



www.ijres.net

Critical Evaluation of the Potential of Large Language Models in Bioscience Higher Education: A Conversation with ChatGPT

Andrew Williams 
University College London, United Kingdom

To cite this article:

Williams, A. (2025). Critical evaluation of the potential of large language models in bioscience higher education: A conversation with ChatGPT. *International Journal of Research in Education and Science (IJRES)*, 11(3), 667-701. <https://doi.org/10.46328/ijres.1302>

The International Journal of Research in Education and Science (IJRES) is a peer-reviewed scholarly online journal. This article may be used for research, teaching, and private study purposes. Authors alone are responsible for the contents of their articles. The journal owns the copyright of the articles. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of the research material. All authors are requested to disclose any actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations regarding the submitted work.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License.

Critical Evaluation of the Potential of Large Language Models in Bioscience Higher Education: A Conversation with ChatGPT

Andrew Williams

Article Info

Article History

Received:

28 November 2024

Accepted:

25 April 2025

Keywords

Artificial intelligence

Large language model

Higher education

Bioscience

ChatGPT

Curriculum

Ethics

Abstract

The possibilities for integrating generative artificial intelligence (AI) and large language models (LLMs) into higher education may revolutionise approaches to pedagogical practices and curriculum design, while LLMs could be transformative in how students approach their learning. This conversation with ChatGPT, and associated critical evaluation, provides an insight into the capabilities and limitations of LLMs and informs on the possibilities for incorporating AI into bioscience education, teaching modalities, assessment and feedback practices and curriculum design, with a particular focus on bioscience education. The conversation highlights how LLMs can facilitate personalized feedback, tutoring, and concise explanations of scientific terms, enhancing student comprehension, self-directed learning, and critical thinking. However, there are concerns regarding the risks of LLMs, such as plagiarism and breaches of academic integrity, algorithmic bias and limitations in contextual understanding. Despite these limitations, AI offers opportunities for enriching undergraduate bioscience curricula by integrating innovative teaching strategies and assessment modalities aligned with subject benchmark statements. Future research directions include exploring ethical implications, equitable access and digital literacy training in higher education settings. Overall, the integration of LLMs in bioscience education offers significant potential for innovative pedagogical approaches and transformative learning experiences.

Introduction

In recent years, generative artificial intelligence (AI) technologies and large language models (LLMs) have catalysed innovative approaches to teaching and learning across diverse disciplines. The release of ChatGPT on November 30, 2022, was received with equal amounts of trepidation and excitement among the higher education community (Rudolph et al., 2023). Natural language processing (NLP), a subtype of AI, enables computers to interpret and mimic human language. ChatGPT and other generative AI tools use advanced NLP technologies, machine learning and deep learning models to create coherent and contextually appropriate output in response to text inputs or prompts. These models are trained on vast resources of data. For example, ChatGPT 3.5 was trained on billions of words and terabytes of data appropriated from the corpus of information existing on the internet (<https://chat.openai.com/>). LLMs use deep learning techniques, including generative pre-trained transformer models (GPT) models and neural networks, to predict word sequences based on contextual relationships and

probability (Bengio et al., 2003).

AI is a multidisciplinary field, with broad definitions that encompass computer science, mathematics, neuroscience and behavioural science. AI can be defined as an artificial system that is capable of independently performing creative functions, solve problems and be able to learn without human interference (Xu et al., 2021). LLMs such as ChatGPT (OpenAI), Gemini (Google) and Claude (Anthropic) fulfil this definition by creating human-quality text, while related technologies like DALL-E, Midjourney and Stable Diffusion (Stability AI) can create realistic images based on simple text prompts (Chan & Colloton, 2024). Moreover, the perceived human quality of interactions with generative AI systems and the interactive nature of LLMs promotes an immersive exchange of information and ideas (Bettayeb et al., 2024; Eysenbach, 2023). The potential of LLMs for enhancing the student learning experience and positively transforming teaching and learning experiences is therefore evident (Bobula, 2024). However, the opportunities that generative AI affords have been balanced with ethical concerns and reservations regarding its negative impact on academic integrity, student learning and cognitive development (Kasneci et al., 2023). There is therefore a need to formulate robust policies and frameworks that support AI integration into higher education and to promote professional development and training that enables teachers to make informed judgements on the practical and ethical adoption of AI technologies (Francis et al., 2025).

The integration of LLMs into higher education contexts represents a challenge for educators as they navigate the complexities of designing curricula that align with evolving industry demands (Fuchs, 2023). The utilization of LLMs emerges as a compelling strategy, as these tools have the capacity to generate vast repositories of knowledge, enabling educators to tailor instructional materials to meet the diverse needs of learners and to support the development of innovative pedagogical approaches (Guo et al., 2024; Li & Wong, 2023). Numerous studies have evaluated the potential of using generative AI to enhance medical education, including personalised learning, creating clinical case studies and medical scenarios, and providing academic writing support (Preiksaitis & Rose, 2023). ChatGPT has also demonstrated its aptitude on several science and medical exams, including the United States Medical Licensing Examinations exam (Kung et al., 2023; Rosoł et al., 2023) and is capable of understanding complex medical knowledge and diagnoses such as surgical information, medical imaging and clinical examples (Ali et al., 2023; Huh, 2023; Oh et al., 2023; Wang et al., 2023). However, concerns were raised regarding bias, hallucinations and academic integrity (Tian et al., 2024). Within the related discipline of bioscience education, the advent of ChatGPT and LLMs may represent an equally promising technology for enhancing pedagogical practices, curriculum design, and student learning and engagement (Baidoo-Anu & Ansah, 2023).

Although the utility of generative AI has been explored in relation to bioscience research (Thapa & Adhikari, 2023), including enhancing access to scientific knowledge, complex data analysis and bioinformatics (Varghese & Chapiro, 2024), little exploration of its potential in bioscience education has been undertaken (Akpan et al., 2025). Bioscience and biomedical research have experienced an explosion in technological and multidisciplinary advancements in recent decades, which are set to transform human understanding of biological processes with real-world impact on medicine and biotechnology. It is therefore imperative that bioscience education keeps pace with advances in technology and addresses complex interdisciplinary problems effectively (Labov et al., 2010). For example, LLMs may facilitate knowledge accessibility by simplifying scientific concepts and terminology,

providing concise explanations, synthesising complex research papers, answering queries and generating summaries for researchers and students (Rane, 2023).

In the following conversation with ChatGPT, we discuss various aspects of bioscience education, with a particular emphasis on key biological concepts, and explore relevant themes in the context of the potential of LLMs in redefining teaching and learning practices. The intention of the conversation was to provide an overview of the capabilities and limitations of existing LLMs as a potential resource for designing biosciences and biology curricula, feedback and assessment strategies, teaching and learning resources and supporting student-centred learning. Furthermore, the commentary aims to support dialogue among educators and bioscience students on how generative AI can positively transform higher education and to reflect on the limits of these technologies in the context of the biosciences.

The conversation between myself and ChatGPT (version 4o; <https://chat.openai.com/>) occurred during July 2024. The specific themes covered were designed to reflect bioscience and bioscience education, using the QAA subject benchmark statement in biosciences as a framework for question proposition (Watson et al., 2023), and consisted of several subject themes including the general role of LLMs in bioscience education, personalized education and feedback, student learning resources, assessment and feedback practices, academic integrity, ethics and bias, and curriculum design. The conversation and textual responses that were generated by ChatGPT have been edited or abridged for brevity, with some repetitive elements removed, and the full transcript is available from the author on request. Specific examples pertaining to bioscience education have been emphasised.

Potential Roles of Large Language Models

The recent introduction of AI-driven chatbots, like ChatGPT, has revolutionized the way automated systems interact with humans. Generative AI chatbots, like ChatGPT, simulate human conversation by generating natural language responses without requiring predefined examples (Ouyang et al., 2022). The performance of ChatGPT can be further improved by human feedback through simple prompts or through chain-of-thought prompting that uses successive reasoning steps to generate expected outcomes or more accurate responses (Wei et al., 2022). This technology has immense potential for applications in various sectors, including higher education, indicating a significant shift in the adoption of human-computer interaction into teaching practices.

The utility of LLMs extends into critical areas like healthcare and biomedical research, as well as bioscience education (Varghese & Chapiro, 2024). In healthcare, generative AI can assist in patient management, providing personalized responses to medical queries, and facilitating communication between patients and healthcare providers (Ayers et al., 2023; Zhang & Kamel Boulos, 2023). Researchers are beginning to utilize the capabilities of LLMs for data analysis, article summarisation and generating hypotheses, although concerns remain about the ethical implications of AI in terms of authorship, academic scholarship and biases (Bagenal, 2024; Blau et al., 2024). The versatility of generative AI has the further potential to transform bioscience education. Higher education institutions can leverage LLMs for a variety of educational tasks, while students are increasingly accepting generative AI as a tool for their studies and approach to learning (Arowosegbe et al., 2024). Although

LLMs show potential for enhancing student participation and providing learning support, concerns regarding plagiarism, over-reliance and the degradation of cognitive and critical thinking skills persist (Baidoo-Anu & Ansah, 2023; Francis et al., 2025). As a starting point in the conversation with ChatGPT, the first questions aimed to address the general roles that generative AI could play in bioscience education (see Table 1).

Question: What is the role of generative artificial intelligence and large language models in bioscience higher education?

ChatGPT: Generative artificial intelligence (AI) and large language models (LLMs) like GPT-4 have the potential to significantly transform bioscience higher education in several ways. Here are some key roles they can play:

Table 1. Examples of Generative AI Uses in Bioscience Education

Generative AI application	Examples of use	Benefits
Personalized Learning and Tutoring	offering explanations, answering questions, and adapting to individual learning paces and styles.	beneficial in complex subjects such as molecular biology, genetics, and biochemistry, where students often need tailored assistance to grasp difficult concepts.
Content Creation and Enhancement	creating and enhancing educational content, including lecture materials, assignments, and exams. It can generate high-quality diagrams, simulations, and interactive modules.	can make learning more engaging and effective.
Research Assistance	help students and researchers with literature reviews, summarizing scientific papers, and suggesting relevant research directions. They can also assist in data analysis and interpretation.	provide insights that might be missed by human researchers.
Language and Communication Support	For non-native English speakers, LLMs can aid in understanding and producing scientific texts; assist in translating complex scientific information.	enhancing student's ability to participate fully in the academic community; translating information into more accessible language.
Administrative and Logistical	automate administrative tasks such as	free up educators to focus more on

Generative AI application	Examples of use	Benefits
Support	grading, scheduling, and communication; help in managing large datasets and providing feedback on student progress.	teaching and mentoring.
Virtual Laboratories and Simulations	create virtual laboratory environments where students can perform experiments and simulations.	valuable when access to physical labs is limited, allowing students to gain hands-on experience in a virtual setting.
Ethical and Societal Implications	teach students about the ethical, legal, and societal implications of advancements in biosciences, including issues related to genetic engineering, biotechnology, and personalized medicine.	foster a deeper understanding of the impact of their work on society.
Collaboration and Interdisciplinary Learning	facilitate collaboration among students and researchers from different disciplines by providing a common platform for knowledge exchange and problem-solving.	integrate information from various fields, helping to create a more interdisciplinary approach to bioscience education.
Continuous Assessment and Feedback	offer continuous assessment and instant feedback to students; real-time feedback can enhance learning outcomes.	helping them understand their strengths and areas for improvement; keep students engaged and motivated.

Challenges and Considerations: While the potential benefits are significant, there are also challenges and considerations to address:

Quality Control: Ensuring the accuracy and reliability of AI-generated content.

Bias and Ethics: Mitigating biases in AI models and addressing ethical concerns related to data privacy and academic integrity.

Faculty Training: Educating faculty on how to effectively integrate AI tools into their teaching.

Resource Availability: Ensuring equitable access to AI tools and technologies for all students.

The output can be broadly categorised into student-focussed and educator-focussed in terms of the application of ChatGPT and LLMs in bioscience education. Personalised learning and tutoring using LLMs have significant potential for helping students understand complex bioscience concepts in areas such as evolution, cell biology,

genetics and biochemistry. Several bespoke generative AI applications are indeed now available for biology students, such as those provided by FlowGPT (<https://flowgpt.com/p/biology-tutor>) and AI MegaLabs (<https://www.aimegalabs.com/biology-tutor-assist-ai/>), while large, online education providers such as Khan Academy (<https://www.khanmigo.ai/>) are also offering AI-powered personal tuition for the biosciences. Some of the concerns around AI generated bias, inaccuracies and hallucinations are also being circumvented by training these applications on tailored data sets, specific to biology and related science subjects (Ahmad et al., 2023). Generative AI may be particularly useful for translating complex biological information and terminology into manageable and accessible language, especially for students undertaking multidisciplinary studies.

Nevertheless, the generative AI output may still require educator supervision or critical appraisal by the user to ensure scientific accuracy, especially in areas that require complex mechanistic or conceptual understanding (Cooper, 2023; Wu, 2023). Rather than liberating more time for teachers to explore concepts with students in depth, the additional supervision or cross-checking of AI generated material may become more burdensome, especially if over-reliance on this technology becomes commonplace. Similarly, generative AI has potential as a personal tutor for individualised learning paths, particularly for continuous assessment and feedback. However, appropriate prompt engineering is an important consideration in this context, especially if asking the LLM to take on a specific role as tutor, for example in role-playing scenarios (Gregorcic & Pendrill, 2023; Mohebi, 2024) or when applying the Socratic method of tuition (see following question and response).

From the perspective of an educator, ChatGPT has considerable potential to assist the development of new course material, for generating teaching content and assessments, enhancing administrative efficiencies and for developing curricula (Eysenbach, 2023; Kayabas, 2024). As highlighted by ChatGPT, training of faculty members, not just with regard to familiarisation of application interfaces and tool capabilities but also on the ethical considerations and limitations of generative AI, may be essential for this approach to be successful (Alasadi & Baiz, 2023; Mah & Groß, 2024). Concerns about privacy and security, particularly where student data is concerned, may also limit the use of generative AI in certain contexts, such as using LLMs to provide feedback on student submissions (Chan, 2023; Elliott & Soifer, 2022).

Personalised Education and Learning Support

Due to their adaptability, the iterative nature of AI interactions and their vast knowledge base, LLMs may be well-suited to supporting personalized education in universities. In this context, generative AI could be used to adapt learning support for individual student's needs, taking into consideration diverse learning styles, different stages of educational developmental and personal preferences (Guo et al., 2024; Li & Wong, 2023). Students can also utilise generative AI to support their autonomous or self-regulated learning, by creating study plans, quiz questions, learning resources and obtaining personalised feedback and tutoring from AI systems (Chang et al., 2023; Jin et al., 2023; Lai, 2024). This may also be relevant in the current higher education climate that is witnessing ever increasing student cohort sizes and may alleviate some of the inherent problems associated with mass education, such as poor engagement, diverse learning requirements, limited resource allocation and diluted faculty contact hours (Dempere et al., 2023; Marginson, 2006). To explore this potential further, ChatGPT was

asked how LMMs can aid personalised education and learning (see Table 2).

Question: I'm interested in learning more about how LLMs can aid personalized education and learning. How would this be achieved?

ChatGPT: Large language models (LLMs) can significantly enhance personalized education and learning in several ways. Here's a detailed look at how this can be achieved:

Table 2. How Can LLMs Aid Personalized Education and Learning?

Personalised learning	Example	Benefits
Tailored Educational Content	customized learning materials based on a student's current knowledge level, learning style, and interests; adaptive textbooks and learning modules; customized problem sets.	adjusts in complexity and depth based on student performance and feedback; tailored to focus on areas where a student needs the most practice.
Interactive Tutoring	virtual tutors; real-time Q&A; guided problem solving.	provide on-demand assistance and explanations; receive instant, detailed explanations; step-by-step guidance, ensuring students understand each stage before moving on.
Feedback and Assessment	immediate and personalized feedback on assignments and exams; automated grading; constructive feedback.	quick and consistent grading of quizzes, tests, and written assignments; detailed feedback highlighting strengths and areas for improvement, tailored to individual performance.
Study Plans and Progress Tracking	create and monitor personalized study plans; goal setting; progress reports.	assisting students in setting realistic academic goals based on their current performance and desired outcomes; regular updates on progress.
Enhancing Engagement	make learning more engaging and interactive; gamification; interactive simulations.	make studying more enjoyable; experiment and learn in a hands-on manner.

Personalised learning	Example	Benefits
Support for Diverse Learning Styles	cater to various learning preferences; visual and audio content; generating diagrams, charts and audio explanations; reading comprehension assistance.	support students who learn better through visual or auditory means; providing simpler explanations to aid understanding.
Language Support	assist non-native speakers and those with language difficulties; translation and simplification; language learning.	translating technical terms and complex sentences into simpler language; improve language skills through interactive exercises and real-time correction.
Encouraging Critical Thinking	challenge students to think critically and creatively; Socratic method; debates and discussions.	encourage deeper thought and understanding; helping students develop and articulate their ideas.

Implementation Strategies - To effectively integrate LLMs into personalized education, consider the following strategies:

1. Integration with Learning Management Systems (LMS): Embed AI tools within existing LMS platforms to provide seamless access to personalized content and tutoring.
2. Data-Driven Insights: Use student data (with proper consent and privacy safeguards) to tailor educational experiences and predict future performance.
3. Training Educators: Provide training for educators to effectively utilize AI tools and interpret AI-driven insights.
4. Feedback Loops: Continuously gather feedback from students and educators to refine and improve AI-generated content and interactions.
5. Scalability and Accessibility: Ensure AI tools are scalable to accommodate large numbers of students and accessible to those with different needs and backgrounds.

Of particular interest is the integration of generative AI technologies into existing virtual learning environments (VLE) or digital learning management systems (LMS). However, this poses significant logistical challenges for higher education institutions and digital education teams. Data security, scalability and accessibility are important considerations, with equal and equitable access a serious challenge. Students have a diverse range of technical abilities and dissimilar educational needs, which would need to be addressed at a programmatic and institutional level (Nikolopoulou, 2025). With regard to bioscience education, any integrated LLM system would also have to align with the subject-specific learning material and programme learning outcomes and therefore be designed appropriately and judiciously.

Interactive tutoring and tailored education content may be of particular benefit to bioscience students. LLMs are able to draw on a huge corpus of information and provide subject-specific explanations of myriad biological topics. However, LLMs do present challenges when navigating content accuracy, content bias and hallucinations (Li et al., 2024). These may have important implications for student learning in the biosciences, as exemplified by ChatGPT responses when prompted to describe the pathogenesis of osteoporosis or metabolic diseases (Alkaissi & McFarlane, 2023). This study highlighted the inadequacies of ChatGPT retrieving genuine scientific references, the challenges of relying on ChatGPT for accurate scientific information and emphasised the need for critical evaluation and fact-checking. Although generated output was considered to cover the basic mechanisms and aid academic writing, ChatGPT and other LLMs may have difficulty explaining complex biomedical and biological concepts, processes and mechanisms. Furthermore, despite LLMs being capable of reproducing vast quantities of knowledge, they are far less proficient at replicating human-level creativity and have no conceptual understanding of generated output (Borger et al., 2023), skills which are highly regarded in the biosciences and related professions. To gain further insight into the potential of LLMs to provide tailored educational content, a specific biological example was requested (see Table 3).

Question: Could you provide more details on 'tailored educational content', please? Could you use the topic of 'human evolution' as an example of how this might be achieved?

ChatGPT: Tailored educational content involves customizing learning materials to meet the specific needs, interests, and learning styles of individual students. Using the topic of "human evolution" as an example, here's how this can be achieved:

Table 3. Examples of Tailored Learning Adjusted to Student Level

Tailored content	Basic Level	Intermediate Level	Advanced Level
Adaptive Textbooks and Learning Modules: Create digital textbooks and modules that adapt in complexity based on a student's understanding and progress.	For beginners, the content would include fundamental concepts like the definition of evolution, an overview of natural selection, and the timeline of human evolution.	For students with some prior knowledge, introduce more detailed information on hominid species, genetic evidence, and evolutionary mechanisms.	For advanced students, delve into current research, debates, and complex theories in human evolution.
Customized Problem Sets and Assignments: Design exercises that focus on the specific areas where a student needs practice.	Simple quizzes and matching exercises to test basic understanding of key concepts and terms.	Short answer questions and basic analytical tasks.	Research assignments and critical analysis of scientific papers.
Interactive Learning	Simple, guided	Moderate complexity,	Complex simulations

Tailored content	Basic Level	Intermediate Level	Advanced Level
Activities: Use interactive modules and simulations to make learning engaging.	interactions to build foundational understanding.	requiring the student to connect multiple concepts.	that require critical thinking and synthesis.
Multimedia Content: Incorporate videos, animations, and infographics to cater to different learning styles.	Simplified multimedia explaining core concepts.	More detailed multimedia resources that include secondary concepts.	Complex multimedia that integrates recent scientific discoveries.
Real-Time Q&A and Feedback: Provide platforms where students can ask questions and receive instant, personalized answers.	Direct, simple answers to foundational questions.	Detailed explanations with connections to broader concepts.	Complex, research-based discussions or debate-style feedback.

Evolution is a core topic in biology higher education curricula and therefore provides a relevant example for the application of LLMs proposed by ChatGPT. The different level descriptions (basic, intermediate and advanced) are appropriate for the design of teaching content and suitable for adaption to differing student capabilities and stages of learning throughout a programme of study. For example, the level descriptions could be applied to an undergraduate BSc Biology or BSc Bioscience programme with each ascending level analogous to year 1 (FHEQ level 4, basic), year 2 (FHEQ level 5, intermediate) and year 3 (FHEQ level 6, advanced) modules of study, as outlined by the UK Quality Assurance Agency (QAA) (QAA, 2024a). These strategies proposed by ChatGPT could function as starting points for educators wishing to incorporate generative AI into their teaching and learning resources and adapted to distinctive student competences appropriate for studying undergraduate biosciences. A similar application is for LLMs to act as virtual tutors to aid student learning. Therefore, ChatGPT was asked to provide more detail on how LLMs could act as interactive tutors and to provide practical examples.

Question: I'm interested in learning more about how LLMs could act as 'interactive or virtual tutors' to aid student learning. Could you explain this in more detail, please, and provide an example with 'human evolution' as the topic?

ChatGPT: Large language models (LLMs) can serve as interactive or virtual tutors by providing real-time, personalized support to students, answering their questions, guiding them through complex topics, and offering feedback. Here's a detailed explanation of how this could work, using "human evolution" as the example topic:

1. Real-Time Q&A - Students can ask the LLM questions at any time and receive immediate, detailed responses.
Example interaction:

Student: "What is the significance of Australopithecus in human evolution?"

LLM: "Australopithecus is significant because it represents one of the early stages of hominid evolution, showing a mix of ape-like and human-like features. It walked upright on two legs, which is a key characteristic of hominids, and provides insight into the transition from tree-dwelling to terrestrial life."

2. Step-by-Step Guidance - The LLM can guide students through complex problems or topics in a step-by-step manner, ensuring they understand each part before moving on. Example interaction:

Student: "Can you help me understand the Out of Africa theory?"

LLM: "Of course! The Out of Africa theory suggests that all modern humans originated from a common population in Africa. Let's break it down:

- a. Origin in Africa: Homo sapiens evolved in Africa around 200,000 years ago.
- b. Migration: Small groups of Homo sapiens began migrating out of Africa around 60,000-70,000 years ago.
- c. Replacement: As these groups moved into Europe and Asia, they replaced local hominid populations like Neanderthals and Denisovans.
- d. Evidence: This theory is supported by genetic evidence showing that all humans share common mitochondrial DNA, indicating a single origin.

3. Personalized Feedback and Explanations - The LLM can provide personalized feedback on assignments and explain mistakes in detail, helping students learn from their errors.

4. Interactive Simulations and Activities - The LLM can facilitate virtual labs or simulations where students can engage in hands-on activities.

5. Adaptive Learning Paths - The LLM can create and adapt learning paths based on the student's progress and areas of difficulty.

6. Encouraging Critical Thinking and Exploration - The LLM can pose thought-provoking questions and encourage students to explore topics in depth, fostering critical thinking. Example interaction:

Student: "Why did Neanderthals go extinct while Homo sapiens survived?"

LLM: "Great question! There are several theories....."

There have been conflicting reports on the potential of ChatGPT and other LLMs to support student's critical thinking skills. While some students demonstrate the potential benefit of leveraging LLMs to benefit critical thinking competencies, other have noted the detrimental effects on student critical thinking (Deng & Yu, 2023; Ruiz-Rojas et al., 2024). Critical thinking is considered to be a higher order cognitive skill, is a core feature of the bioscience benchmark statement and is an important competency for bioscience students (Watson et al., 2023). Generative AI models have been considered to be poor at critical and independent thinking compared to humans, with assessments that incorporate critical thinking criteria considered to be relatively resistant to AI interference (Bharatha et al., 2024; Royer, 2024). Leveraging LLMs to encourage student's critical thinking and cognitive skills is gaining traction in higher education, particularly in the context of personalised tutoring systems and question and answer interactions with generative AI applications (Abas et al., 2023), and as outlined in the responses given by ChatGPT. However, despite advances in LLM algorithms that have been adapted to biological data and concepts, AI output should still be critiqued by educators and students for accuracy.

There are also ethical considerations regarding the extent to which generative AI should be influencing the

educational development of students studying bioscience and related disciplines. These include the risk of algorithmic bias in chatbots perpetuating societal biases, the implications for data privacy and security and the concerns about AI systems becoming epistemic authorities in higher education (Hauswald, 2025; R. T. Williams, 2024). For example, concerns around the use of sensitive patient information raises issues relating to informed consent, confidentiality and data security (Ganesan & Somasiri, 2024; Morandín-Ahuerma, 2023). While AI certainly demonstrates potential for augmenting clinical education and practice, there is a risk of over-reliance that may undermine the development of human clinical and scientific reasoning skills and requirements for professional accountability (Preiksaitis & Rose, 2023; Zhai et al., 2024). The need for human oversight therefore remains, while bioscience and biomedical education programs should encourage learners to critically evaluate AI outputs. Considering the divergent views on the impact of AI on critical thinking skills and cognitive development, ChatGPT was asked how LLMs could be used to encourage student's critical thinking.

Question: You mentioned that LLMs could be used to encourage student's critical thinking. How would this be achieved in the biosciences?

ChatGPT: Encouraging critical thinking in biosciences through the use of Large Language Models (LLMs) can be achieved by leveraging their ability to generate thought-provoking questions, simulate scientific inquiry, facilitate debates, and provide in-depth explanations. Here are several strategies on how this could be implemented:

1. Socratic questioning - LLMs can engage students in Socratic questioning to stimulate deeper thinking and exploration of bioscience concepts.

Student: "Why is genetic variation important in evolution?"

LLM: "Good question. Why do you think genetic variation might be crucial for a species' survival and adaptation?"

Student: "It allows for different traits that can help in changing environments."

LLM: "Exactly. Can you think of an example where genetic variation has played a key role in the survival of a species?"

2. Hypothesis generation and testing - Encourage students to generate hypotheses and think about experimental designs to test them.

Student: "What could have caused the Neanderthal extinction?"

LLM: "That's an interesting question. Let's formulate a hypothesis. One hypothesis could be that climate change affected their food sources. How would you design an experiment or study to test this hypothesis?"

3. Facilitating debates - LLMs can help organize and moderate debates on controversial or complex topics in biosciences.

4. Case study analysis - Present case studies and guide students through analysis and critical evaluation.

LLM: "Here's a case study on antibiotic resistance. Bacteria in a hospital setting have become resistant to multiple antibiotics. What are the potential causes, and how could this issue be addressed?"

5. Conceptual mapping - Help students create conceptual maps to visualize and connect different bioscience concepts.

6. Simulated research projects - Guide students through simulated research projects, from hypothesis to conclusion.

LLM “Imagine you are studying the impact of climate change on polar bear populations. What would be your research question, and how would you go about collecting and analyzing data?”

7. Providing multiple perspectives - Encourage students to consider different perspectives and the complexity of scientific issues.

Student: “What caused the mass extinction of dinosaurs?”

LLM: “There are several hypotheses, including asteroid impact and volcanic activity. Let’s explore the evidence for each. What are the strengths and limitations of these explanations?”

8. Reflective Journaling - Prompt students to reflect on their learning and thought processes.

LLM: “Reflect on what you’ve learned about the role of genetic drift in evolution. How has your understanding changed, and what questions do you still have?”

LLM applications have potential in facilitating learning knowledge exchange, enabling debates and encouraging students to consider different perspectives when engaging with various biological concepts. For example, it has been proposed that generative AI could promote deep thinking and active learning through Socratic dialogue (Ma et al., 2024). This may have benefits in bioscience education by steering students away from an over-reliance of acquiring direct answers and facts from chatbots or tutors and towards an open debate and dialogue. The general approach to active learning and dialogue in bioscience education has been shown to improve scientific reasoning and student understanding of biological concepts (Chien et al., 2024), while the Socratic approach to teaching is considered valuable for promoting critical thinking in medical education (Ho et al., 2023). However, leveraging LLMs in this teaching space does require considered prompt engineering to provide students with an appropriate framework for questioning and to support an effective learning experience.

Creating and Enhancing Education Content

Generative AI is seen by many to have transformative potential in higher education. The concept of AI in education (AIEd) can trace its origins to early computer-assisted instruction systems of the 1960s and 1970s, which evolved into more sophisticated intelligent tutoring systems by the 1980s and 1990s (Casheekar et al., 2024). Over time, advances in machine learning and natural language processing enabled the development of virtual learning environments and AI systems that have allowed university educators to improve both the efficiency and effectiveness of their teaching practices (Holmes & Tuomi, 2022). Teachers and educators are central to the application of AIEd, as it they who are responsible for how and to what extent these tools are applied to student learning (Luckin et al., 2024; Luckin, 2016). AIEd can be applied to both student-centred and educator-centred approaches, with overlapping themes emerging such as the broad deployment of generative AI tools to support student learning and knowledge acquisition. In the context of this section, I emphasise teacher-focused AIEd, and largely ignore the third institute-focused AIEd domain that includes the use of AI for administrative tasks.

Many studies highlight generative AI’s potential in higher education (Adıgüzel et al., 2023; Baidoo-Anu & Ansah, 2023; Bettayeb et al., 2024; Sabzalieva & Valentini, 2023). For example, generative AI may be a valuable tool for educators by facilitating the creation, development and systematic organisation of instructional resources. It

may offer opportunities to analyse and conceptualise complex scientific data to support teaching and student engagement or provide virtual case studies and laboratory scenarios that students can use in a safe learning environment (Nikolopoulou, 2024). In addition, LLMs have been demonstrated to assist in the writing and development of quizzes and exam papers on subject-specific topics that may expedite the development of assessment resources for teachers and offer study assistance to students (Gamage et al., 2023; Michel-Villarreal et al., 2023; Xia et al., 2024). Therefore, ChatGPT was asked to provide more details on how LLMs and AI could assist educators in creating or enhancing educational content (see Table 4).

Question: You previously stated that LLMs and AI can assist educators in creating and enhancing educational content. Could you provide more detail, with a particular focus on biology education?

ChatGPT: LLMs and AI can play a significant role in creating and enhancing educational content for biology education. Here are some detailed ways they can assist educators:

Table 4. Examples of Enhancing Educational Content

Criteria	Examples
Generating Lesson Plans and Educational Materials	<p>Lesson Plans: AI can generate detailed lesson plans, including objectives, key concepts, activities, and assessment methods.</p> <p>Teaching Materials: Create slides, handouts, and worksheets with clear explanations and engaging visuals.</p>
Creating Interactive Simulations and Virtual Labs	<p>Simulations: Design simulations that mimic biological processes, such as enzyme activity, genetic inheritance, or ecological interactions.</p> <p>Virtual Labs: Create virtual laboratory experiments where students can conduct experiments, collect data, and analyze results.</p>
Developing Adaptive Learning Modules	LLMs can create adaptive learning modules that adjust content difficulty based on student performance and understanding.
4. Enhancing Visual and Multimedia Content	AI can generate high-quality visual aids and multimedia content to make complex biological concepts more understandable and engaging.
Crafting Assessment Tools and Rubrics	LLMs can assist in creating diverse assessment tools and detailed rubrics to evaluate student learning effectively.
Providing Supplemental Reading and Research Materials	AI can curate and generate supplemental reading lists and research materials to enhance students' understanding of topics.
Facilitating Collaborative Learning	AI tools can create platforms and activities that encourage collaborative learning and group projects.

Criteria	Examples
Supporting Diverse Learning Styles	LLMs can create content that caters to different learning styles, ensuring all students can engage with the material effectively.
Automating Administrative Tasks	AI can automate administrative tasks related to content creation, allowing educators to focus more on teaching and student interaction.
Enhancing Student Engagement	AI can create gamified content and interactive activities that increase student engagement and motivation.

There are several ways that generative AI technologies could assist bioscience educators in designing learning resources, teaching activities, and improving curricula. The generation of lesson plans and the design of modules using LLMs could expedite the process of module and programme development, which often takes considerable time to plan, draft and finalise. New iterations of generative AI applications can now analyse text, images and data, and generate images, figures, graphs and entire presentations. AI could also facilitate the development of experiential learning resources, such as hands-on activities, virtual labs and simulations, which could be particularly valuable for bioscience students, and those studying other STEM subjects. AI system could facilitate this learning approach by allowing students to gain experience and acquire skills that parallel professional activities such as lab-work, fieldwork and data analysis (Whitworth et al., 2018).

Educators face considerable challenges understanding and incorporating AI technologies into teaching practices and curricula, with attitudes ranging from fully embracing AI to banning AI tools altogether (Concannon et al., 2023). Many educators lack formal training in AI fundamentals or the pedagogical implications of AI technologies. This situation is exacerbated by rapid developments in generative AI capabilities and the continual release of new LLMs and now AI agents; a prime example being the recent release of ChatGPT 3o and Deep Research (<https://openai.com/index/introducing-deep-research/>). Continuous professional development is therefore necessary for educators to attain critical technical and AI literacy. This training should not only equip faculty with the fundamentals of AI applications but also allow them to make informed judgements on AI-generated content, its limitations and the ethical implications (Selwyn, 2019). Educator training is essential for evaluating the impact of AI on assessment, the ability to innovate assessments in light of advances in generative AI capabilities, the increasing student adoption of AI applications, and for the integration of these new technologies into flexible and inventive teaching modalities (Xia et al., 2024).

Educators often encounter physical and technical constraints within institutional teaching environments, such as network access, availability of education technology (EdTech) platforms and audiovisual complexities. This establishes a discord between the intentions of educators and the digital infrastructure available, which interferes with the adoption of EdTech. As EdTech platforms and generative AI technologies rapidly evolve, adapting to new interfaces becomes increasingly challenging. Incorporating these recent technologies into pedagogical frameworks and teaching practices also requires continuous professional and skills development. The solution is

likely to require a cultural change within higher education institutions and the redistribution of limited resources, which the higher education sector needs to balance with robust policies and AI frameworks to minimise ethical and data security risks (Batista et al., 2024; Liu et al., 2023).

Enhancing Inclusivity

Inclusivity in higher education centres on creating environments where students, regardless of their background, identity or circumstance, have equitable access to learning opportunities and success (Smart, 2008). Historically, inclusivity relates to ongoing disparities in access to higher education due to economic, social, cultural or institutional barriers and has several parallels with equity in higher education, whereby measures are established to mitigate inequalities, remove barriers and support the success of diverse student populations (Danowitz & Tuitt, 2011; Goering et al., 2022; Jeffries, 2018). Contemporary student cohorts represent a varied cross-section from society, including those from underrepresented minorities, different socio-economic backgrounds, students with disabilities, LGBTQ+ students, first-generation students, adult learners and international students. Promoting a more inclusive and equitable learning environment can also foster creativity, critical thinking and problem solving, and expose students to diverse viewpoints and the complexities of modern societies (Goering et al., 2022; Munna & Kalam, 2021). Therefore, given the importance of inclusivity and equity in modern universities, ChatGPT was asked how generative AI can help make bioscience education more inclusive (see Table 5).

Question: You stated that generative AI can help make bioscience education more inclusive? How could this be achieved? Could you provide some examples in biology education?

ChatGPT: Generative AI can help make bioscience education more inclusive by creating accessible, personalized, and engaging learning experiences for all students, regardless of their backgrounds, learning preferences, or abilities. Here are several ways this can be achieved, with examples in biology education:

Table 5. Example of Generative AI Can Help Make Bioscience Education More Inclusive

Generative AI application	Example description
Personalized Learning Paths	create customized learning experiences tailored to the individual needs and pace of each student. Adaptive content: adjusts the difficulty and style of content based on student's progress and understanding. Individualized feedback: provides guidance to help students improve.
Multimodal Learning Resources	generate content in various formats to cater to different learning styles and preferences. Visual Aids: diagrams, infographics and videos for visual learners. Audio Descriptions: narrations and podcasts for auditory learners. Interactive Activities: hands-on simulations and experiments for kinesthetic learners.
Language Translation and	translate educational materials into multiple languages and simplify

Generative AI application	Example description
Simplification	complex texts to make them accessible to non-native speakers or students with varying levels of language proficiency.
Support for Students with Disabilities	generate content and tools that are accessible to students with various disabilities. Screen Readers: provide text-to-speech capabilities for visually impaired students.
Inclusive Assessment Methods	create diverse assessment formats that accommodate different learning styles and abilities.
Cultural Relevance and Sensitivity	create content that is culturally relevant and sensitive, ensuring that all students feel represented and respected.
Continuous Access and Flexibility	provide continuous access to learning resources, allowing students to learn at their own pace and on their own schedule.
Interactive and Engaging Content	create interactive and gamified content to keep students engaged and motivated.
Real-Time Support and Tutoring	provide real-time support and tutoring, helping students overcome challenges, where AI guides students through difficult topics.
Collaboration and Peer Learning	facilitate collaboration and peer learning, enabling students to learn from and support each other.

Existing learning resources and teaching material could be further adapted to suit diverse student needs and align with particular learning styles, for example by creating customised content for visual, auditory or kinaesthetic learners, thereby promoting student engagement across diverse cohorts (Alshahrani, 2023; Mohebi, 2024). A particular strength of LLMs is their ability to support students with language translation and simplification, particularly for students whose first language is different to the native language of their host institution. Language applications may also be beneficial for students with particular leaning needs who may benefit from text converted to audio (or vice versa), automatic transcription or captioning of video. LLMs could also support cultural and ethnic diversity, an area where some biology and STEM curricula have historically been deficient (Fernandes, 2021).

Biology is traditionally taught through the lens of a Western and male perspective (Arif et al., 2021) and decolonising bioscience curricula could take inspiration from recent medical curricular reform (Wong et al., 2021). Despite concerns regarding training and algorithmic bias when applying generative AI models in this context (Kasneci et al., 2023), AI has the potential to assist with curricula reform by helping education providers

encompass more diverse aspects of society. For example, generative AI could be prompted to provide biological explanations that incorporate examples from diverse ethnic backgrounds. A contemporary parallel with bioscience research is the recent publication of the panhuman genome project, which aims to more broadly map the genetic diversity of the global human population (Wang et al., 2022). LLMs could therefore assist with incorporating the panoply of human diversity into contemporary bioscience education.

However, concerns have been raised regarding the potential risk of exacerbating inequalities and expanding the digital divide. Despite the perceived democratisation of knowledge acquisition that generative AI provides, those students from certain underrepresented groups have historically suffered from technological deprivation, for example due to lack of internet access, digital skills and language barriers (Raihan et al., 2024). Care should be taken, therefore, to ensure that access to AI technologies is equitable within higher education institutions and that all students are provided with the necessary AI literacy skills required to utilise these tools proficiently. Promoting equitable pedagogies and institutional policies to foster a more inclusive higher education environment will be an ongoing challenge for universities considering the rapid evolution of AI applications (Roshanaei et al., 2023).

Curriculum Design

Due to diverse learner profiles and knowledge levels, teaching in the higher education sector is both challenging and rewarding. Effective university curricula must therefore incorporate robust subject-specific learning material, while maintaining standards for teaching, assessment and quality assurance (Hoecht, 2006). Several proponents of curriculum reform have highlighted the importance of inclusivity and acknowledging the diversity of students with different abilities, ages and educational backgrounds (Chan et al., 2017). The influence of professional standards authorities, such as the Quality Assurance Agency in the UK (QAA, 2024b), offer further support to ensure inclusivity in curriculum design, aligning with both institutional policies and subject benchmark statements, with the overall aim of offering diverse learning opportunities and ensuring coverage of domain- or discipline-specific educational content (Munna & Kalam, 2021). Therefore, ChatGPT was asked how generative AI could assist in aligning undergraduate bioscience curricula with QAA benchmark statements (see Table 6).

Question: I'm also interested in improving the bioscience curriculum at university. In the UK, higher education institutions use the QAA benchmark statement in biosciences as an informative guide. How could generative AI be used to align an undergraduate BSc Bioscience curriculum with the QAA benchmark statement in biosciences?

ChatGPT: Aligning an undergraduate BSc Bioscience curriculum with the QAA (Quality Assurance Agency) benchmark statement in biosciences can be greatly facilitated by the use of generative AI. Here's how generative AI can be utilized to achieve this alignment effectively:

Table 6. Aligning Curriculum with Subject Benchmark Statements

Alignment strategy	Description
Curriculum Mapping and Analysis	Analyze current curriculum and map it against the QAA benchmark statements to identify gaps. Use AI to review existing course content, learning outcomes,

Alignment strategy	Description
	and assessment methods. Align with QAA benchmark statements to ensure comprehensive coverage of required competencies and skills.
Generating Course Content and Resources	Generate and update course content to meet QAA benchmarks, ensuring that all topics and competencies are covered. Generate lecture notes, reading lists, and multimedia resources. Suggest up-to-date academic articles, textbooks, and online resources that align with the benchmarks.
Developing Assessment Methods	Design assessments that align with the QAA benchmarks, ensuring that students are evaluated on all required competencies. Create quizzes, exams, and project assignments that reflect the skills and knowledge outlined in the benchmarks. Develop detailed rubrics that ensure consistent and objective grading.
Personalized Learning Paths	Create personalized learning paths for students to ensure they meet all the benchmark standards by the end of their degree.
Enhancing Pedagogical Approaches	Suggest innovative teaching methods and pedagogical strategies to meet the QAA benchmarks effectively.
Continuous Feedback and Improvement	Facilitate continuous feedback from students and educators to ensure the curriculum remains aligned with QAA benchmarks.
Facilitating Accreditation and Quality Assurance	Assist in preparing for accreditation reviews and ensuring ongoing compliance with QAA benchmarks.
Professional Development for Educators	Provide resources and training for educators to help them align their teaching with the QAA benchmarks.

Aligning new or existing undergraduate programmes with the QAA Subject Benchmark Statement in Biosciences ensures consistent higher education delivery across the sector (Watson et al., 2023). Biosciences is a broad discipline that encompasses various pathways with a common objective of understanding life on Earth. Several disciplines and courses are therefore positioned within the bioscience framework, including biology, biochemistry and physiology, with a collective emphasis placed on developing a range of student skills such as analytical and critical thinking, practical skills, communication, research and collaboration. LLMs could therefore assist educators and course developers to align their programmes with these core student skills and ensure relevant subject areas are included within the curriculum, for example including the theory of evolution; whole organism physiology and anatomy; species interactions; behaviour, population biology and dynamics; ecology and ecosystems biology; biotechnology; and cellular and molecular biology (Watson et al., 2023).

Many contemporary bioscience programmes offer students individualised pathways that allow specialisation. LLMs could offer an efficient way for educators to ensure those pathways remain aligned with the benchmark statement standards within their sub-discipline and provide a tool for continued quality assurance and compliance with QAA, sector and institutional standards of practice. This may also include some insights on training educators and providing suggestions for continued professional development (Jones & Newton, 2024; Lee et al., 2024). As AI technology continues to advance, educator training on recent developments will be essential for appropriate integration in education strategies.

Generic competencies are another critical component of higher education curricula and include critical thinking, problem solving, analytical skills, communication, academic literacy, digital literacy, collaboration, teamwork and life-long learning (Clanchy & Ballard, 1995). These competencies go beyond domain-specific knowledge and equip students with transferable skills applicable across disciplines (Chan et al., 2017). Despite the growing importance of generic competencies, there is a lack of understanding among academics regarding their development and inconsistencies in their deployment within higher education providers (Hyytinen et al., 2019; Tuononen et al., 2022). Generative AI may offer educators a viable solution to align their curricula with these generic competencies and highlight areas of the curriculum where these elements of skills development could be employed (Michel-Villarreal et al., 2023). AI tools may also offer opportunities to align curricula and teaching practices with institutional policies and education frameworks (Wang et al., 2024). However, educator acceptance and poor AI literacy remain a challenge, while curriculum developers must be aware of data privacy and ethical risks when deploying AI tools. The adoption of generative AI and LLMs for educational resource creation should also be approached with caution, as human oversight and agency should remain paramount, ensuring that AI technology complements rather than replaces the human elements of teaching (Chan & Tsi, 2024; Paraschiv, 2024).

Assessment design and delivery is considered a fundamental aspect of programme design. Assessment and feedback is also known to be a key driver of student learning and engagement in higher education, particularly if placed into the context of the principles of ‘assessment for learning’ (Carless, 2005) and ‘constructive alignment’ (Biggs, 1996). Generative AI tools could provide support for educators to align their assessment practices with these principles, to generate ideas for assessment formats and to generate questions that align with programme learning outcomes. Much debate in the literature has focussed on students using generative AI tools, such as ChatGPT, to produce assignment outputs and commit plagiarism or breaches in academic integrity (Benuyenah, 2023; Cotton et al., 2024). Concerns regarding academic integrity and plagiarism are certainly justified given that ChatGPT, and other LLMs, have demonstrated high levels of competencies on several university-level assessment formats (Gilson et al., 2022; Newton & Xiomeriti, 2024; A. Williams, 2024). The contrasting view is that generative AI offers significant potential to diversify assessments, support student learning and promote assessments that evaluate higher order cognitive skills, such as critical thinking, collaboration and problem-solving (Benuyenah, 2023; Gamage et al., 2023). Therefore, I will not comment further on the topic of AI and academic integrity here.

Ethical Challenges and Considerations

The ethics of higher education encompass a broad set of principles that are essential for an inclusive, impartial and intellectually honest academic environment that transcend the basic idea of university as repository of knowledge (Newman et al., 1996; Richardson & Healy, 2019). The commitment to fairness and equity in student admissions, access to resources and the experience of diverse learners are critical for contemporary universities to uphold high standards of ethical practice. Furthermore, transparency and accountability in institutional decision-making processes are key to ensure all students are treated justly. As higher education institutions integrate more digital technology, including AI, into teaching, learning and administrative processes, safeguarding personal information and ensuring data protection is becoming increasingly challenging and an area that necessitates responsible implementation (Roberts et al., 2017; Selwyn, 2019). Another central tenet of higher education is the interaction between teacher and student (Brown & Krager, 1985). However, some consider the emergence of generative AI as a disruptive influence that may harm the student-teacher relationship, while others recognise the benefits of adopting AI technology to complement teacher-led tuition (Chan & Tsi, 2024; Farrelly & Baker, 2023). Therefore, ChatGPT was asked to comment on the ethical challenges and considerations of using generative AI in the higher education context (see Table 7).

Question: You also mention ethical challenges and considerations. Can you explain more about the ethical concerns and bias when using generative AI? Could you explain this in detail, please, and provide some examples relevant to biology education?

ChatGPT: The use of generative AI in education, including biology education, brings significant potential but also comes with several ethical concerns and biases that need careful consideration. Here's a detailed look at these issues:

Table 7. Ethical and Bias Concerns When Using Generative AI in Education

Ethical and Bias Concerns	Description
Academic Integrity	Students may misuse AI tools to generate content that they present as their own work, compromising academic integrity.
Data Privacy and Security	AI systems require large amounts of data to function effectively, raising concerns about the privacy and security of student data.
Informed Consent	Students and educators must be fully aware of how their data is being used by AI systems.
Transparency	The decision-making processes of AI systems should be transparent and understandable to users.
Training Data Bias	AI systems learn from data, and if the training data contains biases, the AI's outputs will reflect those biases. Example: an AI system might not

Ethical and Bias Concerns	Description
	adequately represent biological diversity or educational practices, leading to a biased perspective.
Algorithmic Bias	Even with unbiased data, the algorithms themselves can introduce biases through the ways they process and prioritize information. Example: an AI tool might disproportionately flag students from certain demographic backgrounds if the algorithm is not carefully designed and tested.
Representation Bias	AI systems might not adequately represent all groups, leading to a lack of inclusivity in educational content and resources. Example: A generative AI tool might inadvertently focus primarily on Western scientific discoveries and neglect important contributions from scientists in other parts of the world.

Academic integrity, with regard to students misusing generative AI to produce written assignments, is a major concern across higher education (Sullivan et al., 2023). Some institutions have taken the approach of employing AI detection software to mitigate for this misuse. However, AI detection software is unreliable and produces false positives (Chen et al., 2020). Moreover, entering an arms race between generative AI and AI detection software is likely to be a futile exercise, given the rapid advancement of new iterations of generative AI models. A return to in-person, invigilated assessments and a reduced reliance on online exams, will certainly improve assessment resistance to AI. Various forms of assessment are also considered more refractory to AI misuse, such as *viva voce* examinations, practical and clinical assessments and interactive presentations. These are already evident in many bioscience programmes and have a robust history in bioscience education. However, a more pragmatic and progressive approach may be needed, which favours the integration of AI into existing and new assessment formats (Sabzalieva & Valentini, 2023; Williams, 2025). AI could be incorporated into assessment formats, for example as an ‘AI solution-finder’, ‘AI-generated research proposal’, ‘AI-created learning resource’ or using AI to generate drafts of written assignments (Sharples, 2023). These AI-friendly approaches are additionally beneficial for promoting generative AI and technological literacy among students and fostering critical thinking and creativity.

Caution should be exercised, however, when handling student work and personal data, as several generative AI platforms, such as ChatGPT, use imputed text to further train their models. Data security, use of personal information and ownership of material are important considerations. This may be particularly relevant if educators wish to use generative AI applications to provide feedback on submitted student work. Using protected AI tools or open-source platforms may be necessary to ensure confidentiality and data protection, while mitigating any cybersecurity risks (Chan, 2023).

Despite the possibilities of modifying existing assessment formats or further integrating generative AI into assessment design, concerns persist regarding algorithmic bias and hallucinations (Francis et al., 2025). Inaccuracies or misinformation generated by AI, whether deliberate or unintentional, not only undermines academic rigour, but it also poses serious questions regarding systemic bias. Although several strategies have been proposed to mitigate for algorithmic bias, including technical methods for detecting and mitigating bias, challenges remain for higher education providers, and for students, when navigating the ethical complexities of generative AI output (Barnes & Hutson, 2024; Liao et al., 2023). As unintentional biases may exist in the AI training data, continued vigilance and substantiation of facts are needed. Issues around ownership of AI generated content, intellectual property rights and copyright infringement also persist, which may be compounded by the lack of GDPR and other internationally recognised regulatory agreements placed on AI systems (Berendt et al., 2020). There are also concerns about transparency and accountability regarding AI generated content, with the responsibility of ensuring accuracy falling to the user. The 'black-box problem', which refers to the lack of transparency in how AI models make decisions, further contributes to the lack of accountability, which may require the coordinated training of students on the both the ethical principles and capabilities of generative AI (Bearman & Ajjawi, 2023).

Discussion

Generative AI is considered to be part of the fourth industrial revolution, or industry 4.0, along with advanced robotics, big data analytics and cyber-physical systems (Giannakos et al., 2024; Mian et al., 2020). While industry 4.0 focussed on automation and efficiency, industry 5.0 is now prioritising human collaboration with advanced technology to foster human creativity and productivity (Andres et al., 2022; Hashim et al., 2024). AI in education is therefore a critical component of this revolution. This article examines how generative AI tools can enhance teaching and learning, discussing both the opportunities and hurdles in integrating such technologies into educational environments. Considering the potential impact that LLMs may have on transforming teaching and learning modalities (Michel-Villarreal et al., 2023; Wu, 2023), and the way in which students are adopting AI in their studies and assessments (Arowosegbe et al., 2024; Chan & Hu, 2023), it is important to explore the capabilities and limitations of this technology in higher education settings. This paper presents a conversation with ChatGPT 4o and a critical evaluation of the role of generative AI and LLMs in bioscience higher education.

In particular, chatbots such as ChatGPT offer a user-friendly interface that enables a convincing, human-like dialogue with the user. The iterative and experiential quality of AI chatbots, allows users to access information rapidly and request iterative modifications of generated output. Furthermore, LLM algorithms that leverage natural language processing learn and modify in response to user feedback. In concert with the massive training data set, which is even larger with more recent iterations of GPT models and competitor LLMs, means that a huge corpus of knowledge and human-generated information is available for exploration. Generative AI therefore holds significant promise for exploring ideas for transformative pedagogies in the biosciences (Tamer et al., 2024).

The educational concepts explored in conversation with ChatGPT included using LLMs to support personalized student learning experiences, tailoring educational content, providing personalized feedback and assisting in

learning development. Furthermore, the iterative nature of feedback loops facilitated by LLMs cultivates a process of continuous improvement, empowering students to refine their analytical skills and to develop their understanding of subject-specific concepts (Hounsell, 2003). LLMs may hold promise in supporting problem-solving skills and critical thinking, through simulated scenarios, role-playing, Socratic dialogue and case studies, which represent valuable approaches for student learning in the biosciences and other STEM subjects. The potential of AI to act as a personal tutoring system has been well documented (Adigüzel et al., 2023; Baidoo-Anu & Ansah, 2023), although careful prompt engineering and responsible use are important considerations for both teachers and students, not least to avoid inaccuracies and mitigate for the risks of hallucinations (Kasneci et al., 2023; Rudolph et al., 2023).

However, recent concerns about ChatGPT losing quality and performance over time have emerged, particularly in response to ‘chain-of-thought’ prompting (Chen et al., 2024). This may cause issues with reproducibility and reliability, two important qualities to consider in bioscience education and when using LLMs as a learning support interface. Another concerning feature to consider is the generation of hallucinations, which are a persistent problem with generative AI tools and therefore guidance and training for students and educators is recommended, particularly with regard to scientific writing and as a reliable knowledge base to support student learning (Alkaissi & McFarlane, 2023; Arnold, 2023). Algorithmic bias is important ethical consideration, given the potential of AI systems to absorb and reiterate societal biases (Bolukbasi et al., 2016; Lund & Wang, 2023). In an educational context, this could undermine the value of LLMs as a reliable resource, with responsibility for ensuring partiality assigned to educators and students (R. T. Williams, 2024). The lack of transparency inherent in AI systems, and the associated black box problem, is another ethical consideration that warrants further exploration with regard to fairness and accountability (Garcia Ramos & Wilson-Kennedy, 2024). Issues with bias and digital equity may even go beyond those related to underrepresented student groups to include academic researchers, as highlighted by a recent study suggesting that generative AI may accentuate the widening of the productivity gap between genders (Tang et al., 2025). Although generative AI offers opportunities for the global democratisation of education, this needs to be balanced with responsible ethical frameworks (Yu, 2024; Zhai, 2022). To integrate AI into higher education in a positive and beneficial way may require a deeper critical appraisal of the algorithmic bias challenges, and further education-focussed research in this area would be valuable.

The limitations and ethical considerations of LLMs include challenges related to biased, outdated and incomplete information, complexity of scientific language, lack of contextual understanding, inability to verify information, limited understanding of generated output, and digital equity considerations. However, it was notable that ChatGPT generated balanced responses when prompted on its limitations. The integration of generative AI has been discussed by Sallam *et al.* (2023) in the context of healthcare and allied medical education, with plagiarism, lack of personal interaction, generation of inaccurate content and negative impact on critical thinking development noted as shortcomings (Sallam et al., 2023). Concerns have previously been raised regarding ChatGPT positioning itself as an epistemic authority without appropriate discipline-specific knowledge and expertise (Cooper, 2023). This further highlights the need for appropriate staff and student training in ChatGPT use, and generative AI more broadly, particularly in the context of biosciences education. Overcoming techno-panic and the resistance to adopt emerging technologies among educators would also be important for the integration of AI tools into education

practices (Baidoo-Anu & Ansah, 2023; Garcia Ramos & Wilson-Kennedy, 2024). Further research would be warranted on how best to develop educational frameworks and professional development programs to train bioscience educators on the responsible integration of generative AI technologies.

Central to the conversation surrounding LLMs was the exploration of bioscience curricula development, teaching, assessment and feedback strategies, learning resources that align with biological subject content and quality assurance. By leveraging LLMs, educators can deploy innovative assessment formats, adaptive learning pathways and strategies for personalised feedback and learning support (Baidoo-Anu & Ansah, 2023). There may be further opportunity to explore the development and validation of assessment tools leveraging LLM capabilities for bioscience education and further research could focus on the design and implementation of assessments that incorporate generative AI into their design, particularly in the context of experiential or authentic assessments that are relevant for professional bioscience practices (Salinas-Navarro et al., 2024). This research could also be extended to the impact of generative AI on critical thinking, which is a frequently cited aspect of student learning affected by AI system use (Larson et al., 2024; Ruiz-Rojas et al., 2024).

Conclusion

There is a valuable opportunity to explore the varied potential of LLMs to support numerous aspects of bioscience higher education and related allied healthcare and STEM disciplines. No doubt educators and the higher education sector face significant challenges when trying to adapt to AI, including ethical considerations surrounding the use of LLMs in educational settings, which may be particularly prescient for medical and bioscience education given the humanistic qualities of these subjects. Future research could investigate the ethical implications of LLM-generated content regarding learning resources and assessment practices, including issues related to bias, reliability, privacy and academic integrity (Moorhouse et al., 2023). The impact of LLMs on equitable access and equality among students in higher education could also be explored, including examining the effectiveness of digital literacy training and equitable distribution of digital resources in addressing disparities and promoting inclusive learning environments. Critically, universities have a responsibility to train their students on the technical, ethical and acceptable uses of generative AI and to equip students to become effective members of an AI-driven society (Aoun, 2024; Ocaña-Fernández et al., 2019). In conclusion, the conversation with ChatGPT was intended to stimulate further dialogue among educators and student partners, and to explore the utility of generative AI in the biosciences.

References

- Abas, M. A., Arumugam, S. E., Yunus, M. M., & Rafiq, K. R. M. (2023). ChatGPT and personalized learning: Opportunities and challenges in higher education. *International Journal of Academic Research in Business and Social Sciences*, 13(12). <https://doi.org/DOI:10.6007/IJARBSS/v13-i12/20240>
- Adıgüzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), 429. <https://doi.org/10.30935/cedtech/13152>

- Ahmad, Z., Kaiser, W., & Rahim, S. (2023). Hallucinations in ChatGPT: an unreliable tool for learning. *Rupkatha Journal on Interdisciplinary Studies in Humanities*, 15(4), 1-18. <https://doi.org/10.21659/rupkatha.v15n4.17>
- Akpan, I. J., Kobara, Y. M., Owolabi, J., Akpan, A. A., & Offodile, O. F. (2025). Conversational and generative artificial intelligence and human–chatbot interaction in education and research. *International Transactions in Operational Research*, 32(3), 1251-1281. <https://doi.org/10.1111/itor.13522>
- Alasadi, E. A., & Baiz, C. R. (2023). Generative AI in education and research: Opportunities, concerns, and solutions. *Journal of Chemical Education*, 100(8), 2965-2971. <https://doi.org/10.1021/acs.jchemed.3c00323>
- Ali, R., Tang, O., Connolly, I., Fridley, J., Shin, J., Zadnik Sullivan, P., Cielo, D., Oyelese, A., Doberstein, C., & Telfeian, A. (2023). Performance of ChatGPT, GPT-4, and Google Bard on a neurosurgery oral boards preparation question bank. *Neurosurgery Publish Ahead of Print: 2023–04. Neurosurgery*, 93(5), 1090-1098. <https://doi.org/10.1227/neu.0000000000002551>
- Alkaissi, H., & McFarlane, S. I. (2023). Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*, 15(2). <https://doi.org/10.7759/cureus.35179>
- Alshahrani, A. (2023). The impact of ChatGPT on blended learning: Current trends and future research directions. *International Journal of Data and Network Science*, 7(4), 2029-2040. <https://doi.org/DOI:10.5267/j.ijdns.2023.6.010>
- Andres, B., Sempere-Ripoll, F., Estes, A., & Alemany, M. (2022). Mapping between Industry 5.0 and Education 5.0. *EDULEARN22 Proceedings*, 1, 2921–2926. In.
- Aoun, J. E. (2024). How Higher Ed Can Adapt to the Challenges of AI. https://www.chronicle.com/article/how-higher-ed-can-adapt-to-the-challenges-of-ai?utm_source=Iterable&utm_medium=email&utm_campaign=campaign_10408935_nl_Academe-Today_date_20240710
- Arif, S., Massey, M. D. B., Klinard, N., Charbonneau, J., Jabre, L., Martins, A. B., Gaitor, D., Kirton, R., Albury, C., & Nanglu, K. (2021). Ten simple rules for supporting historically underrepresented students in science. In (Vol. 17, pp. e1009313): Public Library of Science San Francisco, CA USA.
- Arnold, L. (2023). *Striking a Balance: Integrating AI into Assessment Practices*. <https://lydiaarnold.net/2023/06/12/striking-a-balance-integrating-ai-into-assessment-practices/>
- Arowosegbe, A., Alqahtani, J. S., & Oyelade, T. (2024). Students' Perception of Generative AI Use for Academic Purpose in UK Higher Education. *Preprints*, 289-295. <https://doi.org/doi:10.3389/feduc.2024.1463208>
- Ayers, J. W., Poliak, A., Dredze, M., Leas, E. C., Zhu, Z., Kelley, J. B., Faix, D. J., Goodman, A. M., Longhurst, C. A., & Hogarth, M. (2023). Comparing physician and artificial intelligence chatbot responses to patient questions posted to a public social media forum. *JAMA internal medicine*, 183(6), 589-596.
- Bagenal, J. (2024). Generative artificial intelligence and scientific publishing: urgent questions, difficult answers. *The Lancet*, 403(10432), 1118-1120. [https://doi.org/10.1016/S0140-6736\(24\)00416-1](https://doi.org/10.1016/S0140-6736(24)00416-1)
- Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Barnes, E., & Hutson, J. (2024). Navigating the ethical terrain of AI in higher education: Strategies for mitigating

- bias and promoting fairness. *Forum for Education Studies*, 2(2), 1229. <https://doi.org/10.59400/fes.v2i2.1229>
- Batista, J., Mesquita, A., & Carnaz, G. (2024). Generative AI and higher education: Trends, challenges, and future directions from a systematic literature review. *Information*, 15(11), 676. <https://doi.org/10.3390/info15110676>
- Bearman, M., & Ajjawi, R. (2023). Learning to work with the black box: Pedagogy for a world with artificial intelligence. *British Journal of Educational Technology*, 54(5), 1160-1173. <https://doi.org/10.1111/bjet.13337>
- Bengio, Y., Ducharme, R., Vincent, P., & Janvin, C. (2003). A Neural Probabilistic Language Model. *Journal of machine learning research*,
- Benuyenah, V. (2023). Commentary: ChatGPT use in higher education assessment: Prospects and epistemic threats. *Journal of Research in Innovative Teaching & Learning*, 16(1), 134-135. <https://doi.org/10.1108/JRIT-03-2023-097>
- Berendt, B., Littlejohn, A., & Blakemore, M. (2020). AI in education: Learner choice and fundamental rights. *Learning, Media and Technology*, 45(3), 312-324. <https://doi.org/10.1080/17439884.2020.1786399>
- Bettayeb, A. M., Abu Talib, M., Sobhe Altayasinah, A. Z., & Dakalbab, F. (2024). Exploring the impact of ChatGPT: conversational AI in education. *Frontiers in Education*,
- Bharatha, A., Ojeh, N., Fazle Rabbi, A. M., Campbell, M. H., Krishnamurthy, K., Layne-Yarde, R. N., Kumar, A., Springer, D. C., Connell, K. L., & Majumder, M. A. A. (2024). Comparing the Performance of ChatGPT-4 and Medical Students on MCQs at Varied Levels of Bloom's Taxonomy. *Advances in Medical Education and Practice*, 393-400. <https://doi.org/10.2147/AMEP.S457408>
- Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher education*, 32(3), 347-364. <https://doi.org/10.1007/BF00138871>
- Blau, W., Cerf, V. G., Enriquez, J., Francisco, J. S., Gasser, U., Gray, M. L., Greaves, M., Grosz, B. J., Jamieson, K. H., & Haug, G. H. (2024). Protecting scientific integrity in an age of generative AI. *Proceedings of the National Academy of Sciences*, 121(22), e2407886121. <https://doi.org/10.1073/pnas.2407886121>
- Bobula, M. (2024). Generative Artificial Intelligence (AI) in higher education: a comprehensive review of challenges, opportunities, and implications. *Journal of Learning Development in Higher Education*(30). <https://doi.org/10.47408/jldhe.vi30.1137>
- Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., & Kalai, A. T. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in neural information processing systems*, 29.
- Borger, J. G., Ng, A. P., Anderton, H., Ashdown, G. W., Auld, M., Blewitt, M. E., Brown, D. V., Call, M. J., Collins, P., & Freytag, S. (2023). Artificial intelligence takes center stage: exploring the capabilities and implications of ChatGPT and other AI-assisted technologies in scientific research and education. *Immunology and cell biology*, 101(10), 923-935. <https://doi.org/10.1111/imcb.12689>
- Brown, R. D., & Krager, L. (1985). Ethical issues in graduate education: Faculty and student responsibilities. *The Journal of Higher Education*, 56(4), 403-418. <https://doi.org/10.1080/00221546.1985.11780701>
- Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: Principles, Policy & Practice*, 12(1), 39-54. <https://doi.org/10.1080/0969594042000333904>

- Casheekar, A., Lahiri, A., Rath, K., Prabhakar, K. S., & Srinivasan, K. (2024). A contemporary review on chatbots, AI-powered virtual conversational agents, ChatGPT: Applications, open challenges and future research directions. *Computer Science Review*, 52, 100632. <https://doi.org/10.1016/j.cosrev.2024.100632>
- Chan, C. K., Fong, E. T., Luk, L. Y., & Ho, R. (2017). A review of literature on challenges in the development and implementation of generic competencies in higher education curriculum. *International Journal of Educational Development*, 57, 1-10. <https://doi.org/10.1016/j.ijedudev.2017.08.010>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan, C. K. Y., & Colloton, T. (2024). *Generative AI in Higher Education: The ChatGPT Effect*. Taylor & Francis. <https://doi.org/10.4324/9781003459026>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1), 43.
- Chan, C. K. Y., & Tsi, L. H. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, 83, 101395. <https://doi.org/10.1016/j.stueduc.2024.101395>
- Chang, D. H., Lin, M. P.-C., Hajian, S., & Wang, Q. Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability*, 15(17), 12921. <https://doi.org/10.3390/su151712921>
- Chen, L., Zaharia, M., & Zou, J. (2024). How is ChatGPT's behavior changing over time? *Harvard Data Science Review*, 6(2). <https://doi.org/10.1162/99608f92.5317da47>
- Chen, X., Xie, H., & Hwang, G.-J. (2020). A multi-perspective study on artificial intelligence in education: Grants, conferences, journals, software tools, institutions, and researchers. *Computers and Education: Artificial Intelligence*, 1, 100005. <https://doi.org/10.48550/arXiv.2307.09009>
- Chien, C.-C., Chan, H.-Y., & Hou, H.-T. (2024). Learning by playing with generative AI: design and evaluation of a role-playing educational game with generative AI as scaffolding for instant feedback interaction. *Journal of Research on Technology in Education*, 1-20. <https://doi.org/10.1080/15391523.2024.2338085>
- Clanchy, J., & Ballard, B. (1995). Generic skills in the context of higher education. *Higher Education Research and Development*, 14(2), 155-166. <https://doi.org/10.1080/0729436950140202>
- Concannon, F., Costello, E., Farrell, O., Farrelly, T., & Wolf, L. G. (2023). There's an AI for that: Rhetoric, reality, and reflections on EdTech in the dawn of GenAI. *Irish Journal of Technology Enhanced Learning*, 7(1). <https://doi.org/10.22554/ijtel.v7i1.116>
- Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32(3), 444-452. <https://doi.org/10.1007/s10956-023-10039-y>
- Cotton, D. R., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in education and teaching international*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Danowitz, M. A., & Tuitt, F. (2011). Enacting inclusivity through engaged pedagogy: A higher education perspective. *Equity & Excellence in Education*, 44(1), 40-56.

- <https://doi.org/10.1080/10665684.2011.539474>
- Dempere, J., Modugu, K., Hesham, A., & Ramasamy, L. K. (2023). The impact of ChatGPT on higher education. *Frontiers in Education*, 8, 1206936. <https://doi.org/10.3389/feduc.2023.1206936>
- Deng, X., & Yu, Z. (2023). A meta-analysis and systematic review of the effect of chatbot technology use in sustainable education. *Sustainability*, 15(4), 2940. <https://doi.org/10.3390/su15042940>
- Elliott, D., & Soifer, E. (2022). AI technologies, privacy, and security. *Frontiers in Artificial Intelligence*, 5, 826737. <https://doi.org/10.3389/frai.2022.826737>
- Eysenbach, G. (2023). The role of ChatGPT, generative language models, and artificial intelligence in medical education: a conversation with ChatGPT and a call for papers. *JMIR Medical Education*, 9(1), e46885. <https://doi.org/10.2196/46885>
- Farrelly, T., & Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>
- Fernandes, M. (2021). What does decolonising the curriculum mean for STEM subjects? *Compass: Journal of Learning and Teaching in Higher Education*, 14(2), 1-4. <https://doi.org/http://dx.doi.org/10.21100/compass.v14i2.1230>
- Francis, N. J., Jones, S., & Smith, D. P. (2025). Generative AI in higher education: balancing innovation and integrity. *British Journal of Biomedical Science*, 81, 14048. <https://doi.org/10.3389/bjbs.2024.14048>
- Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: is Chat GPT a blessing or a curse? *Frontiers in Education*, 8, 1166682. <https://doi.org/10.3389/feduc.2023.1166682>
- Gamage, K. A., Dehideniya, S. C., Xu, Z., & Tang, X. (2023). ChatGPT and higher education assessments: more opportunities than concerns? *Journal of applied learning and teaching*, 6(2). <https://doi.org/doi:10.37074/jalt.2023.6.2.32>
- Ganesan, S., & Somasiri, N. (2024). Navigating the Integration of Machine Learning in Healthcare: Challenges, Strategies, and Ethical Considerations. *Journal of Computational and Cognitive Engineering*, 4(1), 8-23. <https://doi.org/10.47852/bonviewJCCE42023600>
- Garcia Ramos, J., & Wilson-Kennedy, Z. (2024). Promoting equity and addressing concerns in teaching and learning with artificial intelligence. *Frontiers in Education*,
- Giannakos, M., Azevedo, R., Brusilovsky, P., Cukurova, M., Dimitriadis, Y., Hernandez-Leo, D., Järvelä, S., Mavrikis, M., & Rienties, B. (2024). The promise and challenges of generative AI in education. *Behaviour & Information Technology*, 1-27. <https://doi.org/10.1080/0144929X.2024.2394886>
- Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2022). How does ChatGPT perform on the medical licensing exams? The implications of large language models for medical education and knowledge assessment. *MedRxiv*, 2022.2012. 2023.22283901. <https://doi.org/10.1101/2022.12.23.22283901>
- Goering, A. E., Resnick, C. E., Bradford, K. D., & Othus-Gault, S. M. (2022). Diversity by design: Broadening participation through inclusive teaching. *New Directions for Community Colleges*, 2022(199), 77-91. <https://doi.org/10.1002/cc.20525>
- Gregorcic, B., & Pendrill, A.-M. (2023). ChatGPT and the frustrated Socrates. *Physics Education*, 58(3), 035021. <https://doi.org/10.1088/1361-6552/acc299>
- Guo, H., Yi, W., & Liu, K. (2024). Enhancing Constructivist Learning: The Role of Generative AI in Personalised

- Learning Experiences. *Proceedings of the 26th International Conference on Enterprise Information Systems (ICEIS 2024) 1*, 767-770. <https://doi.org/10.5220/0012688700003690>
- Hashim, M. A. M., Tlemsani, I., Mason-Jones, R., Matthews, R., & Ndrecaj, V. (2024). Higher education via the lens of industry 5.0: Strategy and perspective. *Social Sciences & Humanities Open*, 9, 100828. <https://doi.org/10.1016/j.ssaho.2024.100828>
- Hauswald, R. (2025). Artificial Epistemic Authorities. *Social Epistemology*, 1-10. <https://doi.org/10.1080/02691728.2025.2449602>
- Ho, Y.-R., Chen, B.-Y., & Li, C.-M. (2023). Thinking more wisely: using the Socratic method to develop critical thinking skills amongst healthcare students. *BMC medical education*, 23(1), 173. <https://doi.org/10.1186/s12909-023-04134-2>
- Hoecht, A. (2006). Quality assurance in UK higher education: Issues of trust, control, professional autonomy and accountability. *Higher education*, 51, 541-563. <https://doi.org/10.1007/s10734-004-2533-2>
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542-570. <https://doi.org/10.1111/ejed.12533>
- Hounsell, D. (2003). Student feedback, learning and development. In *Higher education and the lifecourse* (pp. 67-78). SRHE & Open University Press.
- Huh, S. (2023). Can We Trust AI Chatbots' Answers about Disease Diagnosis and Patient Care. *J. Korean Med. Assoc*, 66, 218-222. <https://doi.org/10.5124/jkma.2023.66.4.218>
- Hyytinen, H., Toom, A., & Shavelson, R. J. (2019). Enhancing scientific thinking through the development of critical thinking in higher education. *Redefining scientific thinking for higher education: Higher-order thinking, evidence-based reasoning and research skills*, 59-78. https://doi.org/10.1007/978-3-030-24215-2_3
- Jeffries, R. (2018). *Diversity, equity, and inclusivity in contemporary higher education*. IGI Global.
- Jin, S.-H., Im, K., Yoo, M., Roll, I., & Seo, K. (2023). Supporting students' self-regulated learning in online learning using artificial intelligence applications. *International Journal of Educational Technology in Higher Education*, 20(1), 37. <https://doi.org/10.1186/s41239-023-00406-5>
- Jones, S., & Newton, P. M. (2024). Education and Training Assessment and Artificial Intelligence. A Pragmatic Guide for Educators. *British Journal of Biomedical Science*, 81, 14049. <https://doi.org/10.3389/bjbs.2024.14049>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., & Hüllermeier, E. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kayabas, I. (2024). Innovative Curriculum Development and Content Creation With Generative AI. In *Transforming Education With Generative AI: Prompt Engineering and Synthetic Content Creation* (pp. 462-473). IGI Global. <https://doi.org/DOI: 10.4018/979-8-3693-1351-0.ch022>
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., Madriaga, M., Aggabao, R., Diaz-Candido, G., & Maningo, J. (2023). Performance of ChatGPT on USMLE: potential for AI-assisted medical education using large language models. *PLoS digital health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>

- Labov, J. B., Reid, A. H., & Yamamoto, K. R. (2010). Integrated biology and undergraduate science education: a new biology education for the twenty-first century? *CBE—Life Sciences Education*, 9(1), 10-16. <https://doi.org/10.1187/cbe.09-12-0092>
- Lai, J. W. (2024). Adapting Self-Regulated Learning in an Age of Generative Artificial Intelligence Chatbots. *Future Internet*, 16(6), 218. <https://doi.org/10.3390/fi16060218>
- Larson, B. Z., Moser, C., Caza, A., Muehlfeld, K., & Colombo, L. A. (2024). Critical thinking in the age of generative AI. *Academy of Management Learning & Education*, 23(3), 373-378. <https://doi.org/10.5465/amle.2024.0338>
- Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., Lekkas, D., & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, 6, 100221. <https://doi.org/10.1016/j.caeai.2024.100221>
- Li, K. C., & Wong, B. T.-M. (2023). Artificial intelligence in personalised learning: A bibliometric analysis. *Interactive Technology and Smart Education*, 20(3), 422-445. <https://doi.org/10.1108/ITSE-01-2023-0007>
- Li, Z., Li, F., Fu, Q., Wang, X., Liu, H., Zhao, Y., & Ren, W. (2024). Large language models and medical education: a paradigm shift in educator roles. *Smart Learning Environments*, 11(1), 26. <https://doi.org/10.1186/s40561-024-00313-w>
- Liao, W., Liu, Z., Dai, H., Xu, S., Wu, Z., Zhang, Y., Huang, X., Zhu, D., Cai, H., & Li, Q. (2023). Differentiating ChatGPT-generated and human-written medical texts: quantitative study. *JMIR Medical Education*, 9(1), e48904. <https://doi.org/10.2196/48904>
- Liu, B. L., Morales, D., Roser-Chinchilla, J., Sabzalieva, E., Valentini, A., Vieira do Nascimento, D., & Yerovi, C. (2023). Harnessing the era of artificial intelligence in higher education: a primer for higher education stakeholders. In *Digital Transformation in Higher Education, Part B* (pp. 231-241). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-83608-424-220241009>
- Luckin, R., Rudolph, J., Grünert, M., & Tan, S. (2024). Exploring the future of learning and the relationship between human intelligence and AI. An interview with Professor Rose Luckin. *Journal of applied learning and teaching*, 7(1). <https://doi.org/10.37074/jalt.2024.7.1.27>
- Luckin, R. H., Wayne; Griffiths, Mark and Forcier, Laurie B. (2016). Intelligence unleashed: An argument for AI in education. Retrieved 5th February 2023, from <https://discovery.ucl.ac.uk/id/eprint/1475756/>
- Lund, B. D., & Wang, T. (2023). Chatting about ChatGPT: how may AI and GPT impact academia and libraries? *Library hi tech news*, 40(3), 26-29. <https://doi.org/10.1108/LHTN-01-2023-0009>
- Ma, T., Chen, R., Li, A. T., & Liu, H. (2024). Socratic ChatGPT: Theory, Design, and Empirical Evaluations. *Pacific-Asia Conference on Information Systems*, 12, 1-9.
- Mah, D.-K., & Groß, N. (2024). Artificial intelligence in higher education: exploring faculty use, self-efficacy, distinct profiles, and professional development needs. *International Journal of Educational Technology in Higher Education*, 21(1), 58. <https://doi.org/10.1186/s41239-024-00490-1>
- Marginson, S. (2006). Dynamics of national and global competition in higher education. *Higher education*, 52, 1-39. <https://doi.org/10.1007/s10734-004-7649-x>
- Mian, S. H., Salah, B., Ameen, W., Moiduddin, K., & Alkhalefah, H. (2020). Adapting universities for

- sustainability education in industry 4.0: Channel of challenges and opportunities. *Sustainability*, 12(15), 6100. <https://doi.org/10.3390/su12156100>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R., & Gerardou, F. S. (2023). Challenges and opportunities of generative AI for higher education as explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://doi.org/10.3390/educsci13090856>
- Mohebi, L. (2024). Empowering learners with ChatGPT: insights from a systematic literature exploration. *Discover Education*, 3(1), 36. <https://doi.org/10.1007/s44217-024-00120-y>
- Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5, 100151. <https://doi.org/10.1016/j.caeo.2023.100151>
- Morandín-Ahuerma, F. (2023). Ten UNESCO Recommendations on the Ethics of Artificial Intelligence. *Principios normativos para una ética de la inteligencia artificial*, 86-94. <https://doi.org/10.31219/osf.io/csyux>
- Munna, A., & Kalam, M. A. (2021). Application of theories, principles and models of curriculum design: A literature review. *International Journal of Multidisciplinary and Current Educational Research*, 3(1). <https://repository.uwtsd.ac.uk/id/eprint/1850>
- Newman, J. H., Landow, G. P., Newman, J. E., Turner, F. M., Garland, M. M., & Castro-Klaren, S. (1996). *The idea of a university*. Yale University Press.
- Newton, P., & Xiromeriti, M. (2024). ChatGPT performance on multiple choice question examinations in higher education. A pragmatic scoping review. *Assessment & Evaluation in Higher Education*, 49(6), 781-798. <https://doi.org/10.1080/02602938.2023.2299059>
- Nikolopoulou, K. (2024). Generative Artificial Intelligence in Higher Education: Exploring ways of harnessing pedagogical Practices with the assistance of ChatGPT. *International Journal of Changes in Education*, 1(2), 103-111. <https://doi.org/10.47852/bonviewIJCE42022489>
- Nikolopoulou, K. (2025). Generative artificial intelligence and sustainable higher education: Mapping the potential. *Journal of Digital Educational Technology*, 5(1), ep2506. <https://doi.org/10.30935/jdet/15860>
- Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., & Garro-Aburto, L. L. (2019). Artificial Intelligence and Its Implications in Higher Education. *Journal of Educational Psychology-Propósitos y Representaciones*, 7(2), 553-568. <https://doi.org/http://dx.doi.org/10.20511/pyr2019.v7n2.274>
- Oh, N., Choi, G.-S., & Lee, W. Y. (2023). ChatGPT goes to the operating room: evaluating GPT-4 performance and its potential in surgical education and training in the era of large language models. *Annals of Surgical Treatment and Research*, 104(5), 269. <https://doi.org/10.4174/ast.2023.104.5.269>
- Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., & Ray, A. (2022). Training language models to follow instructions with human feedback. *Advances in neural information processing systems*, 35, 27730-27744.
- Paraschiv, R. (2024). Ethical Implications of Using Artificial Intelligence in Education. *Journal for Freedom of Conscience (Jurnalul Libertății de Conștiință)*, 12(2), 539-549. <https://doi.org/10.5281/zenodo.14860703>
- Preiksaitis, C., & Rose, C. (2023). Opportunities, challenges, and future directions of generative artificial intelligence in medical education: scoping review. *JMIR Medical Education*, 9, e48785.

<https://doi.org/10.2196/48785>

- QAA. (2024a). The frameworks for higher education qualifications of UK degree-awarding bodies. In: Quality Assurance Agency.
- QAA. (2024b). The UK Quality Code for Higher Education 2024. Retrieved 11th February 2025, from <https://www.qaa.ac.uk/the-quality-code/2024#>
- Raihan, M. M., Subroto, S., Chowdhury, N., Koch, K., Ruttan, E., & Turin, T. C. (2024). Dimensions and barriers for digital (in) equity and digital divide: A systematic integrative review. *Digital Transformation and Society*. <https://doi.org/10.1108/DTS-04-2024-0054>
- Rane, N. (2023). Contribution and Challenges of ChatGPT and Similar Generative Artificial Intelligence in Biochemistry, Genetics and Molecular Biology. *Genetics and Molecular Biology (October 16, 2023)*. <https://doi.org/http://dx.doi.org/10.2139/ssrn.4603219>
- Richardson, M., & Healy, M. (2019). Examining the ethical environment in higher education. *British Educational Research Journal*, 45(6), 1089-1104. <https://doi.org/10.1002/berj.3552>
- Roberts, L. D., Chang, V., & Gibson, D. (2017). Ethical Considerations in Adopting a University- and System-Wide Approach to Data and Learning Analytics. In B. Kei Daniel (Ed.), *Big Data and Learning Analytics in Higher Education: Current Theory and Practice* (pp. 89-108). Springer International Publishing. https://doi.org/10.1007/978-3-319-06520-5_7
- Roshanaei, M., Olivares, H., & Lopez, R. R. (2023). Harnessing AI to foster equity in education: Opportunities, challenges, and emerging strategies. *Journal of Intelligent Learning Systems and Applications*, 15(04), 123-143. <https://doi.org/10.4236/jilsa.2023.154009>
- Rosoł, M., Gąsior, J. S., Łaba, J., Korzeniewski, K., & Młyńczak, M. (2023). Evaluation of the performance of GPT-3.5 and GPT-4 on the Polish Medical Final Examination. *Scientific Reports*, 13(1), 20512. <https://doi.org/10.1038/s41598-023-46995-z>
- Royer, C. (2024). Outsourcing humanity? ChatGPT, critical thinking, and the crisis in higher education. *Studies in Philosophy and Education*, 43(5), 479-497. <https://doi.org/10.1007/s11217-024-09946-3>
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of applied learning and teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Ruiz-Rojas, L. I., Salvador-Ullauri, L., & Acosta-Vargas, P. (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. *Sustainability*, 16(13), 5367. <https://doi.org/10.3390/su16135367>
- Sabzalieva, E., & Valentini, A. (2023). ChatGPT and artificial intelligence in higher education: Quick start guide. <https://eduq.info/xmlui/handle/11515/38828>
- Salinas-Navarro, D. E., Vilalta-Perdomo, E., Michel-Villarreal, R., & Montesinos, L. (2024). Using Generative Artificial Intelligence Tools to Explain and Enhance Experiential Learning for Authentic Assessment. *Education Sciences*, 14(1), 83. <https://doi.org/10.3390/educsci14010083>
- Sallam, M., Salim, N. A., Barakat, M., & Ala'a, B. (2023). ChatGPT applications in medical, dental, pharmacy, and public health education: A descriptive study highlighting the advantages and limitations. *Narra J*, 3(1). <https://doi.org/10.52225/narra.v3i1.103>
- Selwyn, N. (2019). Should Robots Replace Teachers? AI and the Future of Education. *John Wiley & Sons, Inc.*

- Sharples, M. (2023). Towards social generative AI for education: theory, practices and ethics. *Learning: Research and Practice*, 9(2), 159-167. <https://doi.org/10.1080/23735082.2023.2261131>
- Smart, J. C. (2008). *Higher education: Handbook of theory and research* (Vol. 23). Springer.
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1), 31-40. <https://doi.org/10.37074/jalt.2023.6.1.17>
- Tamer, S., NAYİR, F., & Bozkurt, A. (2024). Reimagining education: Bridging artificial intelligence, transhumanism, and critical pedagogy. *Journal of Educational Technology and Online Learning*, 7(1), 102-115. <https://doi.org/10.31681/jetol.1308022>
- Tang, C., Li, S., Hu, S., Zeng, F., & Du, Q. (2025). Gender disparities in the impact of generative artificial intelligence: Evidence from academia. *PNAS Nexus*, 4(2). <https://doi.org/10.1093/pnasnexus/pgae591>
- Thapa, S., & Adhikari, S. (2023). ChatGPT, bard, and large language models for biomedical research: opportunities and pitfalls. *Annals of biomedical engineering*, 51(12), 2647-2651. <https://doi.org/10.1007/s10439-023-03284-0>
- Tian, S., Jin, Q., Yeganova, L., Lai, P.-T., Zhu, Q., Chen, X., Yang, Y., Chen, Q., Kim, W., & Comeau, D. C. (2024). Opportunities and challenges for ChatGPT and large language models in biomedicine and health. *Briefings in Bioinformatics*, 25(1), bbad493. <https://doi.org/10.1093/bib/bbad493>
- Tuononen, T., Hyytinen, H., Kleemola, K., Hailikari, T., Männikkö, I., & Toom, A. (2022). Systematic review of learning generic skills in higher education—Enhancing and impeding factors. *Frontiers in education*, 7, 885917. <https://doi.org/10.3389/feduc.2022.885917>
- Varghese, J., & Chapiro, J. (2024). ChatGPT: The transformative influence of generative AI on science and healthcare. *Journal of hepatology*, 80(6), 977-980. <https://doi.org/10.1016/j.jhep.2023.07.028>
- Wang, H., Dang, A., Wu, Z., & Mac, S. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence*, 7, 100326. <https://doi.org/10.1016/j.caeai.2024.100326>
- Wang, S., Zhao, Z., Ouyang, X., Wang, Q., & Shen, D. (2023). Chatcad: Interactive computer-aided diagnosis on medical image using large language models. *arXiv preprint arXiv:2302.07257*. <https://doi.org/10.48550/arXiv.2302.07257>
- Wang, T., Antonacci-Fulton, L., Howe, K., Lawson, H. A., Lucas, J. K., Phillippy, A. M., Popejoy, A. B., Asri, M., Carson, C., & Chaisson, M. J. (2022). The Human Pangenome Project: a global resource to map genomic diversity. *Nature*, 604(7906), 437-446. <https://doi.org/10.1038/s41586-022-04601-8>
- Watson, H., Green, J. R., Nairn, J., Berg, M., Ginger, M., Hammond, K., Hewitt, C., Horrocks, J., Hubbard, K., & Keane, A. (2023). Subject Benchmark Statement Biosciences. https://www.qaa.ac.uk/docs/qaa/sbs/sbs-biosciences-23.pdf?sfvrsn=b570a881_6
- Wei, J., Wang, X., Schuurmans, D., Bosma, M., Xia, F., Chi, E., Le, Q. V., & Zhou, D. (2022). Chain-of-thought prompting elicits reasoning in large language models. *Advances in neural information processing systems*, 35, 24824-24837.
- Whitworth, K., Leupen, S., Rakes, C., & Bustos, M. (2018). Interactive computer simulations as pedagogical tools in biology labs. *CBE—Life Sciences Education*, 17(3), ar46. <https://doi.org/10.1187/cbe.17-09-0208>
- Williams, A. (2024). Comparison of generative AI performance on undergraduate and postgraduate written

- assessments in the biomedical sciences. *International Journal of Educational Technology in Higher Education*, 21(1), 52. <https://doi.org/10.1186/s41239-024-00485-y>
- Williams, A. (2025). Integrating Artificial Intelligence Into Higher Education Assessment. *Intersection: A Journal at the Intersection of Assessment and Learning*, 6(1). <https://doi.org/10.61669/001c.131915>
- Williams, R. T. (2024). The ethical implications of using generative chatbots in higher education. *Frontiers in Education*, 8, 1331607. <https://doi.org/10.3389/feduc.2023.1331607>
- Wong, S. H., Gishen, F., & Lokugamage, A. U. (2021). 'Decolonising the Medical Curriculum ': Humanising medicine through epistemic pluralism, cultural safety and critical consciousness. *London review of education*, 19(1). <https://doi.org/10.14324/LRE.19.1.16>
- Wu, Y. (2023). Integrating generative AI in education: how ChatGPT brings challenges for future learning and teaching. *Journal of Advanced Research in Education*, 2(4), 6-10. <https://doi.org/doi:10.56397/JARE.2023.07.02>
- Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1), 40. <https://doi.org/10.1186/s41239-024-00468-z>
- Xu, Y., Liu, X., Cao, X., Huang, C., Liu, E., Qian, S., Liu, X., Wu, Y., Dong, F., & Qiu, C.-W. (2021). Artificial intelligence: A powerful paradigm for scientific research. *The Innovation*, 2(4). <https://doi.org/10.1016/j.xinn.2021.100179>
- Yu, H. (2024). The application and challenges of ChatGPT in educational transformation: New demands for teachers' roles. *Heliyon*, 10(2). <https://doi.org/10.1016/j.heliyon.2024.e24289>
- Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: a systematic review. *Smart Learning Environments*, 11(1), 28. <https://doi.org/10.1186/s40561-024-00316-7>
- Zhai, X. (2022). ChatGPT user experience: Implications for education. *Available at SSRN 4312418*. <https://doi.org/http://dx.doi.org/10.2139/ssrn.4312418>
- Zhang, P., & Kamel Boulos, M. N. (2023). Generative AI in medicine and healthcare: promises, opportunities and challenges. *Future Internet*, 15(9), 286. <https://doi.org/10.3390/fi15090286>

Author Information

Andrew Williams



<https://orcid.org/0000-0002-0908-0364>

University College London

United Kingdom

Contact e-mail: andrew.williams@ucl.ac.uk
