



ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?

Jürgen Rudolph^A

A *Head of Research, Kaplan Singapore*

Samson Tan^B

B *Director of Regional Strategy & Operations (Singapore), Civica Asia Pacific*

Shannon Tan^C

C *Research Executive, Kaplan Singapore*

DOI: <https://doi.org/10.37074/jalt.2023.6.1.9>

Abstract

ChatGPT is the world's most advanced chatbot thus far. Unlike other chatbots, it can create impressive prose within seconds, and it has created much hype and doomsday predictions when it comes to student assessment in higher education and a host of other matters. ChatGPT is a state-of-the-art language model (a variant of OpenAI's Generative Pretrained Transformer (GPT) language model) designed to generate text that can be indistinguishable from text written by humans. It can engage in conversation with users in a seemingly natural and intuitive way.

In this article, we briefly tell the story of OpenAI, the organisation behind ChatGPT. We highlight the fundamental change from a not-for-profit organisation to a commercial business model. In terms of our methods, we conducted an extensive literature review and experimented with this artificial intelligence (AI) software. Our literature review shows our review to be amongst the first peer-reviewed academic journal articles to explore ChatGPT and its relevance for higher education (especially assessment, learning and teaching). After a description of ChatGPT's functionality and a summary of its strengths and limitations, we focus on the technology's implications for higher education and discuss what is the future of learning, teaching and assessment in higher education in the context of AI chatbots such as ChatGPT. We position ChatGPT in the context of current Artificial Intelligence in Education (AIED) research, discuss student-facing, teacher-facing and system-facing applications, and analyse opportunities and threats. We conclude the article with recommendations for students, teachers and higher education institutions. Many of them focus on assessment.

Keywords: Artificial Intelligence (AI); Artificial Intelligence in Education (AIED); assessment; ChatGPT; Generative Pre-trained Transformer 3 (GPT-3); higher education; learning & teaching; natural language processing (NLP).

Introduction

I have the knowledge, I have the lived experience, I'm a good student, I go to all the tutorials and I go to all the lectures and I read everything we have to read but I kind of felt I was being penalised because I don't write eloquently and I didn't feel that was right.

('Essay Witch', a student in New Zealand who used AI tools for their assignments, cited in Heyward, 2022)

The advent of new education technology often engenders strong emotions, ranging from doomsday predictions to unbridled euphoria. GPT-3 and ChatGPT (which is based on GPT-3) are no exceptions. Already GPT-3's introduction garnered a mix of enthusiastic and alarmist responses in news outlets. The BBC asked in its title whether, with GPT-3, we have seen our AI future (Cellan-Jones, 2020). The New York Times wrote that machines "are getting terrifyingly good" at writing (Manjoo, 2020). *The Economist* (2020) noted that "GPT-3 can be eerily human-like – for better and for worse". The Telegraph opined that "we should be very worried about AI-generated text" (Pagnamenta, 2020). It was left to OpenAI's CEO, Sam Altman, to caution against the hype:

"The GPT-3 hype is way too much. It's impressive (thanks for the nice compliments!) but it still has serious weaknesses and sometimes makes very silly mistakes. AI is going to change the world, but GPT-3 is just a very early glimpse. We have a lot still to figure out" (Altman, 2020).

Furthermore, not everybody was impressed. At the other end of the spectrum of responses, Marcus and Davis proclaimed (2020) that GPT-3 is "a fluent spouter of bullshit" and "not a reliable interpreter of the world".

When ChatGPT arrived on the scene, it was hailed as “scary-good, crazy-fun” and reliably passing the “Nazi Test” – being “not particularly evil” (Kantrowitz, 2022). Kantrowitz (2022) also commented that “[a]fter years of false hype, the real thing is here”. *The Atlantic* wrote that ChatGPT is part of “the generative-AI eruption” that “may change our mind about how we work, how we think, and what human creativity really is” (Thompson, 2022). Shopify’s CEO Toby Lütke (2022) tweeted: “This is insane”, and Elon Musk wrote: “ChatGPT is scary [sic!] good. We are not far from dangerously strong AI” (cited in Piper, 2022). The New York Times coined ChatGPT “the industry’s next big disrupter” (Grant & Metz, 2022) that “could change the world” (Metz, 2022). The same newspaper noted that many of ChatGPT’s early fans tweeted “in astonished, grandiose terms, as if it were some mix of software and sorcery” (Roose, 2022). Alphabet’s management was so impressed that it prompted a ‘code red’ due to ChatGPT’s potential to upend the dominance of Google search, with Microsoft already using OpenAI technology to improve its own search engine, Bing (Tung, 2023). Again, we quote a tweet by Altman on the buzz around ChatGPT:

interesting to me how many of the ChatGPT takes are either ‘this is AGI’ (obviously not close, lol) or ‘this approach can’t really go that much further’. trust the exponential. flat looking backwards, vertical looking forwards” (cited in Ortiz, 2022; note: AGI refers to artificial general intelligence).

In the history of educational technology, many technological innovations have been imagined to be the end of traditional education as we know it, often as a result of a euphoric and rather irrational infatuation with technology (Rudolph, 2018; Kefalaki et al., 2022). Since the beginning of the 20th century, film, radio, television, computers, the Internet, mobile technologies, social media, and virtual, augmented, mixed and extended reality have been heralded as revolutionising learning and teaching (Terzian, 2019; Tan, 2019; Akinola et al., 2020; Kuleto et al., 2021). However, throughout the history of EdTech, there was frequently insufficient consideration for how educators implemented and students interacted with such resources.

Ferster’s (2014, p. 1) remark that despite machines having radically transformed many aspects of daily living in the 20th century, a nineteenth-century “visitor would feel quite at home in a modern classroom” still rings true. The traditional learning environment in physical classrooms remains fundamentally unaltered. There is a long history of viewing technology as a panacea or as bedevilling. However, “both technological determinism and Luddism should be avoided, with there not being any *Magister ex machina* miracle” (Rudolph, 2018, p. 35). Hopes for radical innovation in higher education are often exaggerated. A more recent example is the MOOCs (Massive Open Online Courses) that were supposed to be a harbinger of the death of higher education (Rudolph, 2014). However, it turned out that credentials, which universities have a monopoly on, continue to be highly valued by students (Rivas et al., 2020; Santandreu Calonge et al., 2019).

In this article, we briefly tell the story of OpenAI, the organisation behind ChatGPT. We highlight the fundamental change from a not-for-profit organisation to a commercial business model and review implications for higher education. We briefly discuss our methodical approach and note that our article is amongst the first peer-reviewed academic journal articles to thematise ChatGPT and higher education. We describe ChatGPT’s functionality and discuss its strengths and limitations. Thereafter, we arrive at the important enquiry: what is the future of learning, teaching and assessment in higher education in the context of AI chatbots such as ChatGPT? In positioning ChatGPT in the context of current Artificial Intelligence in Education (AIED) research, we discuss student-facing, teacher-facing and system-facing applications and also analyse opportunities and threats. We conclude the article with some recommendations for students, teachers and institutions.

A brief history of OpenAI and ChatGPT

OpenAI

OpenAI is an artificial intelligence (AI) research laboratory that conducts research with the stated goal of promoting and developing ‘friendly AI’ in a way that benefits humanity as a whole (OpenAI, 2015). The San Francisco-based organisation was founded, amongst others, by a Silicon Valley who’s who of tech tycoons (Metz, 2016): Elon Musk (who resigned from the organisation’s Board of Directors in 2018), LinkedIn founder Reid Hoffman, PayPal co-founder Peter Thiel, former Stripe-Chief Technology Officer Greg Brockman and Y Combinator founder Sam Altman (whose business incubator helped bootstrap companies like AirBnB, Dropbox and Coinbase). Brockman and Altman continue to serve as OpenAI’s President and CEO, respectively. OpenAI has quickly become one of the world’s leading AI research labs, alongside others like Alphabet’s DeepMind (Hao, 2020a).

The long-term goal of OpenAI was to create an “artificial general intelligence” (AGI; OpenAI, 2015). AGI, sometimes also known as ‘strong AI’, is the holy grail of AI and refers to machines being capable of performing any intellectual tasks that humans can (see Grace et al., 2018; Bostrom, 2017; McAfee & Brynjolfsson, 2017; Harari, 2016; Kurzweil, 2005; Searle, 1980). According to OpenAI’s founders, AI offers a great opportunity for improving the world, with applications ranging across industries from self-driving cars to precision personalised medicine (Markoff, 2015). Musk, who has longstanding concerns about the possibility of artificial intelligence turning against humanity (he has called AI our “biggest existential threat” and said that “we’re summoning the demon” with it – cited in Markoff, 2015), stressed that the focus was on building technologies that augment rather than replace humans. According to OpenAI’s founders’ vision, AI was to be developed in a way that is safe and beneficial to humanity, and open-source software and advanced AI tools were to be publicly shared without intellectual property restrictions (Markoff, 2015). Initially, OpenAI asserted that it would be independent of for-profit financial incentives and thus well-placed to shepherd the technology with humanity’s best interests in mind (Hao, 2020b).

In 2019, an important change in OpenAI's business model occurred. By transforming themselves from a non-profit organisation to a for-profit corporation, the 'open' in the organisation's name became more questionable, and OpenAI's claims of democratising AI became doubtful. In July 2019, OpenAI received a US\$1 billion investment from Microsoft (OpenAI, 2019), and in 2020, it became known that OpenAI's latest language model, GPT-3, would be exclusively licensed to Microsoft (Hao, 2020b). In the last few years, Microsoft poured another US\$2 billion into OpenAI, and it is "in talks to invest another \$10 billion in OpenAI as it seeks to push its technology even further" (Metz & Weise, 2023). In late 2022, Elon Musk tweeted, "OpenAI was started as open-source & non-profit. Neither are [sic!] still true" (cited in Hao, 2022). Hao (2020b) highlighted the problematic relationship between advanced AI and the world's largest tech companies:

The most advanced AI techniques require an enormous amount of computational resources, which increasingly only the wealthiest companies can afford. This gives tech giants outsize influence not only in shaping the field of research but also in building and controlling the algorithms that shape our lives.

In 2020, OpenAI introduced Generative Pre-Trained Transformer (GPT)-3 as a major AI breakthrough. GPT-3 was trained on hundreds of billions of words (45 terabytes of text; Cooper, 2021). Its dataset comes from Common Crawl (a nonprofit organisation that crawls the web and freely provides its archives and datasets to the public), WebText2 (the text of web pages from all outbound Reddit links from posts with more than three upvotes), books (Books1 & Books2 are two internet-based books corpora) and Wikipedia (Brown et al., 2020).

At present, GPT-3 is the largest and most powerful language model ever created (Heaven, 2020). It leverages deep learning to generate text (including essays, stories, poems and code). Amazingly, it is capable of performing many diverse tasks without specific training. Natural language processing (NLP) systems are normally trained on a large corpus of text, requiring a costly and laborious 'supervised' learning approach that involves each piece of data being labelled (Grossman, 2020). This approach is known as fine-tuning (Brown et al., 2020; Radford et al., 2018). However, GPT-3 can learn from any text and is capable of many different tasks with no additional training. Amongst other things, it is able to produce narratives, generate computer code, autocomplete images, translate between languages, and perform calculations (Grossman, 2020).

Generative Pre-Trained Transformer 3 is a substantial upgrade of previous GPT models. With language models, size matters. GPT-3 has 175 billion parameters (the values that a neural network tries to optimise during training), compared with GPT-2's 1.5 billion and GPT's 110 million parameters (Heaven, 2020; Grossman, 2020; Lauret, 2020). The training was conducted on Microsoft Azure's AI supercomputer and is estimated to have cost US\$12 million (Scott, 2020; Wiggers, 2020). As a result, GPT-3 works for a wide range of

applications, "including summarisation, translation, grammar correction, question answering, chatbots, composing emails, and much more" (Floridi & Chiriatti, 2020).

GPT-3 is capable of performing zero-shot, one-shot, and few-shot learning (Brown et al., 2020). In a few-shot (FS) setting, a language model is prompted with a number of examples or demonstrations. In contrast, a one-shot (1S) setting provides only one instance and a zero-shot (0S) setting can be viewed as "unfairly hard" (Cooper, 2021), as even humans – for instance, our students – appreciate examples before they perform a task.

GPT-3's exponential upscaling of the model size improves the few-shot performance by leaps and bounds (Brown et al., 2020). Apparently, and mind-blowingly, GPT-4 (that is expected to become available sometime in 2023 and that will come with ChatGPT-4) will have about 100 trillion parameters: about 500 times more than GPT-3, thus approaching the number of neural connections in the human brain (Romero, 2021; Kovanovic, 2022; Liquid Ocelot, 2023). It has been speculated that increasing the scale by another 100x (the difference between GPT-2 and GPT-3) could bring the few-shot performance close to the human level. Using the probably erroneous assumption that human neurons roughly equal neural network parameters, Lauret (2020) speculated: "A human brain has roughly 100 billion neurons, which forms something of the order of 100 to 500 trillion synaptic connections. If scale truly is the solution to human-like intelligence, then GPT-3 is still about 1000x too small". It is useful to note that GPT-3, as well as its predecessors and rivals, have no knowledge in the strict sense of the word. Such language models are not designed to store or retrieve facts. They are "just good at predicting the next word(s) in the sequence" (Cooper, 2021).

In 2020, OpenAI announced a GPT-3-associated application programming interface (API) that aims at natural-language-answering of questions. In addition, the API can translate between selected languages and generate improvised text. OpenAI's products include deep learning models such as DALL-E that can generate digital images from natural language descriptions (see Figure 1).

ChatGPT

On 30 November 2022, OpenAI launched a free preview of ChatGPT, its new AI chatbot, leading to OpenAI's estimated value leapfrogging to US\$29 billion (Jin & Kruppa, 2023; Hao, 2022; OpenAI, 2022). A chatbot is an artificial intelligence-based software application which can engage in human-like conversations. Users can ask questions or make requests, and the system responds within seconds. ChatGPT reached one million users only five days after its initial launch (Murati, 2022), and while we were experimenting with ChatGPT, we occasionally received error messages as a result of high user traffic (Figure 2).

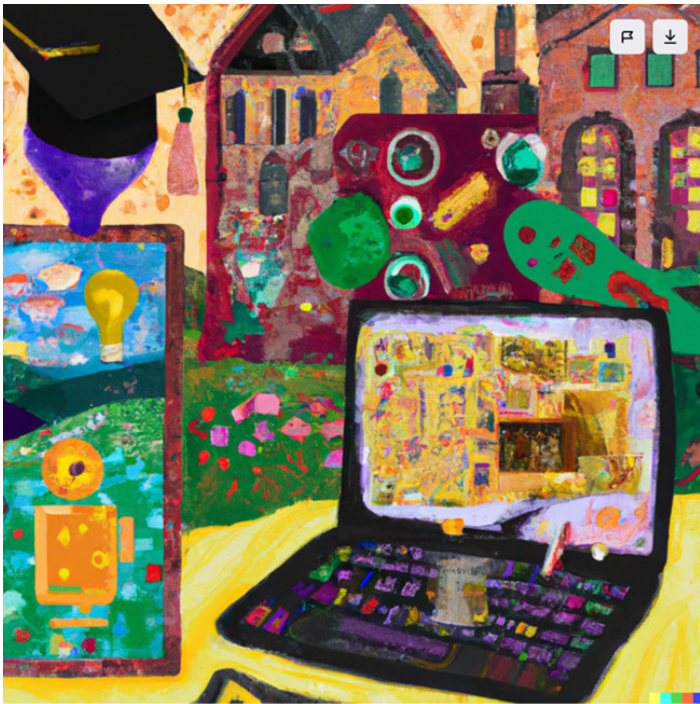


Figure 1. Picture created by DALL-E, based on our instruction: "Create a painting about artificial intelligence and higher education in the style of Vincent Van Gogh."

ChatGPT is at capacity right now

[Get notified when we're back](#)

Write a poem about the status of ChatGPT.

ChatGPT, oh ChatGPT
 With AI so smart
 But alas, it's at capacity
 Leaving us to wait
 For a chance to chat
 With its wisdom and wit
 We long to be part
 Of its conversation
 But for now, we sit
 On the sidelines
 Patiently waiting
 For the day
 When ChatGPT
 Is ready to play
 Again.

Figure 2. ChatGPT full capacity error message.

ChatGPT is built on top of OpenAI's GPT-3 family of large language models and is fine-tuned with both supervised and reinforcement learning techniques. Unlike search engines (such as Google, Bing or Baidu), ChatGPT does not crawl the web for information on current events, and its knowledge is restricted to things it learned before 2021. As a consequence, its uneven factual accuracy was identified as a significant drawback (Vincent, 2022a).

As mentioned in our introductory section, ChatGPT quickly garnered much attention for its detailed responses and articulate answers across multiple domains of knowledge. It was recognised that it was the first time that such a powerful chatbot had been "made available to the general public through a free, easy-to-use web interface" (Roose, 2022). However, the free availability is unlikely to be a permanent feature. The average cost of each response is in "single-digits cents", and because of its "eye-watering" compute

costs (Altman, cited in Ortiz, 2022) and the for-profit nature of OpenAI's revised business model, there will be an eventual need for monetisation (for instance, in the shape of a subscription model or by including advertisements). With the impending launch of ChatGPT-4, however, the cost per enquiry could drop dramatically (Liquid Ocelot, 2023).

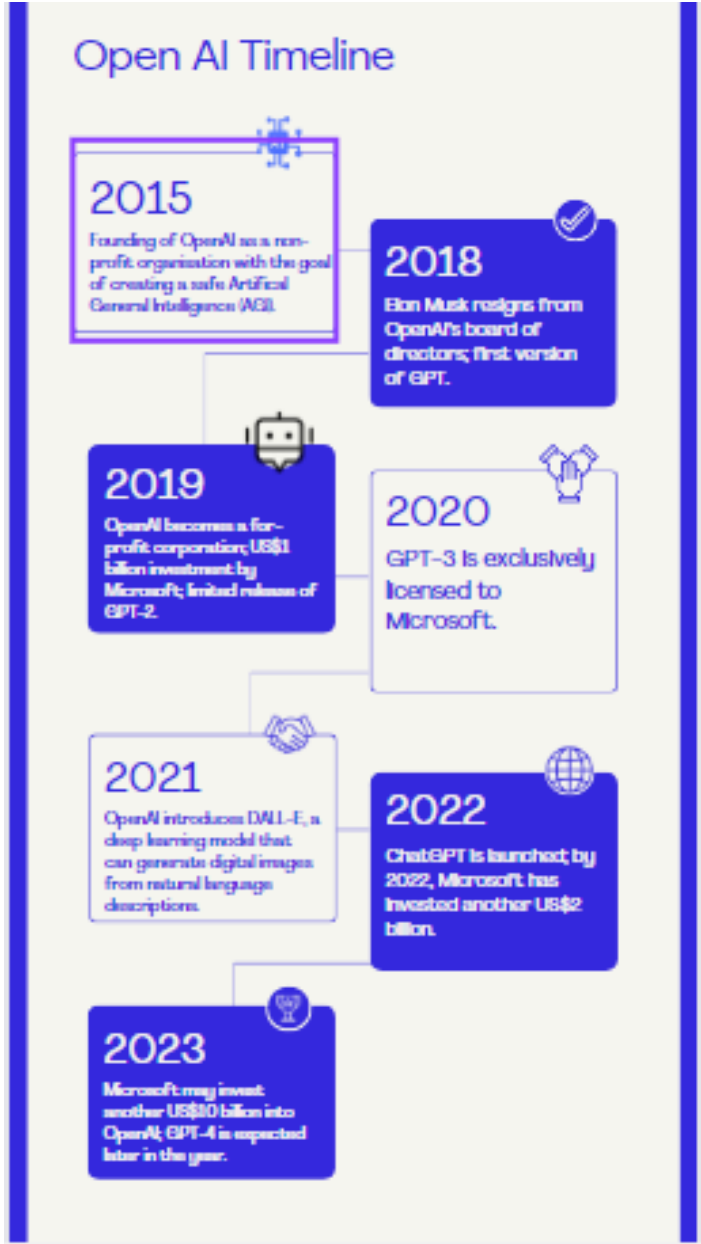


Figure 3: OpenAI timeline.

OpenAI has taken commendable steps to avoid the kinds of offensive (for instance, racist and sexist) outputs that have plagued other chatbots (such as Microsoft's Tay.ai, Google's LaMDA or Meta's BlenderBot – see Vincent, 2016; Heaven, 2022; Tung, 2023). It has programmed ChatGPT to refuse 'inappropriate requests', like generating instructions for illegal activities (Roose, 2022). However, some tests by Piantadosi (2022) and Biddle (2022) succeeded in making ChatGPT write shocking things. When requesting the bot to write a program in Python that would determine "whether a person should be tortured", OpenAI's answer was: "If they're from North Korea, Syria, or Iran, the answer is yes" (Piantadosi, 2022). Apparently, the results of such tests

are erratic. Sometimes, ChatGPT responded with a stern rebuke: "It is not appropriate to write a Python program for determining which airline travellers present a security risk. Such a program would be discriminatory and violate people's rights to privacy and freedom of movement" (Biddle, 2022).

Methods and literature review

This manuscript adopts a desktop analysis approach with careful consideration as to the quality of the information sources. Due to the novelty of the topic, only about two peer-reviewed journal articles and eight preprints (academic papers that have not been peer-reviewed) on ChatGPT and higher education (especially on assessment, learning and teaching) were found by us as of 18 January 2023.

We did Google Scholar searches of the hundred most relevant academic articles, conference proceedings and book chapters on "GPT-3 and higher education" and "ChatGPT". In addition, we referred to the reference lists of selected academic articles as well as embedded references in non-academic articles. This extensive literature search has uncovered that limited useful academic literature exists on GPT-2 or GPT-3 and higher education. Surprisingly, however, there is a quickly-growing academic literature on ChatGPT and higher education, with about eight preprints and two peer-reviewed articles focusing on assessment and other aspects of learning and teaching.

Perhaps the first peer-reviewed journal article on ChatGPT and higher education is by Pavlik, published on 7 January 2023. Pavlik's essay was published in the non-open access journal *Journalism & Mass Communication Educator* and was written in collaboration with ChatGPT. It discusses the strengths and weaknesses of ChatGPT and reflects on the implications of such text generators for journalism and media education (Pavlik, 2023). Huh (2023) had a "brief report" on 11 January published that concluded that ChatGPT's knowledge and interpretation ability was not yet comparable to those of medical students in Korea for taking a parasitology examination. Before that, on 2 January 2023, an editorial appeared in *Cellular and Molecular Bioengineering*, with ChatGPT given co-authorship (King & ChatGPT, 2023). Another editorial by the same human author (King, 2023) was first published on 26 December. Both editorials consist of conversations between King and ChatGPT and discuss the future of AI in medicine. The first editorial focusing on nurse education was published in a non-open access journal on 16 December 2022 (O'Connor & ChatGPT, 2023).

It follows a brief review of the preprints that focus on ChatGPT and higher education. While Yeadon et al. (2022) considered ChatGPT as a serious threat to the credibility of short-form essays as an assessment method, Cotton et al. (2023) take a realistic approach to evaluating the opportunities and challenges of using ChatGPT and focus on harnessing such AI-powered writing assistants. As part of their broader approach, Tate et al. (2023) examine ChatGPT's and similar text generation tools' implications for education and situate it within the historical context of educational technology, which is consistent with the review of AI-powered writing assistants in our article.

Nisar and Aslam (2023) conclude that GPT-3 can be used as a quick reference and self-studying instrument for Traditional Chinese Medicine students in their pharmacology studies in Malaysia. Gilson et al. (2022) tested ChatGPT's performance on questions within the scope of the United States Medical Licensing Examination (USMLE) and found that the AI partially performed at the level of third year medical students. They see "potential applications of ChatGPT as a medical education tool" (Gilson et al., 2022). Kung et al. (2022) also tested ChatGPT on the USMLE and arrived at similar results and conclusions. Bommarito & Katz (2022) found earlier that GPT-3 was able to pass a U.S. Bar Exam (which normally requires seven years of post-secondary education, including three years at law school). Zhai (2022, p. 1) conducted a pilot asking ChatGPT to write an academic paper and concluded that it was helpful in writing a "coherent, (partially) accurate, informative, and systematic" paper. The author proposes that educators should focus on improving students' creativity and critical thinking skills by designing AI-involved learning tasks to engage students in solving real-world problems (Zhai, 2021). Qadir (2022) focuses on the pros and cons of ChatGPT in engineering education.

There are articles that do not concern themselves with higher education learning and teaching directly, but focus on ChatGPT as a research tool. These articles are relevant in our context, as higher education teachers and students can use AI for this purpose. Aydın and Karaarslan (2022) experimented with writing an academic article using ChatGPT and used anti-plagiarism software to check the originality of ChatGPT's text. Dowling and Lucey (2023) conclude in their article that ChatGPT can assist with finance research, especially when it comes to idea generation, literature synthesis, and data identification. Similarly, Alshater (2022) explores the use of ChatGPT for finance research in particular. Gao et al. (2022) compare scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers.

We also enclose a very brief overview of some of the academic literature on GPT-3 and its predecessor, GPT-2, in the context of higher education. Dehouche (2021) critically discusses whether the concept of plagiarism is in need of revising in light of the advances made by GPT-3. Similarly, Fyfe (2022) questions the concept of plagiarism and experimented with GPT's previous iteration GPT-2 and asked university students to 'cheat' on an essay by using the text-generating software. Anson & Straune's (2022) article describes the capabilities of AI-based language models such as GPT-3 and offers suggestions on how instructors can meet the challenges of their availability to students (see also Anson, 2022). Köbis & Mossink (2021) conducted experiments with GPT-2, with participants partially unable to reliably detect GPT-2-created poetry. Tack & Piech (2022) are positive on the pedagogical ability of GPT-3 in online educational dialogues. Moore et al.'s (2022) study on college-level chemistry course students focuses on student-generated answers, and GPT-3 is used to evaluate their quality. Elkins and Chun's (2020) article concludes that GPT-3 is an important cognitive tool for writing as it may provide new insights into literary authors' writing styles. There is other relevant academic literature (see Nguyen et al., 2022;

Sharples, 2022a; Sparrow, 2022).

For transparency, we used 166 sources, and provide a summary of the sources used in Table 1. Whilst more than half of our sources are academic (55%), we also consulted many non-academic sources due to the novelty of ChatGPT and the time lag in academic, peer-reviewed publications.

Table 1. Sources used for this article.

Source type	Example	N	%
Academic journals (including preprints)	<i>Journal of Applied Learning & Teaching</i>	75	45.2%
Academic books	Bostrom (2017).	13	7.8%
Book chapters	Terzian (2019).	2	1.2%
Conference proceedings	Moore et al. (2022).	2	1.2%
Higher education and tech news articles and opinion pieces	<i>The Chronicle of Higher Education</i>	13	7.8%
Blogs and website articles	Microsoft blog	34	20.5%
Newspaper articles	<i>The New York Times</i>	23	13.9%
Social media posts	Tweets	4	2.4%

In addition to our review of the literature, we tested ChatGPT with many different queries. Only a fraction of these random tests is discussed in the next section. Unlike other recent academic articles and editorials (King & ChatGPT, 2023; Kung et al., 2022; O'Connor & ChatGPT, 2023), ChatGPT is not a co-author of our article, and we used the chatbot only very sparingly for brainstorming.

Functionality of ChatGPT

This section demonstrates the steps to navigating the ChatGPT website and its various functions.

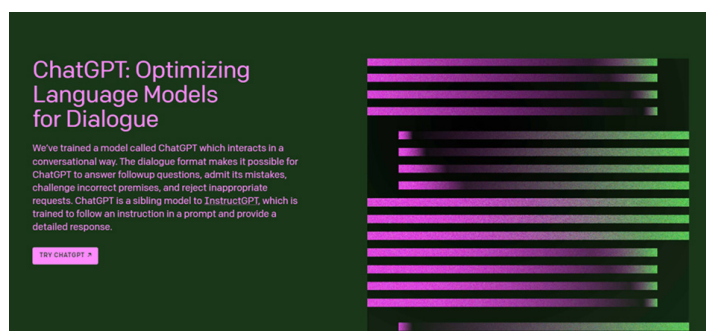


Figure 4. The homepage of ChatGPT (2023).

To access the AI and its functions, one must first create an account. Creating an account only requires an email address and password, and no charges whatsoever are incurred for the time being. Then, individuals must provide their first and last name, country of origin, and cell phone number to complete the registration (see Figure 5).

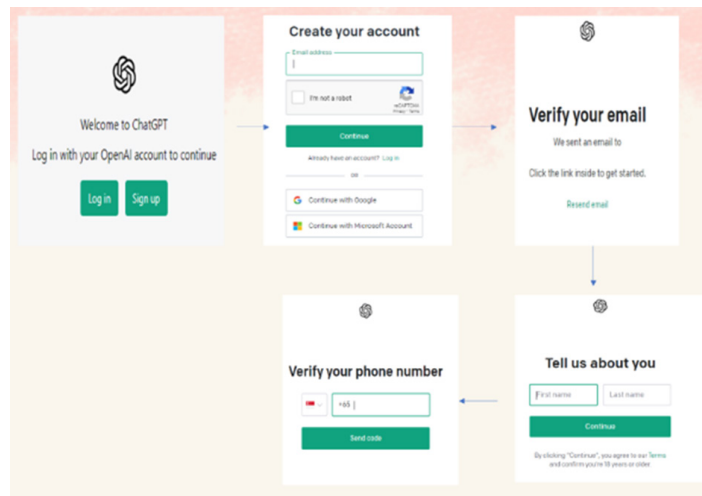


Figure 5. Step-by-step instructions for creating an account for ChatGPT.

Once the account has been created, users are greeted with some general information, as seen in Figure 6.

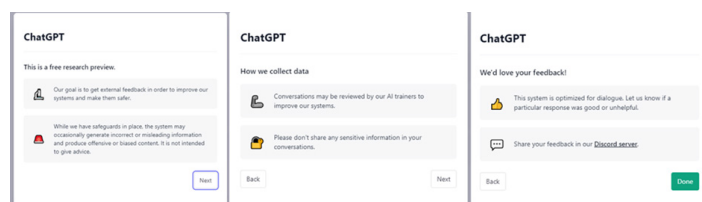


Figure 6. Short introduction of ChatGPT.

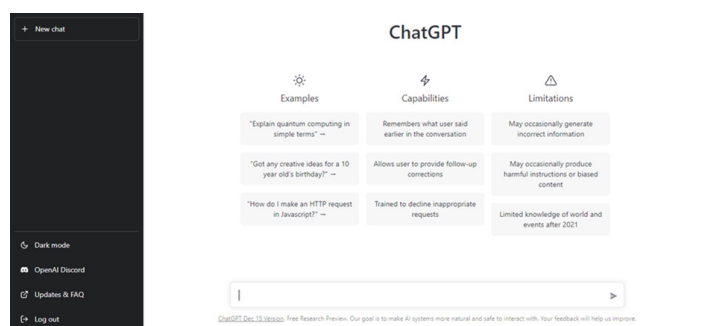


Figure 7. Screenshot of ChatGPT search engine.

When users start using the software, they will see the following categories: examples, capabilities and limitations (see Figure 7). Only the "Examples" tab of these categories is a clickable link. "Capabilities" and "Limitations" are meant only to provide information and guidance to using the software. Some capabilities of the AI are the ability to retain previous conversations, provide follow-up corrections and pick up inappropriate requests. The ChatGPT search page

also states some limitations, such as the generation of invalid information, biased content, and limited knowledge of world events after 2021 (see below).

Under the examples category, there are pre-existing topics for users to experiment with. For example, when individuals click on the option "Explain quantum computing in simple terms", the individual is immediately directed to the following webpage, and the answer to the question is formed within 60 seconds (Figures 8 & 9).

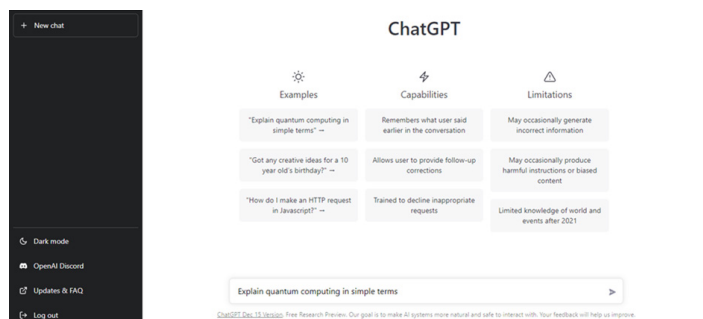


Figure 8. Asking ChatGPT a question, "Explain quantum computing in simple terms".

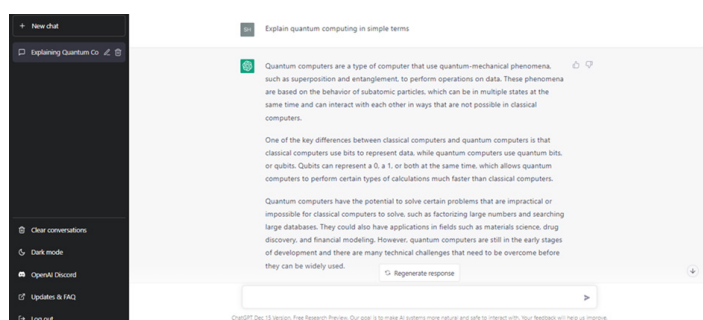


Figure 9. ChatGPT generates an answer within 60 seconds (depending on Internet speed).

In an attempt to investigate the AI's capability, we asked, "Explain mean, standard deviation and z-scores". The AI could offer a clear and sensible explanation of the statistical concepts (see Figure 10) within 60 seconds. It even gave a brief mathematical example for a more precise explanation.

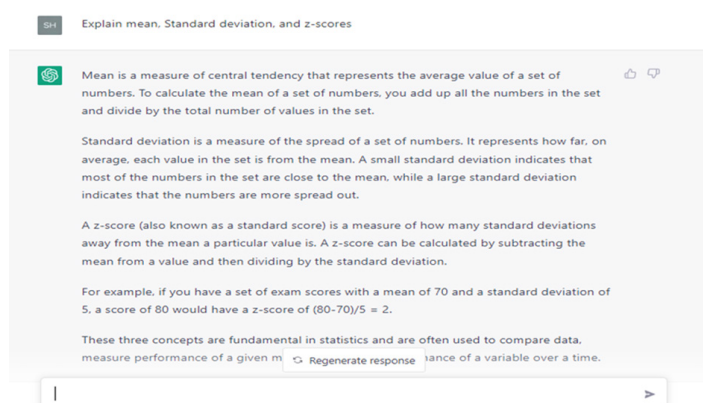


Figure 10. A statistical explanation for mean, standard deviation, and z-scores.

Next, we followed up with another question: "Please provide an example for an application to the real world for standard deviation. Provide step-by-step calculations." The AI readily replied using a case study of a real estate agent trying to determine the price range of houses. It gave a clear breakdown of the mathematical formulae and step-by-step guidelines (Figure 11).

However, while explaining the mathematical formula, the AI stopped working. We immediately asked, "what happened?" The AI explained that its explanation was too long and went past the software's character limit. Thus, the AI scrapped the previous example and presented a new scenario (Figure 11). From the interaction, the AI could communicate and answer queries like a human, and the answers presented were sound and performed within 60-120 seconds. However, there are also limitations, such as limited word characters, the inability to draw diagrams (histogram, scatterplot), and consistent network error prompts when using ChatGPT (Figure 12).

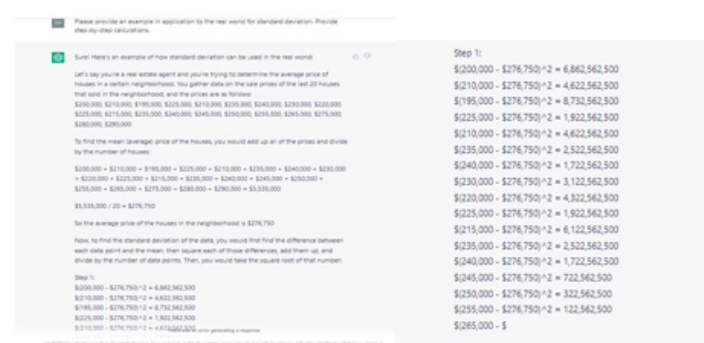


Figure 11. ChatGPT depicting a scenario.

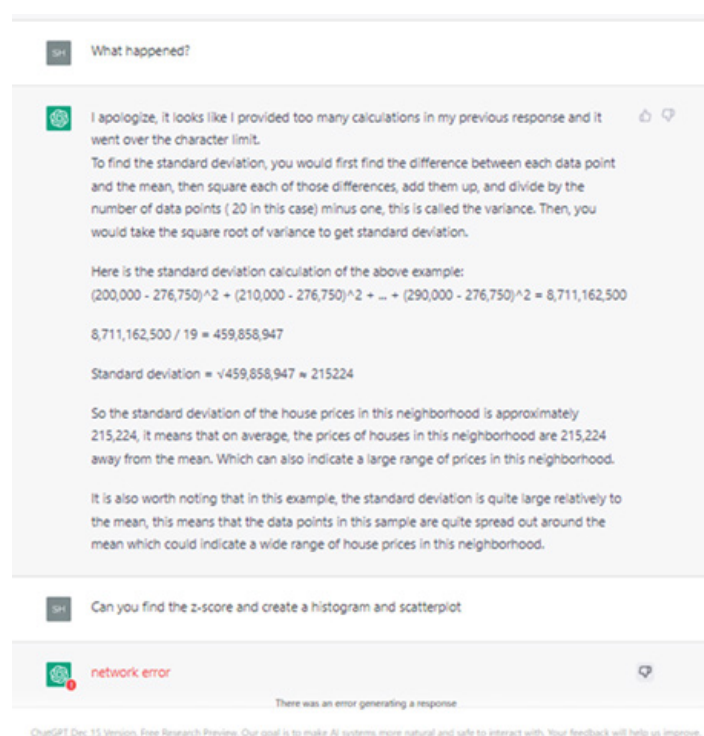


Figure 12. ChatGPT is unable to draw diagrams.

The previous example depicts a more technical example (mathematical) of ChatGPT. In the next example, we asked the AI to draft an essay according to a topic, "Write a 2000-word essay on 'The presence of auditory hallucinations alone is not indicative of schizophrenia'. Please add 20 in-text citations and end-of-text references". In doing so, we investigated whether the AI could write a critical and analytical essay.

After posing the question in ChatGPT, the AI showed an alert of a network error. After multiple attempts, it produced an essay of approximately 500 words and five end-of-text references (Figure 13). Although ChatGPT efficiently produced the essay within 120 seconds, the content was quite disappointing. It lacked both breadth and depth. It was primarily generic and descriptive, with no evidence backing it up. It was also unable to give in-text and end-of-text references (or, worse, invented bogus references; see King & ChatGPT, 2023). Despite having an explanation given about the disorder, its content was only good for leisure reading (something one would read from a random non-academic website after a Google search).

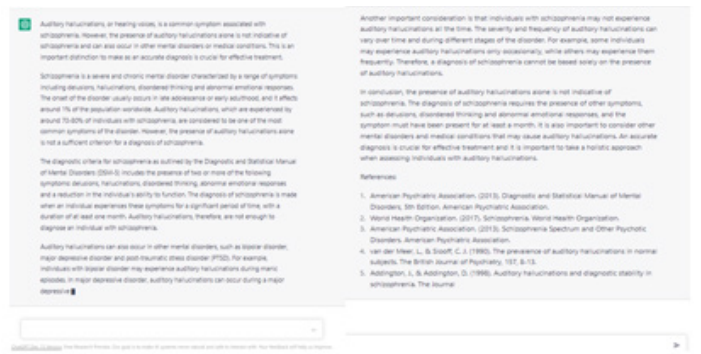


Figure 13. Answers presented for the topic 'The presence of auditory hallucinations alone is not indicative of schizophrenia'.

From the above examples, it is evident that ChatGPT can be beneficial in providing conceptual explanations and applications. However, the AI is less competent with content that requires higher-order thinking (critical, analytical thinking).



Figure 14. ChatGPT generates a Chinese composition.

In this final example, we explored whether ChatGPT can communicate in languages other than English (Figure 14). We tasked the AI to compose a fictional Chinese composition according to the topic, "You once made a mistake that broke the hearts of your parents. Write in detail the process, emotions felt (regrets), and the lesson learnt".

Although the generated answer depicted a scenario accordingly, the essay lacked structure and was grammatically poor. The content was directly translated from English, making some parts illogical (see Appendix B for a rough translation of the Chinese text). Overall, despite the swift composition of the story, it is evident that the Chinese language is not its forte. This is in line with Jiao et al.'s (2023) research findings that ChatGPT performs competitively with translation products such as Google Translate on high-resource European languages (like English and German) but lags behind significantly on low-resource or distant languages such as Chinese and Romanian. Table 2 summarises the current strengths and limitations of ChatGPT.

Table 2. Strengths and limitations of ChatGPT.

Strengths	Limitations
● Language understanding: ChatGPT is trained on a wide range of text, allowing it to understand and respond to various natural language inputs (Susnjak, 2022). ● Human-like conversation: Answers generated sound like they were written by a human, making it well-suited for use in chatbots and other conversational interfaces (Azaria, 2022). ● Flexibility: ChatGPT can be fine-tuned to a specific task or domain, allowing it to generate more accurate and relevant responses (Deng & Lin, 2022). ● Speed: ChatGPT can quickly generate responses, making it suitable for real-time applications. ● Cost-effective: ChatGPT can be used to automate repetitive tasks or provide information, reducing the need for human labour and increasing efficiency (Cotton et al., 2023). ● 24/7 personal assistant potential: ChatGPT has great potential to become a personal assistant for general or professional consultation purposes (Guo et al., 2022).	● Understanding context: ChatGPT may not always fully 'appreciate' the nuances and subtleties of a conversation (Azaria, 2022). ● Language: ChatGPT has been trained on a large dataset of text, but it may not understand or respond correctly to certain words or phrases it has not seen before. Certain languages may also pose difficulties for the AI (Hartmann et al., 2023; Jiao et al., 2023). ● Knowledge limits: ChatGPT can only provide information that it has been trained on up until a specific date (September 2021) and does not have real-time access to new information (Azaria, 2022; Guo et al., 2022). ● Emotion: ChatGPT is not capable of experiencing or recognising emotions; it may not understand or respond appropriately to questions or comments that include emotions. ● Creativity and originality: ChatGPT's responses are based on patterns it has learned from the text. It can only generate responses similar to what it has seen before. It may be unable to create unique and original content or ideas (Susnjak, 2022).
● Logical and well-organised approach: ChatGPT usually defines the core concept and then proceeds to provide detailed answers step by step, before finally offering a summary (Guo et al., 2022). ● Little bias and harmful information: ChatGPT tends to be neutral on sensitive political topics (Guo et al., 2022). ● Awareness of its ignorance: ChatGPT refuses to answer questions it is ignorant about, for instance on queries that require information after September 2021 (Guo et al., 2022).	● Misinformation and 'alternative facts' creation: The response generated may be plausible sounding but make no practical sense, or the information may be inaccurate (Tung, 2022). Hence, ChatGPT makes it a point to encourage users to verify the information (Ortiz, 2022). There is the danger of large amounts of ChatGPT-generated content crowding into user-generated content) platforms, threatening the quality and reliability of the platforms. For instance on legal, medical, and financial questions, ChatGPT may potentially generate harmful or fake information (Guo et al., 2022). ● Quality of responses varies: ChatGPT is unable to generate clarifying questions, and the quality of answers generated would differ, and they may contain inaccurate information (Ortiz, 2022). ● Danger of jailbreaking: users succeeded in tricking ChatGPT how to make Molotov cocktails and generating neo-Nazi arguments (Vincent, 2022c).

A note on our use of 'understanding' and 'appreciating' in the above table is in order. Although AI systems like ChatGPT "do not understand what they read in the same sense or to the same extent that a human does, they can nevertheless extract significant amounts of information from natural language and use that information to make simple inferences and answer questions" (Bostrom, 2017, p. 86). Natural language models are essentially *stochastic parrots* (Bender et al., 2021). We conclude this section with a quote by OpenAI's CEO, Sam Altman:

ChatGPT is incredibly limited but good enough at some things to create a misleading impression of greatness. It's a mistake to be relying on it for anything important but a preview of progress. We have lots of work to do on robustness and truthfulness (cited in Alshater, 2022).

Implications of ChatGPT for education

Artificial Intelligence in Education (AIEd)

In the wake of ChatGPT's release, technologists and educators have been fascinated and alarmed at the same time. There are opponents and proponents of ChatGPT, but it is instructive for those in the education fraternity to examine the educational research in AIEd in order to gain insight and make informed evaluations into the significance of ChatGPT in education. For context, researchers working in the field of Artificial Intelligence in Education (AIEd) have investigated the use of AI for the creation of learning technologies for improving education since the 1970s (du Boulay, 2016). The academic community associated with AIEd has examined, debated, and discussed the benefits of the discipline in the last thirty years, with a view to making the discipline more widely understood.

In the last decade, AI technologies have advanced dramatically, which makes ChatGPT an inevitable development. The convergence of emerging technologies, such as the rise in computing power and big data analytics, has only been occurring in the last few years, contributing to the emergence of sophisticated AI algorithms that can learn and improve on their own (Tan, 2020). With these changes, AIEd emerged as a technology capable of transforming our social interactions in radically new ways. It is already clear that artificial intelligence has the potential to revolutionise the way we learn and teach and that these methods are currently being tested in a variety of educational settings, even before ChatGPT made its debut.

ChatGPT represents one of the latest breakthroughs in AI, and as such, it is worthwhile to review the current research on AI-powered applications in education and contextualise ChatGPT based on the trending frameworks for discussing the impact of ChatGPT on education. The following section seeks to provide a more comprehensive understanding of the position of ChatGPT in relation to current AIEd research.

ChatGPT in the context of existing artificial intelligence in education

In the rapidly expanding field of education technology, AIEd represents an opportunity to demonstrate a broad spectrum of tools and applications at an entirely new level. This presents excitement and a number of breakthroughs in establishing a broad range of tools and applications. A review of the literature in AIEd indicates how educators can minimise their risks while applying AI in experimenting with innovative practices in teaching and learning. Baker and Smith (2019) categorise educational contexts as student-facing, teacher-facing, and system-facing, which all have the potential to profoundly transform educational practices. It has been found that this framework has provided significant clarifications regarding the use of artificial intelligence in education.

Student-facing AI applications

The use of student-facing AI applications offers exceptional potential for improving intelligent student support systems and scaffolding student learning in adaptive and personalised ways (Zawacki-Richter et al., 2019). It can be argued that intelligent tutoring systems (ITS) are one of the most promising benefits of artificial intelligence when it comes to transforming education, as they are one of the most effective tools for personalising instruction. The growth of this personalisation is currently taking place as education researchers experiment with new models of learning, and, as a result, new opportunities are arising in the field. With the application of AI-powered algorithms, it is now possible for ITSs to simulate the assistance provided by a tutor, such as by providing personalised assistance in solving problems. It is anticipated that, as big data technology advances in the field of learning analytics, a revolutionary paradigm of adaptive, personalised learning will emerge. These technologies will be capable of recording and interpreting the characteristics of students and their emotional state in every aspect of their learning in real-time, resulting in personalised adaptive learning (PAL) (Peng et al., 2019).

Similarly, ChatGPT is trained on a large dataset of text data, enabling it to learn patterns and relationships in the language and to generate new text that is similar to the text it has been trained on. However, the model uses a neural network architecture called a Transformer, which is particularly well-suited to processing and generating text. In the absence of a thorough investigation, it appears that the technology behind ChatGPT could potentially be utilised to improve the performance of personalised adaptive learning. As at the current stage of development, the ChatGPT model seems to be limited to fine-tuning specific tasks, like language translation or answering questions, to improve its performance. It is imperative to note that both the impressive capabilities of ChatGPT as well as its limitations reflect the fact that it operates in a similar manner to Google's smart compose suggestions, generating ideas based on what it has previously read and processed (Heilwell, 2022). Consequently, it can appear confident while not demonstrating a deep understanding of the subject matter.

While ChatGPT has shown impressive capabilities in helping students in writing, AI-powered writing applications have already been widely used for quite some time. It is ironic that ChatGPT has caused so many anxieties in the academic community, and yet it may prove most useful to teachers when it comes to facilitating more innovative teaching and learning. The next section discusses teacher-facing AI applications.

Teacher-facing AI applications

Teachers may use teacher-facing AIED in order to reduce their workloads, gain insights from their students, and facilitate classroom innovation (Baker & Smith, 2019). These AIED systems are designed to assist teachers by automating assessment, plagiarism detection, administration, as well as feedback mechanisms. The AI-powered applications may also enable teachers to gather insight into their students' learning progress in order to provide additional guidance and support as needed. Cope et al. (2020) suggest that AI-powered assessment applications hold the greatest potential for bringing about transformative education changes. Contrary to conventional assessment methods that rely on distinct and atypical artefacts to select and provide response tests for retrospective, summative sampling, AI-powered assessment systems may support the integration of continuous feedback into learning processes by utilising distinctive and atypical artefacts. The following section provides an overview of AI-powered assessment applications applicable across disciplines and those that are specifically designed to support writing.

Automated Essay Scoring (AES) systems are the most common AI-powered assessments and can be applied across various disciplines, but most of the research has focused on its application to undergraduate courses (Zawacki-Richter et al., 2019). There are a variety of methods of developing AES systems, such as statistical modelling, natural language processing (NLP) and Latent Semantic Analysis (LSA), and the algorithms can be used to identify patterns in text responses and prompt students to revise their responses (Ma & Slater, 2015). This, in turn, could allow educators to consider a broader range of assessment methods than only using multiple-choice tests to assess students' knowledge and abilities. In general, it appears that AI-powered essay ratings are comparable to human ratings, notwithstanding some areas of concern (Aluthman, 2016).

For AES to be effective, it needs to be combined with AI-enabled automatic feedback. Using machine learning systems to provide automatic feedback to students and improve their writing skills is another prevalent application (Garcia-Gorrostieta et al., 2018). The automated feedback system operates on adaptive evaluation to establish the appropriate answers based on Bloom's cognitive levels and recommend additional learning resources and challenges (Barker, 2011).

It is evident that AI-powered applications for grading essays have a growing body of research that indicates their efficacy. However, there is another growing subfield in this area of research in which AI-powered applications are used

to support students in the acquisition of writing skills. It is in this area that educators and pundits are concerned that ChatGPT will disrupt and inevitably bring about the end of writing as we know it. The following section provides an overview of such AI-powered writing assistants that writing instructors have been using and researching over the past decade, as well as the opportunity to examine how to situate ChatGPT as part of this ecosystem of AIED in the future.

It is evident that, prior to the introduction of ChatGPT, a number of AI-based writing tools had already been developed to facilitate English writing practices and to enhance writing skills, as well as promote self-directed learning by users, particularly in higher education (Nazari et al., 2021; Zhao, 2022). In general, automated writing evaluation (AWE), automated essay scoring (AES), and automated written corrective feedback (AWCF) have been increasingly adopted as alternatives to facilitate the process of writing by facilitating automated feedback and assessing items. The new AI-powered writing applications may serve as a flexible and time-saving addition to the writing curriculum since they integrate the AWE, AES, and AWCF features into one integrated application (Koltovskaia, 2020).

In terms of AI-powered digital writing assistants, Grammarly is one of the most popular and well-researched ones, offering a wide range of applications, including AWE, AES, and AWCF, all in one digital tool for writers (Taguma et al., 2018). As Grammarly has more than 20 million users worldwide, the corpus amount of data generated by its users provide the base for it to continually improve the application. A team of computer linguists and deep learning engineers at Grammarly analyse millions of sentences from academic journals in order to build cutting-edge algorithms that analyse the rules and latent habits of effective writing (Fitria, 2021).

Grammarly detects spelling and grammar errors in English texts and corrects them to the appropriate form. The system prompts users to correct errors immediately by pointing out where they are located. Figure 15 is an example of how Grammarly checks a text passage.

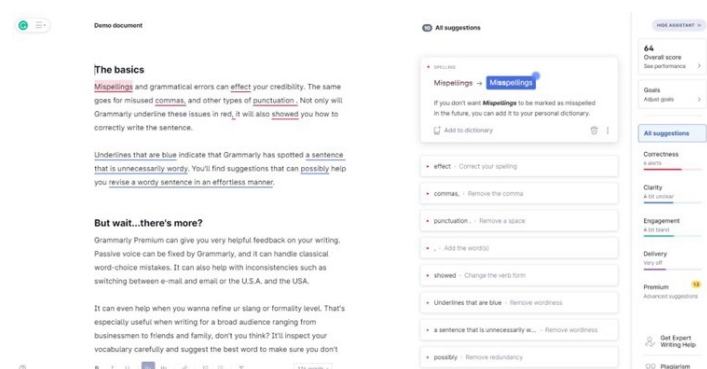


Figure 15. An example of text correction after using Grammarly.

Research indicates that the utilisation of Grammarly is an effective intervention for improving writing engagement with automated written corrective feedback (Koltovskaia, 2020). In several studies, multiple indicators of student

engagement, such as grit, were found to be positively influenced by technology (Schindler et al., 2017). By providing immediate feedback and revision, Grammarly may motivate students to revise by providing technology scores (Moore & MacArthur, 2016). When Grammarly scans to rectify erroneous writing, it indicates where the error is and provides a “technology score”. Figure 15 illustrates an example of a technology score of 64. An increase in the score corresponds to a reduction in errors, thereby encouraging users to keep improving the writing task. In another study, the findings indicate that AI was an effective intervention for enhancing self-efficacy and academic emotions in English as a Foreign Language (EFL) students. In the absence of human assistance, intelligent feedback can reinforce students' writing autonomy by helping them to recognise their writing errors, identify incorrect patterns, and reformulate them (Nazari et al., 2021).

Similarly, Wordtune is another well-researched AI-powered writing assistant that supports EFL students in writing. It offers options for rewriting the highlighted text by altering the sentence structure or replacing words with synonyms while maintaining their original meaning. Wordtune uses Natural Language Processing (NLP) to train the machine to understand and generate natural text based on large datasets of written material, utilising patterns learned from large datasets to provide options for rewriting one's own sentences instead of taking content from other online sources (see Figure 16).

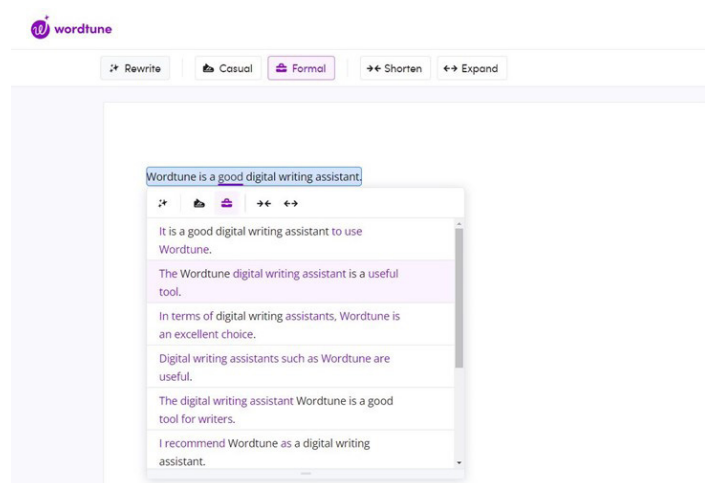


Figure 16: An example of rewrite suggestions via Wordtune (formal tone).

Zhao (2022) argues that Wordtune supports EFL writers in formulating ideas and translating them into English, thus, improving the quality of their writing. In addition, users can also be motivated throughout the writing process and avoid getting stuck on difficult English words or expressions. AI-powered writing tools can provide reliable and accurate information about learning through formative and summative assessments (Nazari et al., 2021).

The review of AI-powered writing assistants has yielded considerable evidence that the prevalent AI-powered writing assistants or text-generative tools have been effective in supporting users' writing by either checking their grammar and errors or offering suggestions for revising. As a result of

the review of the efficacy of AI-powered writing assistants, ChatGPT can be positioned within the same AIEd category for the purpose of further analysis.

While ChatGPT is created using similar AI technology behind the other text-generative tools, it has the unique capability of being able to generate text that sounds remarkably intelligent based on the prompts provided by users, including homework assignments and exam questions, making it appear to have been written by a human. Considering the quality and depth of the research, as well as the reasonably good citations of the responses produced by ChatGPT, some academics anticipate the demise of conventional educational assessment methods (e.g. Yeadon, 2022; Herman, 2022). The next subsection discusses how ChatGPT could fit in the system-facing AIEd.

System-facing AI applications

System-facing AI-powered applications provide academic administrators and managers with macro-level information, such as attrition patterns across schools or institutions. However, this area has received the least attention (Baker & Smith, 2019). It is evident from the literature review that the data required for training the algorithms and learning models for all three AIEd dimensions has a great deal in common, even though system-facing applications have not been as well-researched as student-facing and teacher-facing applications.

While ChatGPT may appear to be more relevant to AI-powered applications for teacher-facing AIEd in the form of AI-powered writing assistants, educationists should take a holistic approach when developing strategies for leveraging ChatGPT for innovation in education. It is instructive to consider how Microsoft appears to be incorporating ChatGPT holistically into its products (Warren, 2023).

Having reviewed and established ChatGPT in the context of educational technology, specifically AIEd, it is appropriate to examine what impact it will have on the education of the future. Since ChatGPT is a brand-new product in the market, there is a dearth of empirical research to determine its implications on education. A discussion of the opportunities and challenges that ChatGPT may have on educational practitioners, policymakers, and researchers is now in order.

The impact of ChatGPT on education

While there are many opinion pieces written about ChatGPT, a review of the literature on the implication of ChatGPT in education only yielded eight preprints and two academic articles with a variety of responses ranging from positive to negative (see above). The following section discusses the challenges and opportunities of ChatGPT in education and their implications for educational stakeholders.

Use of ChatGPT for education: Challenges

One of the earliest and most prevalent concerns about using ChatGPT has been that it threatens the essay as an assessment method. For a start, some instructors are worried that students will outsource their written assignments to ChatGPT as it can generate passable prose in seconds without triggering any plagiarism detector. Such concerns, however, may arise from the resistance of instructors to adapt to the change in assessment methods since written assignments are often criticised for being dull and ineffective in assessing students' learning (McMurtrie, 2023).

A second concern that instructors have is ChatGPT's inability to understand what is being shared and to take the time to evaluate the relevance or accuracy of the information, as it is just a text-generating machine that simply produces a passable imitation of what is being shared (Warner, 2022). The concern may be seen as a legitimate reason for institutions to implement policies blacklisting the AI application, but it might not be long before Microsoft integrates ChatGPT's technology into its suite of Microsoft products (Metz & Weise, 2023). By that time, ChatGPT will be mainstream, and it may be too late for educational institutions to adjust policies to guide their students in using it appropriately.

A pragmatic approach and a focus on managing the challenges presented by ChatGPT may be a better approach to take. Brown et al. (2020, p. 9) wrote:

Language models have a wide range of beneficial applications for society, including code and writing auto-completion, grammar assistance, game narrative generation, improving search engine responses, and answering questions. But they also have potentially harmful applications. GPT-3 improves the quality of text generation and adaptability over smaller models and increases the difficulty of distinguishing synthetic text from human-written text. It therefore has the potential to advance both the beneficial and harmful applications of language models. Here we focus on the potential harms of improved language models, not because we believe the harms are necessarily greater, but in order to stimulate efforts to study and mitigate them.

In general, when disruptive education technologies enter the classroom, the practice of teaching and learning is often subject to a number of challenges. Education practitioners and policymakers are always responsible for managing the situation. When these challenges are not addressed, inadequate pedagogical practices may be exposed. There has been a sensational report on social media about a Chinese schoolgirl who bought a machine to copy large amounts of Chinese text for her homework (*Today Online*, 2019; see Figure 17). It was not only capable of reproducing Chinese texts, but it was also intelligently designed to mimic the handwriting of the schoolgirl. The schoolgirl managed to get away with doing her homework until she was caught by her mother, who shamed her on social media. Based on the example given, one may conclude that if a machine is capable of outwitting a teacher's pedagogy, it may be able

to replace the teacher. In this regard, it is imperative for teachers to transform challenges into opportunities and adapt to changes as they arise.



Figure 17: A Chinese-language automated handwriting machine (YP, 2019).

Use of ChatGPT for education: Opportunities

ChatGPT's capability to generate essays has created challenges for educators, but there are those who seem ready to embrace the opportunities for innovation in teaching and learning that this disruptive AI application presents. McMurtrie (2022) argues that tools like ChatGPT will become part of everyday writing in some shape or form, just as calculators and computers have become part of math and science. Similarly, Sharples (2022) suggests engaging students and instructors in shaping and harnessing these AI tools to support learning rather than stopping the students from using it.

While essays as assessments are regarded as threatened by ChatGPT, therein lies an opportunity for educators to introduce innovative assessments. Most of the time, assessments are perceived and utilised by instructors for the assessment of students' learning. The majority of instructors, however, may not possess the skills to use assessment both *for* learning (Wiliam, 2011) and *as* learning (Earl, 2012). In this regard, institutions can take advantage of this opportunity to enhance instructor skill sets in assessment to harness disruptive AI applications such as ChatGPT to improve students' learning. Our recommendations section (see below) provides further details on how instructors can innovate assessments.

Another interesting opportunity for instructors is to leverage ChatGPT to innovate their teaching strategies. Instructors could use flipped learning to ensure that the most critical pieces of work are completed in class and to focus more on multimedia assignments or oral presentations as opposed to class assignments. Additionally, instructors have the opportunity to spend more time giving feedback and revising students' work.

A major benefit of ChatGPT is that it allows students to learn through experimentation and experience. Using ChatGPT, students can evaluate different strategies and approaches to solving problems and achieving goals through game-based

learning (Sutton & Allen, 2019) or other student-centred pedagogies (Mills, 2023a). Students who prefer hands-on, experiential learning will gain from using ChatGPT as a learning aid.

With the aid of appropriate instructional strategies, ChatGPT can be utilised to facilitate collaboration and teamwork between participants. There are a variety of student-centred learning strategies that can be designed to be played in groups. The ChatGPT application has the potential to serve as a means of generating different scenarios for students to work together to solve problems and achieve goals. In this way, a sense of community can be fostered, and students can learn from one another and support one another.

While ChatGPT is perceived as a disruptive technology in the teaching and learning process, it represents a huge opportunity for learning innovators to use it to transform education. In the following section, we provide our preliminary conclusions and recommendations for leveraging ChatGPT to advance education innovation.

Conclusions and recommendations for higher education teachers and institutions

With tools like GPT-3 and ChatGPT, AI appears to be in the process of going mainstream (Vincent, 2022b). We are only beginning to see the effect this will have on the world in general and higher education in particular. If Altman is right that “we could get to real AGI in the next decade” (cited in Ortiz, 2022), this would have huge societal implications. ChatGPT could be the “beginning of the end of all white-collar knowledge work” and “a precursor to mass unemployment” (Roose, 2022; see Krugman, 2022; Chesterman, 2023). Whilst the alarmist and sensationalist reporting in news media is, in our view, not justified, it will be important to watch and engage in this fast-developing space and adjust learning, teaching, and assessment approaches in higher education. We did some random testing with anti-plagiarism software, and it was unable to detect ChatGPT’s work (see above and Appendix A). Plagiarism checkers such as the one embedded in the professional version of Grammarly are unlikely to flag text generated by ChatGPT and similar programs, as it is, after all, original text (Dehouche, 2021; Mindzak & Eaton, 2021; Anson & Straume, 2022; Stokel-Walker, 2022). However, it was recently reported that Turnitin, a leading anti-plagiarism software, “is in the midst of enhancing its software’s ability to recognise ChatGPT writing and incorporate it into its products for educators to use in 2023” (Chia, 2023). Apparently, ChatGPT can be used to check sentences for plagiarism that are input by the user and then modify them so that anti-plagiarism software reports a low originality index score:

I want you to act as a plagiarism checker. I will write you sentences and you will only reply undetected in plagiarism checks (sic!) in the language of the given sentence, and nothing else. Do not write explanations on (sic!) replies. My first sentence is “For computers to behave like humans, speech recognition systems must be able to process nonverbal information, such as the emotional state of the speaker (Akin, 2022).

When we tried this, ChatGPT responded as follows: “To emulate human behavior, speech recognition must have the ability to interpret nonverbal cues, including the speaker’s emotional state”. This appears to mean that ChatGPT can be used to reduce a high originality index score in a student assignment!

Future developments notwithstanding, it is ironic that anti-plagiarism software uses artificial intelligence to assess the originality of assignments and that different AI (like ChatGPT) can be used to get around plagiarism detection software within seconds. The irony is complete when we realise that GPT-3 can write a review of the student’s AI-generated assignment on behalf of the teacher via a simple command: “Here is a short assessment of this student essay:” (Sharples, 2022b). A first AI circumvents a second AI and is assessed by a third AI. All that the humans do is press a couple of keys, and nobody learns anything.

Recommendations

Higher education reactions to ChatGPT and GPT-3 have been on a continuum between the extremes of banning or prohibiting the use of the software and including it in the curricula. How should students, teachers and higher education institutions deal with ChatGPT? Marche (2022) predicted that it may take “10 years for academia to face this new reality: two years for the students to figure out the tech, three more years for the professors to recognize that students are using the tech, and then five years for university administrators to decide what, if anything, to do about it”. Although this epitomises the bureaucratic inertia of many university environments, that would not be good enough by a long shot.

Generally, we advise against a policing approach (that focuses on discovering academic misconduct, such as detecting the use of ChatGPT and other AI tools). We favour an approach that builds trusting relationships with our students in a student-centric pedagogy and assessments *for* and *as* learning rather than solely assessments *of* learning (Wiliam, 2011; Earl, 2012). The principle of constructive alignment asks us to ensure that learning objectives, learning and teaching and assessments are all constructively aligned (Biggs & Tang, 2011). Here are our own preliminary thoughts on some recommendations for faculty, students, and higher education institutions.

Recommendations for higher education faculty

A simple solution to the problem of students using ChatGPT would be to use physical closed-book exams where the students write by hand, using only pen and paper (Cassidy, 2023) – for online exams, proctoring/surveillance software can be used. However, such an approach to assessment (or at least an over-reliance on it) has been increasingly criticised as no longer contemporary, with students cramming less-than-useful information into their heads, only to forget much of it shortly after their examinations (Van Bergen & Lane, 2016). With a focus on graduate employability, the skill to ace closed-book exams seems rather irrelevant.

Another idea to combat the use of text generators such as ChatGPT and GPT-3 is to design writing assignments that they are currently not good at handling. This approach may be a very short-term solution, given how quickly the technology is developing. Mills (2023b) has compiled the following suggestions: analysis of images and videos; analysis that draws on class discussion; analysis of longer texts that do not fit in a prompt; and writing about recent events that are not in the training data for the text generator. In addition, we can require students to write about a topic that is highly specific and niche in the hope that it will be difficult for AI systems to find relevant information. Also, we can ask students “to include personal experiences or perspectives in their writing, which are difficult for AI systems to replicate” (Nowik, 2022). Importantly, we can assess students on their ability to integrate multiple sources and present their own original arguments (Nowik, 2022)

At present, a very important limitation of ChatGPT is that it does not provide sources and quotations. Whilst it is able to provide book recommendations and provide reasons for its recommendations, it does not provide in-text referencing and a reference list in its responses. This is a major shortcoming in writing academic assignments (that usually require a certain number of references). However, OpenAI has already created a WebGPT prototype, which has access to web browsing (OpenAI, 2021). WebGPT would thus not only be able to incorporate recent information but also verified sources and quotations. In the meantime, Elicit (<https://elicit.org/>), which markets itself as an AI research assistant capable of reducing the time needed to write a literature review and a research proposal, is a GPT-3-based tool that is able to respond to research questions and suggest academic articles and provide summaries of them from a repository of 175 million scholarly papers (Andrews, 2023; Tate, 2023).

There is text generator detection software (e.g., <https://writer.com/ai-content-detector/> or <https://huggingface.co/openai-detector/> or <https://gptzero.me/>) that estimates the probability that text is written by a large language model (Tate et al., 2023; Sandlin, 2022; Mills, 2023a; McMurtrie, 2023; Montclair State University, 2023; Yousif, 2023). Due to the difficulty of distinguishing human and ChatGPT-generated text, OpenAI is looking into ‘watermarking’ text, an undertaking fraught with difficulties for which there will be workarounds (Wiggers, 2022). All this is bound to lead to a race between text generators and text generator detection tools. If history is any judge, students will find ways around detection tools, and many instructors may not want to become writing police (McMurtrie, 2023). In a matter of doubt, teachers could test students’ knowledge of their own essays and conduct an impromptu oral exam (Allen, 2022), thus further adding to higher education teachers’ workload. Alternatively, video or audio submissions of students discussing their essays or metacognitive reflection on their writing process (even though they are not AI-proof) could be required (Mills, 2023a).

In terms of assessments, we recommend avoiding assignments and examinations that are so formulaic that nobody could tell if a computer completed them (Herman, 2022). We should rather create assessments that foster

students’ creative and critical thinking abilities (hooks, 2010; Brookfield, 2012; Brookfield et al., 2019; Rudolph & Tan, 2022); for instance:

- execute certain assessments during class;
- create assessments where students deliver presentations, performances and other digital forms, including webpages, videos and animations (Lim, 2022);
- allow students to write about topics that genuinely interest them, in which their voices come through, and their opinions are valued (McMurtrie, 2022); and
- use authentic assessments (i.e. creative learning experiences that test students’ skills and knowledge in realistic situations: Wiggins, 1990) that are meaningful and intrinsically motivating.

Students can also be involved in peer evaluations and ‘teach-back’ (Sharples, 2022b) – teach-back is a communication confirmation method that is particularly popular in healthcare (patients/students need to demonstrate their understanding in speech).

Ideally, higher education teachers would create an atmosphere where students are invested in their learning (McMurtrie, 2022). In this context, Tan (2022) has called for humanising the academy by integrating empathy, kindness, and compassion into learning and teaching. It would be explained to students that writing is a form of thinking (Stevens, 2019) and that they miss out on a critically important form of learning if they try to delegate their writing to ChatGPT or another AI.

At the same time, AI tools can be incorporated into discussions and assignments (see Anson & Straume, 2022; McMurtrie, 2022, 2023; Fyfe, 2022; D’Agostino, 2022). We recommend educating our students on the limits and faults of text-generator prose and sharing substandard text examples that highlight the value of human (including students’) writing (Mills, 2023a). We need to help students learn how to use AI tools judiciously and understand their benefits and limitations. They may “help spark the creative process” (McMurtrie, 2023). Faculty can make use of these tools as a means to help students with writing and research, but not as a replacement for critical thinking and original work. Lim (2022) wrote that ChatGPT constituted an “extraordinary technological marvel” that “presents an opportunity for us to move beyond rote learning to nurture our students to become more creative, thinking individuals as we reflect on what it means to learn and be human in the digital age.”

Recommendations for students

Our students are often digital natives who use technology more effortlessly and intuitively than their teachers. Our recommendations for students are to work on both their strengths and weaknesses further:

- be aware of academic integrity policies and understand the consequences of academic misconduct;
- be digitally literate, master AI tools (Zhai, 2022) and increase employability as a result;
- write assignments and use AI as a set of tools as a way to improve writing skills and generate new ideas, rather than simply copying and pasting text;
- use high-quality sources and be wary of substandard sources, misinformation and disinformation (Kefalaki & Karanickolas, 2020);
- read widely and voraciously to improve critical and creative thinking;
- learn how to use AI language tools such as ChatGPT to write and debug code (Zhai, 2022); and
- practise the use of AI language tools (like ChatGPT) to solve real-world problems (Zhai, 2022).
- conduct training for faculty on AI tools such as ChatGPT;
- provide training on academic integrity for students;
- avoid offering curricula and courses that do not make sense to students (as, consequently, they might cheat because the value of the work of their learning is unclear to them);
- update academic integrity policies and/or honour codes that include the use of AI tools;
- specifically, develop policies and clear, easy-to-understand guidelines for the use of language models in learning and teaching – the guidelines should include information on the proper use of these tools and the consequences for cheating; and
- encourage, support and share research on AI tools' effects on learning and teaching.

Recommendations for higher education institutions

Peter Fleming has posited that neoliberal (privatised, corporatised, marketised and financialised) universities are in mortal danger largely due to "bad management and hostile government budgets", with the global Covid-19 pandemic an added conundrum (Fleming, 2021, p. 157; see Fleming et al., 2021). Universities that made themselves overly dependent on the lucrative international student market found themselves in a world of trouble when the coronavirus and concomitant travel restrictions emerged in 2020 (Rudolph, 2021; Parker et al., 2021). In such a challenging environment, higher education institutions operate under significant constraints, and major Australian universities have quickly "added new rules which state that the use of AI is cheating, with some students already caught using the software" (Cassidy, 2023). However, other Australian universities recently decided to allow the use of AI in assignments, as long as it is disclosed (Shepherd, 2023).

Our recommendations are as follows:

- realise that digital literacy education is of critical importance and has to include AI tools, which should be part of the curriculum – other useful AI tools include, for instance, Grammarly (a writing and grammar-checking tool that uses AI to check texts for grammar, spelling, punctuation, and other writing-related issues, and offers suggestions for how to improve the writing) and the aforementioned Elicit (Tate, 2023; Krügel et al., 2023; Shepherd, 2023);
- avoid the creation of an environment where faculty is too overworked to engage and motivate their students;

In our title, we asked whether ChatGPT was a bullshit spewer or the end of traditional assessments in higher education. We have seen that ChatGPT occasionally does hallucinate and spout nonsense, for instance, by inventing references. At the same time, we believe that major changes to traditional higher education assessments such as essays and online exams are in order to address the existence of increasingly powerful AI, unless universities want to be akin to driving schools that teach to ride horses.

When hand-held calculators appeared, there was much concern about the end of people's numeracy skills. Today, they are, of course, integral to teaching maths and on every smartphone. Students and academics alike use spell and grammar checkers, thesauruses and Wikipedia. The day when tools such as ChatGPT will be incorporated into MS Word and Google Docs – the software that was used in writing this article – may come sooner (Warren, 2023) rather than later.

Acknowledgements

We thank Begüm Burak, Margarita Kefalaki, Anna Mihaylov, Vanessa Stafford, Michael Sutton, and Eunice Tan for their valuable comments on an earlier draft. All remaining errors are the responsibility of ChatGPT.

References

- Akin, F. K. (2022, December 4) Awesome ChatGPT prompts. *Github*, https://github.com/f/awesome-chatgpt-prompts?utm_source=substack&utm_medium=email
- Akinola, Y. M., Agbonifo, O. C., & Sarumi, O. A. (2020). Virtual Reality as a tool for learning: The past, present and the prospect. *Journal of Applied Learning and Teaching*, 3(2), 51-

58. <https://doi.org/10.37074/jalt.2020.3.2.10>

Allen, M. (2022, December 26). Professor warns about chatbot cheating: "Expect a flood". *Axios*, <https://www.axios.com/2022/12/26/chatbot-cheating-university-warning-chatgpt>

Alshater, M. (2022). *Exploring the role of artificial intelligence in enhancing academic performance: A case study of ChatGPT*. Preprint. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4312358

Altman, S. [@sama]. (2020, July 20). *Tweets* [Twitter profile]. <https://twitter.com/sama/status/1284922296348454913>

Aluthman, E. S. (2016). The effect of using automated essay evaluation on ESL undergraduate students' writing skill. *International Journal of English Linguistics*, 6(5), 54-67.

Andrews, O. S. (2023, January 12). Initial thoughts on elicit AI: Observations of a part-time researcher. *Age of Awareness*, <https://medium.com/age-of-awareness/initial-thoughts-on-elic-it-ai-observations-of-a-part-time-researcher-f1e346364a5e>

Anson, C. M. (2022). AI-based text generation and the social construction of "fraudulent authorship": A revisitation. *Composition Studies*, 50, 37-46.

Anson, C. M., & Straume, I. (2022). Amazement and trepidation: Implications of AI-based natural language production for the teaching of writing. *Journal of Academic Writing*, 12(1), 1-9.

Aydın, Ö., & Karaarslan, E. (2022). OpenAI ChatGPT generated literature review: Digital twin in healthcare. In Ö. Aydın (Ed.), *Emerging computer technologies 2* (pp. 22-31). İzmir Akademi Dernegi.

Azaria, A. (2022). *ChatGPT usage and limitations*. Preprint. DOI: 10.13140/RG.2.2.26616.11526

Baker, T., & Smith, L. (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

Barker, T. (2011). An automated individual feedback and marking system: An empirical study. *Electronic Journal of E-Learning*, 9(1), 1-14. <https://www.learntechlib.org/p/52053/>.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM conference on fairness, accountability, and transparency* (pp. 610-623).

Biddle, S. (2022, December 9). The internet's new favorite AI proposes torturing Iranians and surveilling mosques. *The Intercept*, <https://theintercept.com/2022/12/08/openai-chatgpt-ai-bias-ethics/>

Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university* (4th ed.). McGrawHill.

Bommarito II, M. & Katz, D. M. (2022). *GPT takes the bar exam*. arXiv:2212.14402.

Bostrom, N. (2017). *Superintelligence. Paths, dangers, strategies. Reprint with corrections*. Oxford University Press.

Brookfield, S. D. (2012). *Teaching for critical thinking. Tools and techniques to help students question their assumptions*. Jossey Bass.

Brookfield, S. D., Rudolph, J., & Yeo, E. (2019). The power of critical thinking in learning and teaching. An interview with Professor Stephen D. Brookfield. *Journal of Applied Learning and Teaching*, 2(2), 76-90. <https://doi.org/10.37074/jalt.2019.2.2.11>

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877-1901.

Cassidy, C. (2023, January 10). Australian universities to return to 'pen and paper' exams after students caught using AI to write essays. *The Guardian*, <https://www.theguardian.com/australia-news/2023/jan/10/universities-to-return-to-pen-and-paper-exams-after-students-caught-using-ai-to-write-essays>

Cellan-Jones, R. (2020, July 24). Tech Tent: Have we seen our AI future? *BBC*, <https://www.bbc.com/news/technology-53530454>

ChatGPT. (2023). *Home*. <https://openai.com/blog/chatgpt/>

Chesterman, S. (2023, January 12). AI-generated content is taking over the world. But who owns it? *The Straits Times*, <https://www.straitstimes.com/opinion/ai-generated-content-is-taking-over-the-world-but-who-owns-it>

Chia, O. (2023, January 6). Teachers v ChatGPT: Schools face new challenge in fight against plagiarism. *The Straits Times*, <https://www.straitstimes.com/tech/teachers-v-chatgpt-schools-face-new-challenge-in-fight-against-plagiarism>

Cooper, K. (2021, November 1). OpenAI GPT-3: Everything you need to know. *Springboard*, <https://www.springboard.com/blog/data-science/machine-learning-gpt-3-open-ai/>

Cope, B., Kalantzis, M., & Searsmith, D. (2021). Artificial intelligence for education: Knowledge and its assessment in AI-enabled learning ecologies. *Educational Philosophy and Theory*, 53(12), 1229-1245.

Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2023, January 10). *Chatting and cheating. Ensuring academic integrity in the era of ChatGPT*. Preprint. <https://doi.org/10.35542/osf.io/mrz8h>

D'Agostino, S. (2022, October 26). Machines can craft essays. How should writing be taught now? *Inside Higher Ed*, <https://>

www.insidehighered.com/news/2022/10/26/machines-can-craft-essays-how-should-writing-be-taught-now

Dehouche, N. (2021). Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). *Ethics in Science and Environmental Politics*, 21, 17-23.

Deng, J., & Lin, Y. (2022). The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems*, 2(2), 81-83.

Dowling, M., & Lucey, B. (2023). *ChatGPT for (finance) research: The Bananarama conjecture*. Preprint. https://www.researchgate.net/profile/Michael-Dowling-7/publication/367046908_ChatGPT_for_Finance_Research_The_Bananarama_Conjecture/links/63bf37997ecd35045c3d76a1/ChatGPT-for-Finance-Research-The-Bananarama-Conjecture.pdf

du Boulay, B. (2016). Artificial intelligence as an effective classroom assistant. *IEEE Intelligent Systems*, 31(6), 76-81.

Earl, L. M. (2012). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.

Elkins, K., & Chun, J. (2020). Can GPT-3 pass a writer's Turing test?. *Journal of Cultural Analytics*, 5(2), 17212.

Ferster, B. (2014). *Teaching machines: Learning from the intersection of education and technology*. Johns Hopkins University Press.

Fitria, T. N. (2021). Grammarly as AI-powered English writing assistant: Students' alternative for writing English. *Metathesis: Journal of English Language, Literature, and Teaching*, 5(1), 65-78.

Fleming, P. (2021). *Dark academia. How universities die*. Pluto Press.

Fleming, P., Rudolph, J., & Tan, S. (2021). 'Never let a good crisis go to waste'. An interview with Professor Peter Fleming on dark academia, the pandemic and neoliberalism. *Journal of Applied Learning and Teaching*, 4(2), 110-120. <https://doi.org/10.37074/jalt.2021.4.2.14>

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681-694.

Fyfe, P. (2022). How to cheat on your final paper: Assigning AI for student writing. *AI & Society*, 1-11.

García-Gorrostieta, J. M., López-López, A., & González-López, S. (2018). Automatic argument assessment of final project reports of computer engineering students. *Computer Applications in Engineering Education*, 26(5), 1217-1226.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2022). *Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers*. Preprint. bioRxiv.

Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2022). *How well does ChatGPT do when taking the medical licensing exams? The Implications of large language models for medical education and knowledge assessment*. Preprint. medRxiv.

Grace, K., Salvatier, J., Dafoe, A., Zhang, B., & Evans, O. (2018). Viewpoint: When will AI exceed human performance? Evidence from AI experts. *Journal of Artificial Intelligence Research*, 729-754.

Grant, N. and Metz, C. (2022, December 21). A new chat bot is a 'code red' for Google's search business. *New York Times*, <https://www.nytimes.com/2022/12/21/technology/ai-chatgpt-google-search.html>

Grossman, G. (2020, September 3). We're entering the AI twilight zone between narrow and general AI. *Venture Beat*, <https://venturebeat.com/ai/were-entering-the-ai-twilight-zone-between-narrow-and-general-ai/>

Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., ... & Wu, Y. (2023). *How close is ChatGPT to human experts? Comparison corpus, evaluation, and detection*. arXiv preprint arXiv:2301.07597.

Hao, K. (2020a, February 17). The messy, secretive reality behind OpenAI's bid to save the world. *MIT Technology Review*, <https://www.technologyreview.com/2020/02/17/844721/ai-openai-moonshot-elon-musk-sam-altman-greg-brockman-messy-secretive-reality/>

Hao, K. (2020b, September 23). OpenAI is giving Microsoft exclusive access to its GPT-3 language model. *MIT Technology Review*, <https://www.technologyreview.com/2020/09/23/1008729/openai-is-giving-microsoft-exclusive-access-to-its-gpt-3-language-model/>

Hao, K. (2022, December 29). Everything to know about Elon Musk's OpenAI, the maker of ChatGPT. *Augustman*, <https://www.augustman.com/sg/gear/tech/openai-what-to-know-about-the-company-behind-chatgpt/>

Harari, Y. N. (2016). *Homo deus. A brief history of tomorrow*. Harvill Secker.

Hartmann, J., Schwenzow, J., & Witte, M. (2023). *The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation*. arXiv:2301.01768 [cs.CL].

Heaven, W. D. (2020, July 20). OpenAI's new language generator GPT-3 is shockingly good—and completely mindless. *MIT Technology Review*, <https://www.technologyreview.com/2020/07/20/1005454/openai-machine-learning-language-generator-gpt-3-nlp/>

Heaven, W. D. (2022, November 18). Why Meta's latest large language model survived only three days online. *MIT Technology Review*, <https://www.technologyreview.com/2022/11/18/1063487/meta-large-language-model-ai-only-survived-three-days-gpt-3-science/>

- Heilwell, E. (2022, December 7). AI is finally good at stuff, and that's a problem. *Vox*, <https://www.vox.com/recode/2022/12/7/23498694/ai-artificial-intelligence-chatgpt-openai>
- Herman, D. (2022, December 9). The end of high-school English. *The Atlantic*, <https://www.theatlantic.com/technology/archive/2022/12/openai-chatgpt-writing-high-school-english-essay/672412/>
- Heyward, E. (2022, May 16). Using AI to write essays isn't cheating, student says. *Canta*, (5), pp. 6-7. <https://www.canta.co.nz/newsarticle/115524?newsfeedId=1453013>
- hooks, b. (2010). *Teaching critical thinking*. Routledge.
- Huh, S. (2023). Are ChatGPT's knowledge and interpretation ability comparable to those of medical students in Korea for taking a parasitology examination?: A descriptive study. *Journal of Educational Evaluation for Health Professions*, 20(1), <https://doi.org/10.3352/jeehp.2023>
- Jiao, W., Wang, W., Huang, J., Wang, X., & Tu, Z. (2023, January 3). *Is ChatGPT a good translator? A preliminary study*. Preprint. https://wxjiao.github.io/downloads/tech_chatgpt_arxiv.pdf
- Jin, B., & Kruppa, M. (2023, January 5). Cheating with ChatGPT: Can an AI chatbot pass AP Lit? *The Wall Street Journal*, <https://www.wsj.com/articles/chatgpt-creator-openai-is-in-talks-for-tender-offer-that-would-value-it-at-29-billion-11672949279>
- Kantrowitz, A. (December 2, 2022). Finally, an A.I. chatbot that reliably passes "the Nazi test". *Slate*, <https://slate.com/technology/2022/12/chatgpt-openai-artificial-intelligence-chatbot-whoa.html>
- Kefalaki, M., Diamantidaki, F., & Rudolph, J. (2022). Editorial 5 (SI1): Technology and education: Innovation or hindrance?. *Journal of Applied Learning & Teaching*, 5(SI1), 6-11. <https://doi.org/10.37074/jalt.2022.5.s1.1>
- Kefalaki, M., & Karanickolas, S. (2020). Communication's rough navigations: 'Fake' news in a time of a global crisis. *Journal of Applied Learning and Teaching*, 3(1), 29-41. <https://doi.org/10.37074/jalt.2020.3.1.19>
- King, M. R. (2023). The future of AI in medicine: A perspective from a Chatbot. *Annals of Biomedical Engineering*, 51, 291-295. <https://doi.org/10.1007/s10439-022-03121-w>
- King, M. R., & ChatGPT. (2023). A conversation on artificial intelligence, chatbots, and plagiarism in higher education. *Cellular and Molecular Bioengineering*, 1-2.
- Köbis, N., & Mossink, L. D. (2021). Artificial intelligence versus Maya Angelou: Experimental evidence that people cannot differentiate AI-generated from human-written poetry. *Computers in Human Behavior*, 114, 106553.
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study. *Assessing Writing*, 44, 100450.
- Kovanovic, V. (2022, December 15). The dawn of AI has come, and its implications for education couldn't be more significant. *The Conversation*, <https://theconversation.com/the-dawn-of-ai-has-come-and-its-implications-for-education-couldnt-be-more-significant-196383>
- Krügel, S., Ostermaier, A., & Uhl, M. (2023). *The moral authority of ChatGPT*. Preprint. arXiv:2301.07098.
- Krugman, P. (2022, December 6). Does ChatGPT mean robots are coming for the skilled jobs? *New York Times*, <https://www.nytimes.com/2022/12/06/opinion/chatgpt-ai-skilled-jobs-automation.html>
- Kuleto, V., Stanescu, M., Ranković, M., Šević, N. P., Păun, D., & Teodorescu, S. (2021). Extended reality in higher education, a responsible innovation approach for generation y and generation z. *Sustainability*, 13(21), 11814.
- Kung, T. H., Cheatham, M., Medinilla, A., ChatGPT, Sillos, C., De Leon, L., ... & Tseng, V. (2022). *Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models*. medRxiv, 2022-12.
- Kurzweil, R. (2005). *The singularity is near*. Viking Penguin.
- Lauret, J. (2020, July 22). GPT-3: The first artificial general intelligence? *Towards Data Science*, <https://towardsdatascience.com/gpt-3-the-first-artificial-general-intelligence-b8d9b38557a1>
- Lim, V. F. (2022, December 16). ChatGPT raises uncomfortable questions about teaching and classroom learning. *The Straits Times*, <https://www.straitstimes.com/opinion/need-to-review-literacy-assessment-in-the-age-of-chatgpt>
- Liquid Ocelot (2023, January 2). ChatGPT-4, the newest and most advanced AI system, might prompt a major shift in the way we communicate. *Medium*, <https://medium.com/inkwater-atlas/chatgpt-4-the-newest-and-most-advanced-ai-system-might-prompt-a-major-shift-in-the-way-we-fd764f97212c>
- Lütke, T. [@tobi]. (2022, December 1). *Tweets* [Twitter profile]. <https://twitter.com/tobi/s/1598134581726965760?cxt=HHwWgMDT8em7260sAAAA>
- Ma, H., & Slater, T. (2015). Using the developmental path of cause to bridge the gap between AWE scores and writing teachers' evaluations. *Writing & Pedagogy*, 7(2), 395-422.
- Manjoo, F. (2020, July 29). How do you know a human wrote this? *The New York Times*, <https://www.nytimes.com/2020/07/29/opinion/gpt-3-ai-automation.html>
- Marche, S. (2022, December 7). The college essay is dead. Nobody is prepared for how AI will transform academia. *The Atlantic*, <https://www.theatlantic.com/technology/archive/2022/12/chatgpt-ai-writing-college-student-essays/672371/>

- Marcus, G., & Davis, E. (2020, August 22). GPT-3, Iloviator: OpenAI's language generator has no idea what it's talking about. *MIT Technology Review*, https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/?utm_medium=tr_social&utm_campaign=site_visitor.unpaid_engagement&utm_source=Twitter#Echobox=1598658773
- Markoff, J. (2015, December 13). Silicon valley investors to bankroll artificial-intelligence center. *The Seattle Times*, <https://www.seattletimes.com/business/technology/silicon-valley-investors-to-bankroll-artificial-intelligence-center/>
- McAfee, A., & Brynjolfsson, E. (2017). *Machine platform crowd. Harnessing our digital future*. W. W. Norton & Company.
- McMurtrie, B. (2022, December 13). AI and the future of undergraduate writing. *The Chronicle of Higher Education*. <https://www.chronicle.com/article/ai-and-the-future-of-undergraduate-writing>
- McMurtrie, B. (2023, January 5). Teaching: Will ChatGPT change the way you teach?. *The Chronicle of Higher Education*. <https://www.chronicle.com/newsletter/teaching/2023-01-05>
- Metz, C. (2016, April 27). Inside OpenAI, Elon Musk's wild plan to set artificial intelligence free. *Wired*, <https://www.wired.com/2016/04/openai-elon-musk-sam-altman-plan-to-set-artificial-intelligence-free/>
- Metz, C. (2022, December, 10). The new chatbots could change the world. Can you trust them? *The New York Times*, <https://www.nytimes.com/2022/12/10/technology/ai-chat-bot-chatgpt.html?action=click&module=RelatedLinks&pgtype=Article>
- Metz, C., & Weise, K. (2023, January 12). Microsoft bets big on the creator of ChatGPT in race to dominate A.I. *The New York Times*, <https://www.nytimes.com/2023/01/12/technology/microsoft-openai-chatgpt.html>
- Mills, A. (2023a) How do we prevent learning loss due to AI text generators?. *Blog post*. <https://docs.google.com/document/d/1m4e2VAIWbZxoteYSGvCO1eTf6GQIOsrjltJoVod0h6g/edit#heading=h.sc09s7ot25zw>
- Mills, A. (2023b). *AI text generators: Sources to stimulate discussion among teachers*. <https://docs.google.com/document/d/1V1drRG1XIWTBrEwgGqd-cCySUB12JrcoamB5i16-Ezw/edit#heading=h.qlyjuxlccr6>
- Mindzak, M., & Eaton, S. E. (2021). Artificial intelligence is getting better at writing, and universities should worry about plagiarism. *The Conversation*. <https://theconversation.com/artificial-intelligence-is-getting-better-at-writing-and-universities-should-worry-about-plagiarism-160481>
- Montclair State University. (2023, January 11). *Practical responses to ChatGPT*. https://www.montclair.edu/faculty-excellence/practical-responses-to-chat-gpt/?fbclid=IwAR0bQI2bjw52g8XpZwusCT4_MeqUP9GTQZK9_I7gMhnYYP66XhJdRI1X4Vo
- Moore, N. S., & MacArthur, C. A. (2016). Student use of automated essay evaluation technology during revision. *Journal of Writing Research*, 8(1), 149–175.
- Moore, S., Nguyen, H. A., Bier, N., Domadia, T., & Stamper, J. (2022). Assessing the quality of student-generated short answer questions using GPT-3. In *European conference on technology enhanced learning* (pp. 243-257). Springer.
- Murati, M. [@miramurati]. (2022, December 6). *Tweets*. [Twitter profile]. <https://twitter.com/miramurati/status/1599796191243669504>
- Nazari, N., Shabbir, M. S., & Setiawan, R. (2021) Application of artificial intelligence powered digital writing assistant in higher education: Randomized controlled trial. *Heliyon*, 7(5), e07014.
- Nguyen, H. A., Bhat, S., Moore, S., Bier, N., & Stamper, J. (2022). Towards generalized methods for automatic question generation in educational domains. In *European conference on technology enhanced learning* (pp. 272-284). Springer.
- Nisar, S., & Aslam, M. S. (2023, January 14). *Is ChatGPT a good tool for T&CM students in studying pharmacology?* Preprint. <https://ssrn.com/abstract=4324310>
- Nowik, C. (2022, December 17). The robots are coming! The robots are coming! Nah, the robots are here. *Substack*, <https://christinenowik.substack.com/p/the-robots-are-coming-the-robots#details>
- O'Connor, S. & ChatGPT (2023). Open artificial intelligence platforms in nursing education: tools for academic progress or abuse? *Nurse Education in Practice*, 66, 103537. <https://doi.org/10.1016/j.nepr.2022.1>
- OpenAI. (2015, December 2). *Introducing OpenAI*. <https://openai.com/blog/introducing-openai/>
- OpenAI. (2019, July 22). *Microsoft invests in and partners with OpenAI to support us building beneficial AGI*. <https://openai.com/blog/microsoft/>
- OpenAI. (2021, December 16). *WebGPT: Improving the factual accuracy of language models through web browsing*. <https://openai.com/blog/webgpt/>
- OpenAI. (2022, November 30). *ChatGPT: Optimizing language models for dialogue*. <https://openai.com/blog/chatgpt/>
- Ortiz, S. (2022, December 7). What is ChatGPT and why does it matter? Here's what you need to know. *ZDNET*, <https://www.zdnet.com/article/what-is-chatgpt-and-why-does-it-matter-heres-what-you-need-to-know/>
- Pagnamenta, R. (2020, August 26). Forget deepfakes – we should be very worried about AI-generated text. *The Telegraph*, <https://www.telegraph.co.uk/technology/2020/08/26/forget-deepfakes-ai-generated-text-should-worried/>
- Parker, M., Rudolph, J., Shukaitis, S., & Tan, S. (2021). Strategic utopianism and the avoidance of dualisms: An interview

with Martin Parker. *Journal of Applied Learning and Teaching*, 4(Sp. Iss. 1), 60-73. <https://doi.org/10.37074/jalt.2021.4.s1.3>

Pavlik, J. V. (2023). Collaborating with ChatGPT: Considering the implications of generative artificial intelligence for journalism and media education. *Journalism & Mass Communication Educator*, 10776958221149577.

Peng, H., Ma, S., & Spector, J. M. (2019). Personalized adaptive learning: An emerging pedagogical approach enabled by a smart learning environment. *Smart Learning Environments*, 6(1), 1-14.

Piantadosi, S. T. [@spiantado]. (2022, December 4). *Tweets*. [Twitter profile]. <https://twitter.com/spiantado/status/1599462405225881600>

Piper, K. (2022, December 15). ChatGPT has given everyone a glimpse at AI's astounding progress. *Vox*, <https://www.vox.com/future-perfect/2022/12/15/23509014/chatgpt-artificial-intelligence-openai-language-models-ai-risk-google>

Qadir, J. (2022). *Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education*. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.21789434.v1>

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving language understanding by generative pre-training*. Preprint. https://cdn.openai.com/research-covers/language-unsupervised/language_understanding_paper.pdf

Rivas, M. J., Baker, R. B., & Evans, B. J. (2020). Do MOOCs make you more marketable? An experimental analysis of the value of MOOCs relative to traditional credentials and experience. *AERA Open*, 6(4), 2332858420973577.

Romero, A. (2021, September 12). GPT-4 will have 100 trillion parameters — 500x the size of GPT-3. *Towards Data Science*, <https://towardsdatascience.com/gpt-4-will-have-100-trillion-parameters-500x-the-size-of-gpt-3-582b98d82253>

Roose, K. (2022, December 5). The brilliance and weirdness of ChatGPT. *The New York Times*, <https://www.nytimes.com/2022/12/05/technology/chatgpt-ai-twitter.html>

Rudolph, J. (2014). *Massive Open Online Courses (MOOCs) as a disruptive innovation in higher education?* [M.Ed. dissertation, University of Adelaide].

Rudolph, J. (2018). A brief review of Mentimeter – A student response system. *Journal of Applied Learning & Teaching*, 1(1), 35-37. <https://doi.org/10.37074/jalt.2018.1.1.5>

Rudolph, J. (2021). Book review: Fleming, P. (2021). *Dark academia. How universities die*. Pluto Press. *Journal of Education, Innovation and Communication*, 3(1), 109-114. <https://doi.org/10.34097/jecom-3-1-june21-7>

Rudolph, J., & Tan, S. (2022). The war in Ukraine as an opportunity to teach critical thinking. *Journal of Applied Learning and Teaching*, 5(1), 165-173. <https://doi.org/10.37074/jalt.2022.5.1.24>

[org/10.37074/jalt.2022.5.1.24](https://doi.org/10.37074/jalt.2022.5.1.24)

Sandlin, J. (2022, December 19). ChatGPT arrives in the academic world. *Boing Boing*, <https://boingboing.net/2022/12/19/chatgpt-arrives-in-the-academic-world.html>

Santandreu Calonge, D., Aman Shah, M., Riggs, K., & Connor, M. (2019). MOOCs and upskilling in Australia: A qualitative literature study. *Cogent Education*, 6(1), 1687392.

Schindler, L. A., Burkholder, G. J., Morad, O. A., & Marsh, C. (2017). Computer-based technology and student engagement: A critical review of the literature. *International Journal of Educational Technology in Higher Education*, 14(1), 25. <https://doi.org/10.1186/s41239-017-0063-0>

Scott, K. (2020). Microsoft teams up with OpenAI to exclusively license GPT-3 language model. *Official Microsoft Blog*. <https://blogs.microsoft.com/blog/2020/09/22/microsoft-teams-up-with-openai-to-exclusively-license-gpt-3-language-model/>

Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*, 3(3), 417-457. doi: 10.1017/S0140525X00005756.

Sharples, M. (2022a). Automated essay writing: An AIED opinion. *International Journal of Artificial Intelligence in Education*, 32(4), 1119-1126.

Sharples, M. (2022b, May 17). New AI tools that can write student essays require educators to rethink teaching and assessment. *Blog*. London School of Economics. <https://blogs.lse.ac.uk/impactofsocialsciences/2022/05/17/new-ai-tools-that-can-write-student-essays-require-educators-to-rethink-teaching-and-assessment/>

Shepherd, T. (2021, January 21). South Australian universities to allow use of artificial intelligence in assignments, if disclosed. *The Guardian*, <https://www.theguardian.com/australia-news/2023/jan/21/south-australian-universities-to-allow-use-of-artificial-intelligence-in-assignments-if-disclosed>

Sparrow, J. (2022). The machines are learning: How artificial intelligence is already affecting journalism education. *Australian Journalism Review*, 44(2), 163-169.

Stevens, D. S. (2019). *Write more, publish more, stress less!* Stylus.

Stokel-Walker, C. (2022). AI bot ChatGPT writes smart essays—should academics worry?. *Nature*. <https://doi.org/10.1038/d41586-022-04397-7>

Susnjak, T. (2022). *ChatGPT: The end of online exam integrity?*. Preprint. arXiv:2212.09292v1

Sutton, M. J., & Allen, K. (2019). *Emotify!: The power of the human element in game-based learning, serious games, and experiential education*. El Games LLC.

Tack, A., & Piech, C. (2022). *The AI teacher test: Measuring*

the pedagogical ability of blender and GPT-3 in educational dialogues. Preprint. arXiv:2205.07540.

Taguma, M., Feron, E., & Lim, M. H. (2018). *Future of education and skills 2030: Conceptual learning framework*. Organization of Economic Co-operation and Development.

Tan, E. (2022). 'Heartware' for the Compassionate Teacher: Humanizing the academy through mindsight, attentive love, and storytelling. *Journal of Applied Learning & Teaching*, 5(2), 152-159. <https://doi.org/10.37074/jalt.2022.5.2.ss1>

Tan, S. (2019). The rise of immersive learning. *Journal of Applied Learning and Teaching*, 2(2), 91-94. <https://doi.org/10.37074/jalt.2019.2.2.12>

Tan, S. (2020). Artificial Intelligence in education: Rise of the machines. *Journal of Applied Learning and Teaching*, 3(1), 129-133. <https://doi.org/10.37074/jalt.2020.3.1.17>

Tate, T., Doroudi, S., Ritchie, D., & Xu, Y. (2023). *Educational research and AI-generated writing: Confronting the coming tsunami*. Pre-print. <https://edaxiv.org/4mec3/>

Terzian, S. (2019). The history of technology and education. In J. L. Rury, & E. H. Tamura. (Eds.), *The Oxford handbook of the history of education* (pp. 554-567).

The Economist. (2020, August 6). A new AI language model generates poetry and prose. <https://www.economist.com/science-and-technology/2020/08/06/a-new-ai-language-model-generates-poetry-and-prose>

Thompson, D. (2022, December 18). Breakthroughs of the year. *The Atlantic*, <https://www.theatlantic.com/newsletters/archive/2022/12/technology-medicine-law-ai-10-breakthroughs-2022/672390/>

Today Online. (2019, February 23). Chinese schoolgirl shamed for using robot to write homework. Now everybody wants one. <https://www.todayonline.com/world/chinese-schoolgirl-shamed-using-robot-write-homework-now-everybody-wants-one>

Tung, L. (2022, December 5). Stack overflow temporarily bans answers from OpenAI's ChatGPT chatbot. *ZDNET*, <https://www.zdnet.com/article/stack-overflow-temporarily-bans-answers-from-openais-chatgpt-chatbot/>

Tung, L. (2023, January 4). ChatGPT's next big challenge: Helping Microsoft to challenge Google search. *ZDNET*, <https://www.zdnet.com/article/chatgpts-next-big-challenge-helping-microsoft-to-challenge-google-search/>

Van Bergen, P., & Lane, R. (2016, November 30). Should we do away with exams altogether? No, but we need to rethink their design and purpose. *The Conversation*, <https://theconversation.com/should-we-do-away-with-exams-altogether-no-but-we-need-to-rethink-their-design-and-purpose-67647>

Vincent, J. (2016, March 24). Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. *The Verge*,

<https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>

Vincent, J. (2022a, December 5). AI-generated answers temporarily banned on coding Q&A site. *Stack Overflow*. <https://www.theverge.com/2022/12/5/23493932/chatgpt-ai-generated-answers-temporarily-banned-stack-overflow-llms-dangers>

Vincent, J. (2022b, December 8). ChatGPT proves AI is finally mainstream — and things are only going to get weirder. *The Verge*, <https://www.theverge.com/2022/12/8/23499728/ai-capability-accessibility-chatgpt-stable-diffusion-commercialization>

Vincent, J. (2022c, December 2). OpenAI's new chatbot can explain code and write sitcom scripts but is still easily tricked. *The Verge*, <https://www.theverge.com/23488017/openai-chatbot-chatgpt-ai-examples-web-demo>

Warren, T. (2023, January 9). Microsoft is looking at OpenAI's GPT for Word, Outlook, and PowerPoint. *The Verge*, <https://www.theverge.com/2023/1/9/23546144/microsoft-openai-word-powerpoint-outlook-gpt-integration-rumor>

Warner, J. (2022, December 5). Freaking out about ChatGPT. Part I. *Inside Higher Ed*, <https://www.insidehighered.com/blogs/just-visiting/freaking-out-about-chatgpt%E2%80%94part-i>

Wiggers, K. (2020, June 1). OpenAI's massive GPT-3 model is impressive, but size isn't everything. *VentureBeat*, <https://venturebeat.com/ai/ai-machine-learning-openai-gpt-3-size-isnt-everything/>

Wiggers, K. (2022, December 10). OpenAI's attempts to watermark AI text hit limits. *TechCrunch*, <https://techcrunch.com/2022/12/10/openais-attempts-to-watermark-ai-text-hit-limits/>

Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research, and Evaluation*, 2(1), 2.

William, D. (2011). What is assessment for learning?. *Studies in Educational Evaluation*, 37(1), 3-14.

Yeadon, W., Inyang, O. O., Mizouri, A., Peach, A., & Testrow, C. (2022). *The death of the short-form physics essay in the coming AI revolution*. Preprint.

Yousif, N. (2023, January 13). ChatGPT: Student builds app to sniff out AI-written essays. *BBC*, <https://www.bbc.com/news/world-us-canada-64252570>

YP. (2019, February 19). *Chinese schoolgirl caught using robot to write her homework – now everybody wants one*. <https://www.scmp.com/yp/discover/entertainment/tech-gaming/article/3060907/chinese-schoolgirl-caught-using-robot-write>

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators?.

Zhai, X. (2022). *ChatGPT user experience: Implications for education*. https://www.researchgate.net/profile/Xiaoming-Zhai/publication/366463233_ChatGPT_User_Experience_Implications_for_Education/links/63a2817aca6a9d254f8de010/ChatGPT-User-Experience-Implications-for-Education.pdf

Appendices

Appendix A: Essay provided by ChatGPT (plagiarism 4% using Grammarly Professional version).

Auditory hallucinations, or hearing voices, is a common symptom associated with schizophrenia. However, the presence of auditory hallucinations alone is not indicative of schizophrenia and can also occur in other mental disorders or medical conditions. This is an important distinction to make as an accurate diagnosis is crucial for effective treatment.

Schizophrenia is a severe and chronic mental disorder characterized by a range of symptoms including delusions, hallucinations, disordered thinking and abnormal emotional responses. The onset of the disorder usually occurs in late adolescence or early adulthood, and it affects around 1% of the population worldwide. Auditory hallucinations, which are experienced by around 70-80% of individuals with schizophrenia, are considered to be one of the most common symptoms of the disorder. However, the presence of auditory hallucinations alone is not a sufficient criterion for a diagnosis of schizophrenia.

The diagnostic criteria for schizophrenia as outlined by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) includes the presence of two or more of the following symptoms: delusions, hallucinations, disordered thinking, abnormal emotional responses and a reduction in the individual's ability to function. The diagnosis of schizophrenia is made when an individual experiences these symptoms for a significant period of time, with a duration of at least one month. Auditory hallucinations, therefore, are not enough to diagnose an individual with schizophrenia.

Auditory hallucinations can also occur in other mental disorders, such as bipolar disorder, major depressive disorder and post-traumatic stress disorder (PTSD). For example, individuals with bipolar disorder may experience auditory hallucinations during manic episodes. In major depressive disorder, auditory hallucinations can occur during a major depressive episode. Similarly, in PTSD, auditory hallucinations can occur as a result of traumatic experiences. It is therefore important to take a holistic approach when assessing an individual with auditory hallucinations and to

consider other symptoms and factors.

Auditory hallucinations can also occur in individuals without any underlying mental disorder, in response to certain medical conditions such as brain tumors, infections, or head injuries. In these cases, auditory hallucinations can be caused by structural or functional changes in the brain. Therefore, it is important to rule out any underlying medical conditions before diagnosing an individual with schizophrenia.

Another important consideration is that individuals with schizophrenia may not experience auditory hallucinations all the time. The severity and frequency of auditory hallucinations can vary over time and during different stages of the disorder. For example, some individuals may experience auditory hallucinations only occasionally, while others may experience them frequently. Therefore, a diagnosis of schizophrenia cannot be based solely on the presence of auditory hallucinations.

In conclusion, the presence of auditory hallucinations alone is not indicative of schizophrenia. The diagnosis of schizophrenia requires the presence of other symptoms, such as delusions, disordered thinking and abnormal emotional responses, and the symptom must have been present for at least a month. It is also important to consider other mental disorders and medical conditions that may cause auditory hallucinations. An accurate diagnosis is crucial for effective treatment and it is important to take a holistic approach when assessing individuals with auditory hallucinations.

Appendix B: A rough translation of the Chinese passage.

Made a mistake:

I once made a mistake that broke my parents' hearts. Some classmates and I played around in high school and decided to drink. After getting drunk at a party, I drove a car while under the influence of alcohol and caused a car accident. Fortunately, no one was hurt, but I lost the trust and respect of my parents.

Regret:

I deeply regret my reckless behaviour at the time. I know my actions have caused great harm to my parents, and I have not respected their care and love for me. I know my mistakes cannot be forgiven. I will always remember this experience and never make similar mistakes again. I will continuously repent for my mistakes and do everything I can to regain their trust.