AI meets academia: the impact of ChatGPT on transforming assessment design in UK universities

AI Meets Academia: The Impact of ChatGPT on Transforming Assessment Design in UK Universities

Introduction

This study is crucial in understanding the impact of Artificial Intelligence (AI) such as ChatGPT in educational settings and its implications for the future of teaching and assessment methodologies. AI has evolved over the last thirty years, currently, there are more than 1300 AI tools, with various new applications emerging (Nikolic et al., 2023) in all sectors of the economy. In as much as various types of AI applications have been developed, their adoption is not congruent across industries and even within the same sector. The education sector is not spared either, as most universities have not openly discussed how educators can adopt the applications in facilitating learning and curriculum development. A Chat Generative Pre-Trained Transformer (ChatGPT) developed in 2022 by OpenAI exacerbated the debates on how and whether learners could utilise AI in assessments. ChatGPT is a powered chatbot with an advanced natural language processing system to respond to questions and engage in conversations with users. The potential of ChatGPT to revolutionise teaching and learning strategies and assessment designs presents academics with an opportunity to rethink their teaching and assessment practices. The chatbot also reached approximately 100 million users, two months after its inauguration (Hu, 2023) and this provided new challenges and threats to higher education, in particular academics as they now question the authenticity of student assessments.

Furthermore, the media played a massive role in increasing the popularity of ChatGPT, therefore, higher education institutions must integrate AI applications in teaching and designing assessments. AI applications are not new to education as age-old literature search engines use these applications. However, machine learning, which is the principle used in AI, results in smarter AI applications presenting a constant need for review of their effectiveness in promoting educational objectives and potentially challenging existing policy debates and regulatory frameworks. While ChatGPT presents unprecedented access to information, it also raises concerns about traditional academic rules and norms. It remains unclear whether current assessment approaches in higher education consider students' access to AI tools, evolve with the emergence of smarter AI applications, mitigate associated risks, and align with technological advancements. Failing to address these issues may disrupt the learning process, challenging traditional educational and pedagogical philosophies and hinder the development of critical thinking skills among students who heavily rely on AI tools for assessments. Despite these concerns, limited research has been conducted to explore the

impact of ChatGPT on assessments. This research paper therefore assesses the influence of technological advancements, specifically ChatGPT, on the design of assessments in higher education institutions in the United Kingdom. The research will also explore how practice aligns with policies, in particular assessment policies given the emergence of newer and smarter AI technologies which continue to evolve through machine learning. The information gathered from the participants allows for rich data from educators who are practitioners in higher education institutions.

Advancements in AI technologies have changed how businesses operate and have also opened prospects for pedagogical reforms (Cheng et al., 2016; Dessi et al., 2019). Amid rapidly evolving reforms, there are calls for the responsible and ethical application of AI in all sectors. It is, however, important to highlight that ChatGPT is not the first and only chatbot that has been developed. Various programs can engage in human-like conversations such as Bert (Google), XiaoIce (Microsoft), and Blender (Facebook) (Agomuoh 2023). However, ChatGPT is viewed as the most advanced chatbot that has ever been created because of its ability to produce outstanding texts in seconds (Mhlanga, 2023). Brown et al. (2020) indicated that GPT-3 is ten times better than any previous non-sparse language model. According to Grand View Research (2019), the AI market size is anticipated to be approximately $390.9 billion by 2025 and applications such as natural language processing, intelligent decision-making, and robotic automation will influence the increase of AI market size. Due to ChatGPT's newness, limited research has focused on its impact on assessment design.

The benefits of ChatGPT have been extensively explored, these include the ability of ChatGPT to respond to user prompts conversationally and naturally (OpenAI, 2023). Therefore, students can enquire about any topic, and they can engage in open discussions, write, and edit reports; generate codes and provide tutoring by explaining codes; provide samples of data for databases and analysis and translate texts to other languages (Halaweh, 2023). However, as for educators, ChatGPT can be used in enhancing pedagogical practice, designing learning assessments and creating brainstorming ideas (Sok and Heng, 2023).

Ivanov and Soliman (2023) explored the implications of ChatGPT for tourism education and learning. They conceded that the application is an effective tool because it can "reject inappropriate requests, challenge inaccurate responses and keep track of what the user stated previously in the chat for follow-up queries". Before the conception of ChatGPT, Sharples et al. (2016) explained that AI had the potential to enhance teaching and learning, develop students' critical thinking, and foster innovation, and interpersonal skills. These skills would assist students to better prepare for the future. Anders (2022) echoed the same sentiments. Strzelechi, (2023) researched the use of ChatGPT in higher education and underscored that

ChatGPT provides an opportunity for higher education institutions to re-evaluate the purpose of assessments and how it can enhance learning. Thus, facilitators of learning need to think about the ways they design assessments by developing assessments that incorporate creativity and go beyond writing.

Even though many researchers have documented the positive impacts of ChatGPT, the chatbot can be viewed as a threat to the education system. Yeadon et al. (2023) provided evidence that a student can write an entire essay using ChatGPT and achieve a First-Class grade. Some researchers argue that this is the power of AI applications in democratising education and access to information. However, it throws into question the authorship of the piece of scholarly work, and whether AI should be cited as a reference or an author (Liebrenz et al., 2023). This has prompted various higher education institutions to devise strategies to reduce the issues of plagiarism and academic dishonesty. Nonetheless, the current plagiarism software does not detect the use of ChatGPT in assessments. Khalil and Er (2023)'s research on determining whether plagiarism tools could detect essays written using ChatGPT revealed that 80% of the essays showed a high degree of originality. This means universities must devise strategies to detect any academic misconduct. This also presents institutions with the question of academic integrity, which has always been debated about contract writing (Ellis et al., 2019). Sok and Heng (2023) also highlighted the risks of ChatGPT regarding academic integrity problems, unfair learning assessment, factual inaccuracies of text, misinformation, and a very thin line with disinformation.

Peters et al. (2022) argue that the emergence of technology-aided plagiarism detection is both a blessing and a curse equally to the student and the education system. Although the need to weed out plagiarism is a pedagogical reform process to regulate academic writing practices, the authors contend that the fast-paced technological advances have necessitated educating students about its ethical and legal implications. Hill et al. (2021) cast the view that the education sector needs to reform itself to keep up with the many faces of plagiarism not only for educational qualifications to remain relevant but also for graduates of the system not to cheat at their workplaces. Assessment reforms remain the highlight, but the problem may not be tackled by one sector and requires the global community to act in unison. However, Hill et al. (2022) further argue that universities have become increasingly their own worst enemy in addressing the problem of plagiarism. The intense competition among universities has redefined education as a commodity where students get value for their money and do not necessarily have to earn a degree through the rigours of intellectual labour. University ranking systems, branding, and penetration of international markets have necessitated a focus on consumer satisfaction to retain a competitive edge. It has become a conflict of interest for

universities to police and take action against their customers whom they intend to please to earn good reviews essential for marketing, jeopardising the reform of policies and assessment strategies to curtail plagiarism.

Although designing authentic assessments could improve the issue of academic integrity, ChatGPT presents educators with an opportunity to be more innovative in assessment design by embracing the emergence of AI. There, however, seems to be a gap between the adoption of AI technologies in teaching and learning and the rapid advancements in technology. The slow adaptation of education to technological advancement was exposed during the outbreak of the COVID-19 pandemic and technological deficiencies within the education sector have been well documented. These range from over-reliance on traditional face-to-face learning and assessment (Guangul et al., 2020), violations of academic integrity due to poor monitoring tools (Amzalag et al., 2022) and the continued proliferation of contract plagiarism and ghostwriting (Hamza et al., 2022). Universities are slow in adopting technology in general as seen with a few universities which adopted Proctortrack software during COVID-19 (Guangul et al., 2020) and currently seem to be slow in adopting AI in assessments, instead, some universities have banned students from using any AI in their assessments (Yau & Chan, 2023). This is because educators are not technologically savvy to use AI applications in facilitating assessments (Ng et al. 2023) and are not well-versed in checking AI academic dishonesty. Furthermore, they are poorly informed on whether and how AI could be effectively adopted in assessments. This has been associated with and is a contributing factor to technostress among educators, affecting the adoption of new technologies by educators (Khlaif et al. 2023). García-Peñalvo, (2023) argues that banning ChatGPT use will not prevent students from using it. Therefore, Mhlanga's (2023) research emphasised the significance of using ChatGPT responsibly and ethically. In the same vein, Halaweh (2023) concluded that educators should allow students to use ChatGPT because it provides them with an opportunity to develop ideas and improve their writing, otherwise, they will use it anyway. Hence, universities could develop policies that provide clear instructions and guidelines to students and educators on how to use ChatGPT in assignments and assessment design respectively. Ivanov and Soliman (2023) state that "few universities have publicly announced any policy toward the application of ChatGPT, but those who have done it explicitly forbid its use". This research therefore examines whether ChatGPT influenced educators to modify their assessments and universities to update their assessment policies. This will ultimately help ascertain whether ChatGPT has prompted educators to adopt more creative and innovative approaches in assessment design.

Materials and Methods

Research Design

This research adopted a pragmatic approach, which permitted the researchers to use both qualitative and quantitative methods. The research was conducted in three phases, the objective of combining the two methodologies (quantitative and qualitative) was to triangulate the methodologies by directly comparing the quantitative statistical results and qualitative findings. It employed convergent parallel research design to explore the influence of ChatGPT, an AI-based chatbot, on assessment design in higher education (Cresswell & Plano Clark, 2018). Therefore, the two datasets were collected simultaneously, analysed separately, and the results were synchronously interpreted. This enhanced the validity and reliability of the findings.

The research aimed to understand whether higher education academics have modified their teaching strategies, particularly in assessment design, due to the emergence of AI technologies like ChatGPT. Furthermore, the research also aimed to explore how practice aligns with academic guidelines and policies, in particular assessment policies given the emergence of newer and smarter AI technologies which continue to evolve through machine learning. Thus, it was plausible to combine both quantitative and qualitative methodologies. Adopting this approach allowed researchers to gain a more comprehensive understanding of ChatGPT's role in shaping assessment strategies in selected UK institutions of higher learning by triangulating different data types (Saunders et al. 2023).

Research questions

The research questions for the study included the following:

What are the existing assessment practices that involve or are influenced by AI, particularly ChatGPT, in your educational institution?

Has there been any changes in assessment policies or guidelines due to the integration of AI tools such as ChatGPT?

What mechanisms or strategies are currently in place to ensure that the usage of AI tools in assessment design aligns with institutional policies and ethical guidelines?

Participants

This research involved 32 academics from three faculties/fields in three higher education institutions. To ensure a diverse representation of faculties/fields, a purposive sampling technique was employed. Thus, questionnaires were distributed through Jisc online surveys. The academics worked in the science department (biomedical sciences) and the tourism/events/hospitality departments; this was essential in analysing whether the impacts of ChatGPT varied from discipline to discipline. For the interviews one participant in each faculty was interviewed at their place of work or via Teams. These are participants who were involved

in policymaking or implementing the assessment policies/guidelines, particularly around ethical conduct. Thus, the participants selected and interviewed include:
Participant 1: A senior lecturer who has worked in higher education for five years. The participant is also an AI lead for the department and is responsible for designing AI guidelines and policies for the department.
Participant 2: An associate professor (teaching and learning) who has worked in higher education for more than fifteen years and is also responsible for designing and reviewing academic guidelines and policies.
Participant 3: A non-academic who has designed and provided academic guidelines to academics for more than 15 years. The participant is also responsible for training academics of the current assessment guidelines including AI.

All the participants were informed of the research before their participation. Hence, their participation was voluntary and an opportunity to withdraw from the research was provided. The academics who were involved in the interview were not included in the survey.

Data Collection

Data was collected through a combination of online surveys and interviews. Jisc Online Surveys was used to collect quantitative data, and a self-administered structured questionnaire was designed to gather comprehensive information about the participants' experiences and views on the use of ChatGPT in assessment design. A survey questionnaire was sent to participants through their work emails. Interviews provided qualitative insights into educators'; perceptions, attitudes, and experiences related to ChatGPT's influence on assessment design.

For the quantitative data collection, a link was sent to the three heads of departments of the three universities/faculties who then shared the link to all the staff in their departments. The questionnaire collected data about participants' demographic information (highest level of education, current role, teaching experience, institution type, and field of expertise). The survey questionnaire also included sections such as:
Personal experience with ChatGPT: Knowledge of ChatGPT, usage in personal work, and views on its impact on assessment strategies.
Assessment practices: Types of assessments conducted, recent changes to assessments, and factors influencing these changes.

As for the qualitative methodology, the participants were purposely selected and an invitation to participate in the interview was sent to the selected participants and a participant information

sheet was attached to the invitation. An interview schedule was designed, and semi-structured interviews were conducted to gather in-depth qualitative data. Three interviews were held, each lasting between thirty to forty minutes. The interviews explored participants' experiences, motivations, and opinions regarding the use of ChatGPT in assessment design. It also explored the policies and guidelines designed by the institutions to respond to the emergence of Generative AI. The data collection for the survey and interviews were conducted simultaneously

Data analysis

Quantitative and qualitative data were analysed separately. Quantitative data from the questionnaire were analysed using Statistical Package Social Sciences (SPSS) Version 28. Descriptive statistics were used to summarise responses' central tendencies, dispersion, and distribution. Additionally, hypothesis testing was conducted using the Independent-Samples Kruskal-Wallis Test to examine differences in responses across different demographic categories, such as age range, gender, level of education, and teaching role. Qualitative data, particularly regarding the open-ended questions about assessment changes and perceptions of ChatGPT's impact, were analysed using Thematic Content Analysis (TCA). This started by reading and reviewing the transcripts and post-interview notes to familiarise with the data. To organise the data in a systematic and meaningful approach, the data was highlighted and coded by hand. Coding the responses to identify common themes and patterns, provided deeper insights into the educators' perspectives on the impact of AI in assessment design. This process contributed to identifying core consistencies and meanings of the qualitative data. The interview schedule was sectioned into five categories, these included:

1. Participant Background
2. Current Assessment Practices and Policies concerning AI.
3. Impact of AI Tools on Assessment Design
4. Mechanisms and Strategies to Mitigate/Identify AI Technology Usage in Assessments
5. Future Perceptions

From the above categories each portion of text that was relevant to the research or that was closely aligned to the research question was coded and recurring themes were identified. The transcripts were re-read to identify individual units. Similar and dissimilar units were then grouped and re-grouped, and the categories were also re-labelled. Several themes were identified.

Results and Discussion

The results and discussion section were combined (quantitative and qualitative) to offer a wider outlook of how Higher Education has been impacted by ChatGPT. This enables direct comparison and integration of the two different approaches. Hence, discrepancies or confirmations between the datasets have been addressed. The study investigated the use of ChatGPT by participants most of which were under 50 years. Table. 1 below shows the demographic and occupational characteristics of participants.

Table 1: Study participants' demographic and occupational characteristics

Age range	n (%)
Under 25 years	4 (12.5)
25-34 years	7 (21.9)
35-44 years	10 (31.3)
45-54 years	9 (28.1)
55 years or above	2 (6.3)
Gender	
Female	14 (43.8)
Male	18 (56.3)
Education level	
Bachelor's degree	1 (3.10)
Master's degree	10 (31.3)
Doctoral degree	21 (65.6)
Teaching Experience	
≤5 Years	4 (12.5)
>5 to ≤10 Years	6 (18.8)
>10 to ≤15 Years	7 (21.9)
>15 Years	15 (46.9)
Job category (Current role)	
Lecturer	9 (28.1)
Senior Lecturer	19 (59.4)
Associate /Professor	4 (12.5)
Field of expertise	
Sciences	15 (46.9)
Humanities	17 (53.1)
Taught class size	
Small	13 (40.6)

Medium	16 (50.0)
Large	3 (9.4)

Class size definition in terms of students per class: Small <40, medium 41-80, and large >80

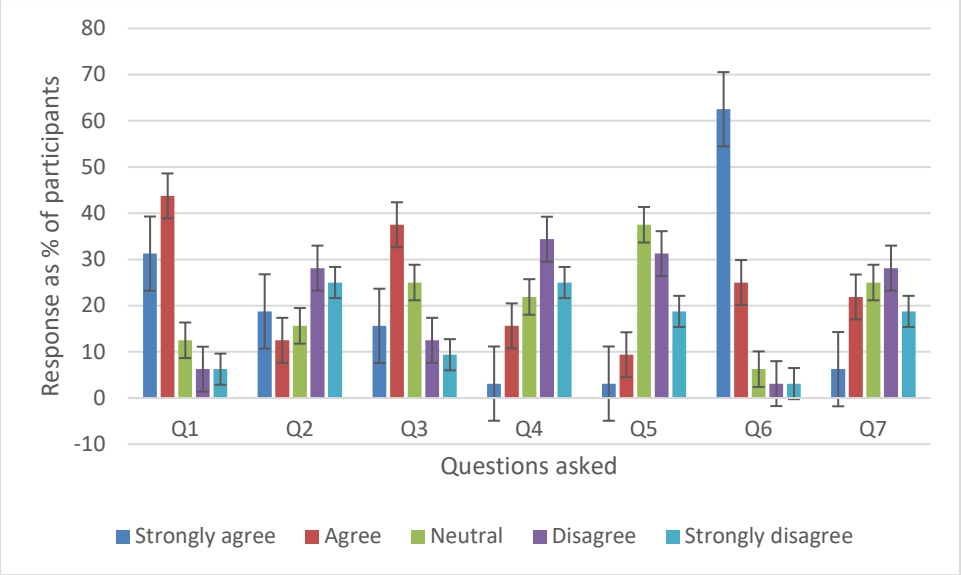
The research focused on facilitators of learning from higher education institutions. Most participants were between the ages of 35 and 54 years with a small proportion (12.5%) being under 25 years. It is therefore plausible to expect this staff group to have the ability to comprehend ChatGPT and its functionalities (Table.1). This assertion is also supported by the fact that all but one participant had a postgraduate qualification with most of them being PhD holders.

It is pertinent to note that according to Brown (2011) and Prensky (2001) these are Generation X and Generation Y who are digital immigrants and digital natives respectively. These two generations exhibit differences in their familiarity with digital technology (Brown 2011). Generation Y, the first digital native generation, is typically competent in and can comprehend ChatGPT and its functionalities. The same, however, cannot be said for many members of Generation X and HEIs need to be aware of the need to capacitate this group of academics as they have not been exposed to technology to the same extent as the younger Generation Y.

Pedagogical approaches which include delivery of teaching, facilitating learning and assessment of learners was at the core of what the participants have been doing. All the participants had more than 5 years of HEI experience with almost half of them (46.9%) having been in higher education for more than 15 years.

The participants indicated an interesting view on the extent to which they are knowledgeable about ChatGPT. Figure 1 below shows participants' opinions on CHATGPT knowledge on assessment.

Figure 1: Participants' Views on CHATGPT knowledge on assessment



Key for Figure 1

Q1. I have knowledge of ChatGPT
Q2. I have used ChatGPT in my own personal work
Q3. ChatGPT is a supportive tool for educators
Q4. I can teach students how to comprehend the large volume of content generated by AI
Q5. ChatGPT improves my critical thinking skills
Q6. Training on the responsible use of AI will be beneficial to me
Q7. I have knowledge of other AI based Chatbots other than ChatGPT

There is a stark contrast between teaching experience and how knowledgeable the academics are on AI. Although more than 65% of the participants knew about ChatGPT, almost 50% of the participants did not know about other AI chatbots. Yet, there are more than 1300 AI tools, with various new applications emerging (Nikolic et al., 2023). Moreover, more than 85% of these would value training on ChatGPT.

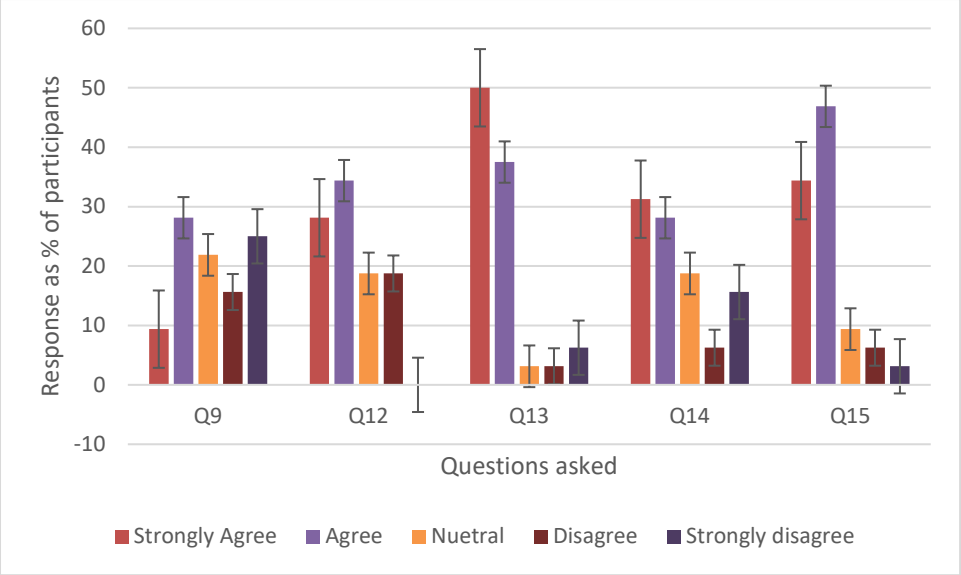
The findings of this research are consistent with those reported by Ceras and Balcioğlu (2023), whose study identified the need for better guidelines as well as training. The findings suggest that there is an appetite for training, and it is important to capitalise on this appetite to ensure the feasible implementation of assessment policies that address the use of ChatGPT and other AI chatbots, to adequately respond to the academic disruption precipitated by AI (Khan et al, 2023). Recent research by González-Calatayud, Prendes-Espinosa, and Roig-Vila (2021) highlighted a gap in research on the pedagogical implications of AI. Moreover, their study findings suggest that most AI studies that investigated the application and impact of AI in

educational assessment neglect pedagogical considerations. Furthermore, AI applications in assessments are predominantly limited to formative evaluations. This finding suggests that while AI's applicability spans various fields, its potential might be restricted due to the limited scope of its current implementations. This research also highlights the need for further training on how AI can be effectively explored and ethically utilised in educational settings.

From the interviews conducted, participants responded to questions on knowledge and AI impact on assessment. Three participants shared the same sentiment that training on AI use should be readily available. The key themes that emerged from interviews were training and leadership involvement. When asked whether any training has been offered to academics to equip them on the use of AI; participant 3 indicated that: *the trouble is we are trying to understand and figure out how AI works ourselves so no training has been offered so far! But we liaising with our teaching and learning technology team to develop some sort of training.*

Both participants 1 and 2 indicated that providing training to academics is essential, however, the students should also be trained on how to use ChatGPT. Participant 2 further expounded that: *the truth is students are using ChatGPT to write their assignments so is it not better to teach them how to use it responsibly?.* This aligns with García-Peñalvo,(2023); Mhlanga, (2023); & Halaweh (2023) who concluded that students should be trained on how to use ChatGPT responsibly and ethically instead of banning it. The main challenge highlighted was the difficulty in identifying AI generated work because of the lack of training.
The interview participants concurred that they need more understanding of AI for them to be able to provide clear training guidelines. Participant 1 stated that leadership involvement would provide academics with a clear direction on how to design authentic assessments and this also requires an understanding of how ChatGPT works. Participant 3 added that: *given that ChatGPT was launched more than a year ago, we should ahead of the game or least have a policy.* This research therefore agrees with Guangul et al., (2020) who highlighted that universities are slow in adopting technology in general. Figure 2 highlights the use of ChatGPT in student assessments

Figure 2: Use of ChatGPT in student assessments



Key for figure 2

Q9. I encourage/ will encourage my students to use ChatGPT
Q12. I feel my assessments may be vulnerable to non-human authorship so I have modified them
Q13. The emergence of ChatGPT has challenged the role of traditional human endeavours in assessments
Q14. Students should acknowledge the use of ChatGPT as a source or co-author in their assessments
Q15. It is difficult to estimate the disruptive nature of ChatGPT in my assessments

The use of ChatGPT by students has been documented by Ngo (2023), and it highlighted benefits, albeit according to students. According to students, the benefits included saving time and providing personalised feedback. In this research, 45% of the respondents were amenable to the idea of encouraging students to use ChatGPT but another 40% were opposed to this notion. This near equidistant disparity in notion reflects the demographics of our participants who were distributed between digital immigrants (Gen X) and digital natives (Gen Y). It is pertinent to note that whether students are encouraged or not, they are already embracing ChatGPT and other chatbots in their learning. Sullivan et al (2023) indicated that one-third of students were using ChatGPT for their assessment, in particular, essay writing and this figure is likely to increase.

Despite the observation from students' use of ChatGPT, there was more concordance on the idea that assessments may be vulnerable to AI authorship. This response was also confirmed by one on the fact that ChatGPT has challenged the role of traditional human endeavours in assessments. More than 85% of respondents agree that the emergence of ChatGPT has posed a challenge in the way traditional assessments were run. The disruption ChatGPT has caused to the assessments processes was also notable with more than 80% of participants

highlighting that it is difficult to estimate the disruptive nature of ChatGPT in assessments. This is compounded by the lack of effective adoption of AI in assessments which was described by Khlaif et al (2023). All the three interview participants indicated that students should be taught how to use ChatGPT, and participant 1 expanded that the use AI tools should be integrated into the curriculum to keep up with the changing technology.

The research also explored the impact AI had on assessment design and administration. Some progress has been made with more than 40% of participants noting a change in assessment strategies. However, 56.2% of the participants have not changed their assessment strategies despite the increase in the use of ChatGPT in assignments by students. Participants were asked what they perceived as recent drivers in changes to assessment strategies. An analysis of their responses is presented in Table 3.

Table 3: Drivers of changes made to assessment strategies

Drivers of changes made to assessments	
Reason for change	Numbers (%)
Feedback	6 (18.8)
Knowledge of AI use	4 (12.5)
Curriculum Review	8 (25)
Own initiative	14 (43.8)

The changes were, however, initiated by individuals (43.8%) and not part of an institutional drive. Twenty-five percent of participants indicated curriculum review as a driver for the change, with cycles for review coming up; it is expected that AI use will become more prominent in both teaching and assessment.

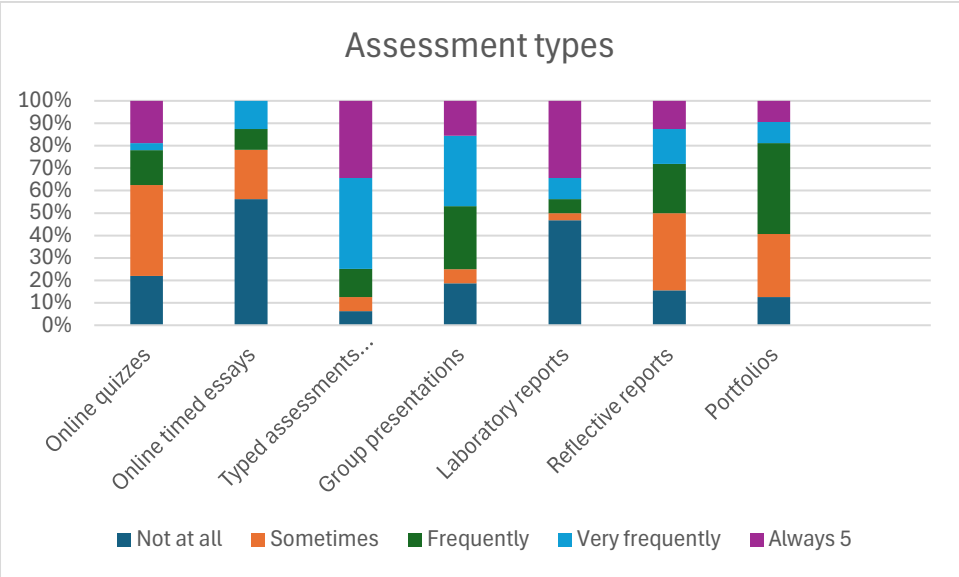
Some assessments are more prone to AI misuse than others and it was pertinent to note that typed assessments are still popular with more than 70% still using them. A survey by Sullivan et al., (2023) indicated that one-third of students were using ChatGPT for essay writing. It is therefore pertinent to suppose a significant number of students could still be using this technology in their assessments. ChatGPT and similar AI chatbots generate content that can easily be copied and pasted (Transformer., Osmanovic., & Steingrimsson; 2022) making it easy to plagiarise and use in typed assessments.

On the other hand, timed essays are seldom used with almost 80% respondents reporting that these are not being used. When asked whether ChatGPT has influenced a change in assessments, all the three interview participants acknowledged that there is a need to review

the current assessments, however, it seems this is up to the individual educators not an institutional stance. Participant 1 added that it is the responsibility of educators to prepare students for the future and emphasised that there is a mixed reaction on the adoption of AI in assessments. Therefore, the mixed reactions on the use of AI can create unequal perception and message to the students.

Figure 3 below shows a breakdown of all the assessments mentioned in the survey instrument.

Figure 3: Commonly used Assessment types



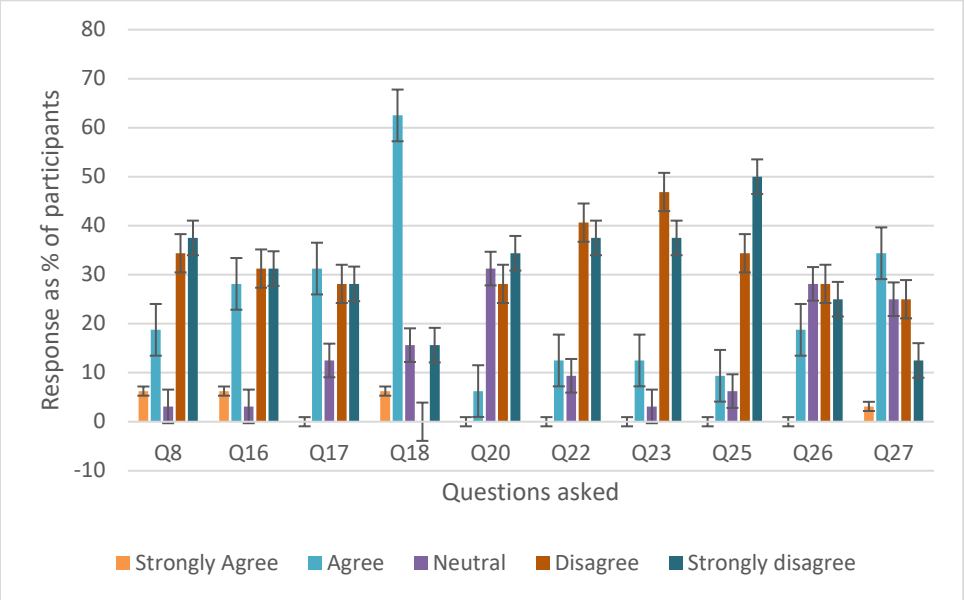
It is easy to evade commonly used plagiarism detection software such as Turnitin, Yeadon et al (2023) alluded to the fact that AI-generated content produces similarity scores between 2 +/- 1%(Grammarly) and 7 +/- 2%(Turnitin). With such low scores, essay-based assessments are therefore becoming redundant unless more innovative approaches to assessing essays are introduced. All the interviewees indicated that plagiarism tools such as Turnitin have become redundant as AI is smarter than Turnitin. The key theme that emerged was authenticity.

Participant 1 stated that:

the emphasis is on academics focusing on redesigning their assessments, taking into account that we are beyond just demonstration and knowledge, but taking into account, you know, authentic real-world assessments looking at metacognition, I think that's a large and important area.

Participant 3 added that: *ChatGPT is very new to us but what we have done is develop a few workshops on authentic assessment design for the next academic year.* However, participant 2 that the current assessments do not need redesigning because the department does not see how ChatGPT has an impact on them. The department focuses on practical laboratory reports. This shows that the impact of ChatGPT varies from the field of study. For the departments focusing on sciences, the impact is quite limited in comparison to humanities.

Figure 4: Policies and guidelines on the use of AI in assessment



Key for figure 4

Q8. Leadership has put measures in place to ensure responsible use of AI
Q16. The institution provides information on the role of ChatGPT to the education system
Q17. Academic development sessions focussing on the constructive and disruptive nature of AI are delivered
Q18. There is recognition for the potential disruptive role of ChatGPT to assessments by the institution
Q19. ChatGPT has been adopted in assessments to encourage creativity rather than prediction in assessments
Q20. The Institution is promoting AI to improve student outcomes and targeted interventions during assessments
Q21. There are ethical considerations in the adoption of AI based chatbot in assessments
Q22. There are clear guidelines for the adoption of ChatGPT into assessments for educators
Q23. There are clear guidelines for the adoption of ChatGPT into assessments for students
Q24. The use ChatGPT has been considered in assessments for students with a need reasonable adjustment
Q25. Institution equips educators to deal with the disruptive nature of ChatGPT in assessments
Q26. There is a policy to deal with potential copyright infringements caused by ChatGPT in assessments
Q27. Guidelines on cases of academic misconduct during assessments include the inappropriate use of ChatGPT

It has become evident as highlighted by Ceras and Balcioğlu (2023) that there is a need for assessment regulations that are in keeping with AI and the current information technology developments. In this research, 80% of respondents report having heard or experienced reports of students using ChatGPT in their assessments. A similar proportion of respondents, however, mentioned that there are no clear guidelines available for adopting ChatGPT in their institutions.

The key themes that emerged from the interviews are lack of communication, guidelines, and policies. This research found that there is a lack of communication between the leadership and educators in terms of whether AI should be adopted in teaching, learning, and assessments. This could cause a lot of confusion to students as each educator could provide their own suggestion. Participant 2 stated that the lack of a general AI policy has been a challenge for the institution to develop AI guidelines and policies. The participant stated that: *maybe we should wait and see what others are doing, let's wait and see!* Participant 1 also added that: *the policy we designed is vague, ambiguous, and difficult to ascertain whether we should use AI.* None of the interviewed institutions indicated whether ChatGPT should be banned or not. This lack of clear guidelines could also have an impact on academic integrity and unfairness in assessments.

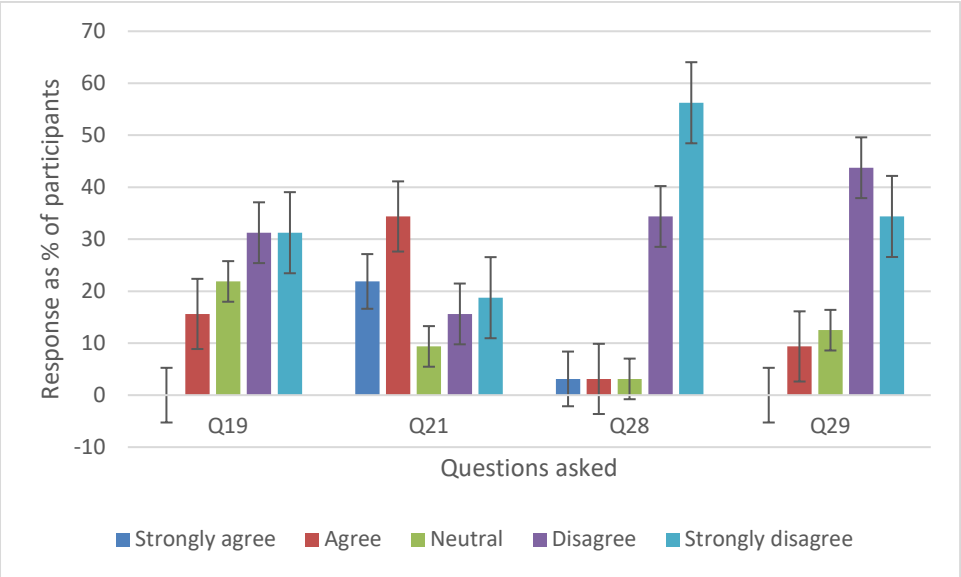
This sentiment was echoed by Dai et al. (2023) who highlighted that there is a need for clear guidance on AI. More than 80% of respondents also noted that students are not getting guidance on how they can effectively use ChatGPT to improve their learning experience. Without this guidance for both the learners and educators, it becomes difficult to standardise approaches to plagiarism-related academic offenses (Dai, Liu, & Lim, 2023). This conclusion agrees with Participant 2 who highlighted that the number of academic offenses increased in the department and added that: *but how do we prove that the students used AI, this has immense pressure and stress on the lecturers and students.* This aligns with Khlaif et al. (2023)'s research who confirmed that the lack of how AI could be effectively adopted in assessments is a contributing factor to technostress among educators. Thus, some academics may choose to overlook some academic offences. Furthermore, it supports Ng et al., (2023)'s notion that educators are not technologically savvy to use AI applications in facilitating assessments and are not well-versed in checking AI academic dishonesty.

One area HEIs can use to drive change in working practice is through the use of guidelines and regulatory frameworks that support and guide procedures on AI use and assessment strategies. In this research, it became evident that institutions have not attempted to address this area. About 85% of the participants confirmed that their institutions do not have policies in place to deal with copyright infringements caused by ChatGPT. This highlights the lack of

preparedness by HEIs to deal with emerging AI chatbots that have a direct impact on assessments.

A slightly higher proportion (35%) confirmed that guidelines on assessment misconduct address instances where ChatGPT has been used inappropriately. These low percentages further confirm Dai et al. (2023) sentiments that there is a need for clear guidance on AI use. It is therefore plausible to expect more concerted efforts in this regard from HEIs if assessments and assessment designs are to fully embrace ChatGPT and other AI chatbots.

Figure 5: Adoption of AI in assessment design



Key for figure 5

19 ChatGPT has been adopted in assessments to encourage creativity rather than prediction in assessments
21 There are ethical considerations in the adoption of AI based chatbot in assessments
28 Digital infrastructure and software currently used to check for similarity in assessments can detect inappropriate use of ChatGPT
29 I have used ChatGPT, and other AI based chatbots to design my assessments

There has been a notable lack of drive to adopt ChatGPT in assessments to encourage creativity. The lack of drive has been confirmed by this research with 0% of the participants strongly agreeing that ChatGPT has been adopted in assessment to encourage creativity. This lack of agreement among participants emphasises the need for HEIs to change and start paying more attention to the emergence of ChatGPT and other AI chatbots. These changes would also require the introduction of infrastructure that supports the proposed changes. In this regard, digital and Information Technology (IT) and digital infrastructure is of paramount importance. Despite Guangul et al., (2020) highlighting the importance of up-to-date IT

systems in higher education to match advances in technology, this research found that institutions lack digital and IT infrastructure that can detect inappropriate use of ChatGPT. Ninety percent of respondents share the same sentiments that current IT infrastructure cannot support the use of ChatGPT or AI tools in teaching and learning.

In addition to the lack of suitable IT infrastructure, respondents noted that Higher Education Institutions need to adjust current assessments in the face of ChatGPT. With more than 80% of the respondents alluding to this notion, it is worth exploring how best HEIs can develop their workforce, so they are competent in dealing with AI and assessments. The key theme that emerged is innovativeness. This is linked to assessment design. All the participants indicated they had not designed any training on assessment design. Participant 1 however, stated that:

So, if we get students good at understanding how to use the tools, our own assessment design will evolve with that…. it has to be 2 halves of the same coin or two sides of the same coin. And if we simply use ChatGPT, your AI tools to enhance assessment design and we weren't empowering students to respond to those, then there's a deficit there, in my view.

More than half of the participants (>55%) agreed or strongly agreed that there are ethical considerations to be mindful of when adopting ChatGPT in assessments. Recent studies have also highlighted the critical need for clear ethical use guidelines, to guide and explore the educational benefits of chatbots such as ChatGPT, prevent its misuse, and uphold academic integrity while mitigating its potential impacts in the education sector (Ipek et al. 2023; Abdulai & Hung, 2023).

Although 43.8% of the academics have changed their assessments to respond to ChatGPT, most of them did not use ChatGPT to design their assessments. Only 9% indicated that they have used ChatGPT or other AI chatbots to design their assessments. This could be influenced by the lack of training and/or the difficulty in estimating the disruptive nature of ChatGPT in assessments. Although all the interview participants indicated the importance of innovativeness, it seems traditional assessment strategies are still being used in all three institutions. Participant 3 indicated that the future of higher education would require learners to have more ownership of their learning and less involvement of academics as gatekeepers. Currently academics as gatekeepers use assessments as a tool to determine whether learning has taken place. Therefore, with time, learners would need to take this responsibility and naturally, this would allow them to use AI effectively and responsibly. This sentiment is aligned with Participant 2's interpretation that empowering learners to use AI tools in their assessments could enhance the quality of student work. This suggests that educators should harness the power of AI to create more innovative assessment designs.

Conclusions

ChatGPT's potential to transform teaching methods, learning strategies, and assessment designs offers academics a unique opportunity to reimagine and innovate their teaching and evaluation approaches. However, higher education institutions have been slow in implementing technological innovations. While many academics are concerned by the influence of AI- on assessment strategies and have made some strides in redesigning their assessments to still attain the intended assessment outcomes, there are no documented institutional strategies to their knowledge. This is exacerbated by a lack of training in AI for academics. It is incumbent upon institutions to offer guidance on the responsible and ethical use of AI in general and for these institutions to stay relevant and survive the fast-paced ever-changing face of digital technology. Students are already using AI in different forms for assessment purposes and as usual, higher education institutions seem to be playing catch-up.

Limitations

Sampling for this research involved 32 academics from three faculties. Although this was an adequate sample but could be more diverse by incorporating a broader range of faculties. Therefore, to further this research, future work could focus on more universities from different regions to give a more generalised view of AI and assessment design.

References

Abdulai, A. and Hung, L. (2023) 'Will CHATGPT undermine ethical values in nursing education, research, and practice?', *Nursing Inquiry*, 30(3). doi:10.1111/nin.12556.

Agomuoh, F. (2023). The 6 biggest problems with ChatGPT right now. *Digital trends,* [online], available at: https://www.digitaltrends.com/computing/how-to-use-open ai-chatbot-text-generation-chatbot/〉 accessed: (24 August 2023).

Amzalag, M., Shapira, N., & Dolev, N. (2022). Two Sides of the Coin: Lack of Academic Integrity in Exams During the Corona Pandemic, Students' and Lecturers' Perceptions. *Journal of academic ethics*, *20*(2), 243–263. https://doi.org/10.1007/s10805-021-09413-5

Anders, B.A. (2022), *"Why ChatGPT is such a big deal for education",* [online], available at: https://scalar.usc.edu/works/c2c-digital-magazine-fall-2022—winter-2023/why-chatgpt-is-bigdeal-education (accessed 10 February 2023).

Brown, T. (2011) ‘Are you a *digital native* or a *digital immigrant*? being client centred in the Digital Era’, *British Journal of Occupational Therapy*, 74(7), pp. 313–313. doi:10.4276/030802211x13099513660992.

Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems,* 33, pp.1877–1901.

Cheng, X., Li, Y., Sun, J., & Huang, J. (2016). Application of a novel collaboration engineering method for learning design: A case study*. British Journal of Educational Technology*, 47(4), 803–818.

Creswell, J.W. and Plano Clark, V.L. (2018) *Designing and conducting mixed methods research*. Thousand Oaks, CA: SAGE.

Ceras, C. C., & Balcioğlu, Y. S. (2023). AI and ethics in education: ChatGPT program example. *International Symposium on Communication and Technology with Philosophical Dimensions*. Marmara University, Istanbul, Turkey.

Dai, Y., Liu, A. and Lim, C.P. (2023) *Reconceptualizing Chatgpt and Generative AI as a student-driven innovation in Higher Education* [Preprint]. doi:10.35542/osf.io/nwqju.

Dessi, D., Fenu, G., Marras, M., & Recupero, D. R. (2019). Bridging learning analytics and Cognitive Computing for Big Data classification in micro-learning video collections. *Computers in Human Behavior,* 92, 468–477.

Ellis, C., Van Haeringen, K., Harper, R., Bretag, T., Zucker, I., Mcbride, S., Rozenberg, P., Newton, P. & Saddiqui, S. (2019) ‘Does authentic assessment assure academic integrity? evidence from contract cheating data’, *Higher Education Research & Development*, 39(3), pp. 454–469. doi:10.1080/07294360.2019.1680956.

García-Peñalvo, F. J. (2023). The perception of artificial intelligence in educational contexts after the launch of ChatGPT: Disruption or panic? *Education in the Knowledge Society,* 24, e31279. https://doi.org/10.14201/eks.31279

González-Calatayud, V., Prendes-Espinosa, P. and Roig-Vila, R. (2021) ‘Artificial Intelligence for Student Assessment: A Systematic Review’, *Applied Sciences*, 11(12), p. 5467. doi:10.3390/app11125467.

Grand View Research. (2019). Artificial intelligence market size, share & trends analysis report by solution, by technology (deep learning, machine learning), by end use, by region, and segment forecasts, 2019–2025. Available at: https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market/ Accessed: (17 July 2023)

Guangul, F.M., Suhail, A.H., Khalit, M.I. *et al.* Challenges of remote assessment in higher education in the context of COVID-19: a case study of Middle East College. *Educ Asse Eval Acc* 32, 519–535 (2020). https://doi.org/10.1007/s11092-020-09340-w

Halaweh, M. (2023) 'Chatgpt in education: Strategies for responsible implementation', *Contemporary Educational Technology*, 15(2). doi:10.30935/cedtech/13036.

Hamza, M. A., Al Assadi, F. R., Khojah, A. A., AlHanaki, R. M., Alotaibi, N. T., Kheimi, R. M., Salem, A. H., & Marar, S. D. (2022). Contract Cheating and Ghostwriting among University Students in Health Specialties. *Journal of empirical research on human research ethics : JERHRE*, *17*(5), 536–544. https://doi.org/10.1177/15562646221128418

Hill, G., Mason, J., & Dunn, A. (2021). Contract cheating: an increasing challenge for global academic community arising from COVID-19. *Research and practice in technology enhanced learning*, *16*(1), 24. https://doi.org/10.1186/s41039-021-00166-8

Hu, K. (2023). ChatGPT sets record fastest growing user base – analyst note. *Reuters.* (Accessed 4 March 2023).

İpek, Z.H., Gozum, A., Papadakis, S., & Kallogiannakis, M. (2023) 'Educational applications of the CHATGPT AI system: Systematic review research', *Educational Process International Journal*, 12(3). doi:10.22521/edupij.2023.123.2.

Ivanov, S. and Soliman, M. (2023) 'Game of algorithms: ChatGPT implications for the future of tourism education and research', *Journal of Tourism Futures*, 9(2), pp.214–221. doi:10.1108/jtf-02-2023-0038.

Khan, R. A., Jawaid, M., Khan, A. R., & Sajjad, M. (2023). ChatGPT - Reshaping medical education and clinical management. *Pak J Med Sci, 39*(2), 605-607. https://doi.org/10.12669/pjms.39.2.7653

Khlaif, Z. N., Sanmugam, M., Joma, A. I., Odeh, A., & Barham, K. (2023). Factors Influencing Teacher's Technostress Experienced in Using Emerging Technology: A

Qualitative Study. *Technology, Knowledge and Learning*, *28*(2), 865–899. https://doi.org/10.1007/s10758-022-09607-9

Liebrenz, M., Schleifer, R., Buadze, A., Bhugra, D., Smith, A. Generating scholarly content with ChatGPT: ethical challenges for medical publishing. *The Lancet Digital Health,* 5(3): e105-e106. doi.org/10.1016/S2589-7500(23)00019-5.

Mhlanga, D. (2023). Open AI in education, the responsible and ethical use of ChatGPT towards lifelong learning SSRN. https://doi.org/10.2139/ssrn.4354422

Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W., & Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational technology research and development: ETR & D*, *71*(1), 137–161. https://doi.org/10.1007/s11423-023-10203-6

Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G.M., Grundy, S., Lyden, S., Neal, P., & Sandison, C. (2023) 'CHATGPT versus Engineering Education Assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this Generative Artificial Intelligence Tool to Investigate Assessment Integrity', *European Journal of Engineering Education*, 48(4), pp. 559–614. doi:10.1080/03043797.2023.2213169.

OpenAI (2023). Available online: https://openai.com (accessed on 15 May 2023).

Peters, M.A., Jackson, L., Hung, R., Mika, C., Buchanan, R.A., Tesar, M., Besley, T., Hood, N. 2022. The case for academic plagiarism education: A PESA Executive collective writing project, *Educational Philosophy and Theory,* 54:9, 1307-1323, DOI: 10.1080/00131857.2021.1897574

Prensky, M. (2001) 'Digital Natives, digital immigrants part 1', *On the Horizon*, 9(5), pp. 1–6. doi:10.1108/10748120110424816.

Saunders, M.N.K., Lewis, P. and Thornhill, A. (2023) *Research methods for business students*. Harlow, England: Pearson.

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), 1–10. https://doi.org/10.37074/jalt.2023.6.1.17

Sharples, M., De Roock, R., Ferguson, R., Gaved, M., Herodotou, C., Koh, E., … Wong, L. H. (2016) innovating pedagogy 2016*: Exploring new forms of teaching ...* Available at:

https://www.researchgate.net/publication/311607636_Innovating_Pedagogy_2016_Exploring_new_forms_of_teaching_learning_and_assessment_to_guide_educators_and_policy_makers (Accessed: 19 July 2023).

Sok, S. and Heng, K. (2023) 'ChatGPT for Education and research: A review of benefits and risks', *SSRN Electronic Journal* [Preprint]. doi:10.2139/ssrn.4378735.

Strzelechi, A. (2023) 'To use or not to use ChatGPT in higher education? A study of students' acceptance and use of Technology', *Interactive Learning Environments*, pp. 1–14. doi:10.1080/10494820.2023.2209881.

Ngo, T. T. A. (2023). The perception by university students of the use of ChatGPT in education. International Journal of Emerging Technologies in Learning (Online), 18(17), 4.

Transformer, G.G.P, Osmanovic Thunström, A., & Steingrimsson, S. (2022). Can GPT-3 write an academic paper on itself, with minimal human input? https://hal.archives-ouvertes.fr/hal-03701250/document

Yau, C., & Chan, K. (2023). University of Hong Kong temporarily bans students from using ChatGPT, other AI-based tools for coursework. South China Morning Post. 17 February 2023.

Yeadon, W., Inyang, O., Mizouri, A., Peach, A. & Testrow, C.P. (2023) 'The death of the short-form physics essay in the Coming Ai Revolution', *Physics Education*, 58(3), p. 035027. doi:10.1088/1361-6552/acc5cf.