. <https://doi.org/10.1080/10494820.2019.167488>
- UNESCO. (2019). *Beijing consensus on artificial intelligence and education*. UNESCO.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*, 1–11.

- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157–178.
- Venkatesh, V., Thong, J. Y. L., & Xu, X. (2016). Unified theory of acceptance and use of technology: A synthesis and the road ahead. *Journal of the Association for Information Systems*, 17(5), 328–376.
- Wang, P. (2019). On defining artificial intelligence. *Journal of Artificial General Intelligence*, 10(2), 1–37. <https://doi.org/10.2478/jagi-2019-0002>
- Wang, X., Liu, Q., Pang, H., Tan, S. C., Lei, J., Wallace, M. P., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, 194, 104703. <https://doi.org/10.1016/j.compedu.2022.104703>
- Wingard, J. (2023). ChatGPT: A threat to higher education? *Forbes*. Retrieved 10th Jan 2023, from <https://www.forbes.com/sites/jasonwingard/2023/01/10/chatgpt-a-threat-to-higher-education>
- Xu, W., & Ouyang, F. (2022). A systematic review of AI role in the educational system based on a proposed conceptual framework. *Education and Information Technologies*, 27, 4195–4223. <https://doi.org/10.1007/s10639-021-10774-y>
- Zawacki-Richter, O., Marin, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(39). <https://doi.org/10.1186/s41239-019-0171-0>
- Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J. B., Yuan, J., & Li, Y. (2021). A review of Artificial Intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021, 1–18. <https://doi.org/10.1155/2021/8812542>

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

## Authors and Affiliations

Matt Bower<sup>1</sup>  · Jodie Torrington<sup>1</sup>  · Jennifer W. M. Lai<sup>1</sup>  · Peter Petocz<sup>2</sup>  · Mark Alfano<sup>3</sup> 

✉ Jennifer W. M. Lai  
jennifer.lai@mq.edu.au

Matt Bower  
matt.bower@mq.edu.au

<sup>1</sup> Macquarie School of Education, Macquarie University, Sydney, NSW 2109, Australia

<sup>2</sup> Graduate Research Academy, Macquarie University, Sydney, NSW 2109, Australia

<sup>3</sup> Department of Philosophy, Macquarie University, Sydney, NSW 2109, Australia