

2024

EJEL Volume 22, Issue 2



Editors

Marija Cubric and Paul Griffiths

Published by Academic Publishing
International Limited

Curtis Farm, Kidmore End, Nr Reading, RG4
9AY, United Kingdom

karen.harris@academic-publishing.org

eISSN: 1479-4403

EJEL Volume 22, Issue 2

Special Issue: Artificial Intelligence (AI) in Education Part 1 Contents

Exploring the Role of ChatGPT as a Peer Coach for Developing Research Proposals: Feedback Quality, Prompts, and Student Reflection <i>Johannes Cronjé</i>	01-15
Cheaters or AI-Enhanced Learners: Consequences of ChatGPT for Programming Education <i>Niklas Humble, Jonas Boustedt, Hanna Holmgren, Goran Milutinovic, Stefan Seipel and Ann-Sofie Östberg</i>	16-29
ChatGPT and Assessment in Higher Education: A Magic Wand or a Disruptor? <i>Maira Klyshbekova and Pamela Abbott</i>	30-45
Have Courage to Use your Own Mind, with or without AI: The Relevance of Kant's Enlightenment to Higher Education in the Age of Artificial Intelligence <i>Alice Watanabe</i>	46-58
Educators' Academic Insights on Artificial Intelligence: Challenges and Opportunities <i>Jayaron Jose and Blessy Jayaron Jose</i>	59-77
Enhancing English as a Foreign Language (EFL) Learners' Writing with ChatGPT: A University-Level Course Design <i>Yu-Ching Tseng and Yi-Hsuan Lin</i>	78-97
Exploring the Feasibility and Efficacy of ChatGPT3 for Personalized Feedback in Teaching <i>Irum Naz and Rodney Robertson</i>	98-111
ChatGPT in Education – Understanding the Bahraini Academics Perspective <i>Amal Alrayes, Tara Fryad Henari and Dalal Abdulkarim Ahmed</i>	112-134

Exploring the Role of ChatGPT as a Peer Coach for Developing Research Proposals: Feedback Quality, Prompts, and Student Reflection

Johannes Cronjé

Faculty of Informatics and Design, Cape Peninsula University of Technology, South Africa

cronje@cput.ac.za

<https://doi.org/10.34190/ejel.21.5.3042>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: This paper describes the results of an archival desk-study that analyzed worksheets produced by four students using ChatGPT as a coach. ChatGPT is a Generative Pre-Trained Large Language model that can write comprehensively in various languages and styles. It was discovered that it could pass university level physics exams and perform at the level of a third-year medical student. Fourth-year students in Information Technology are required to produce a 3000-to-5000-word research report as part of the requirements for graduation, and their first meeting of the year consists of a workshop following the GROW (Goal, Reality, Opportunities, Will) coaching process. Logistical considerations in this predominantly e-learning based course prevented such teamwork. The two-month-old ChatGPT, however, presented an opportunity to determine the extent to which a chatbot could be used as a peer coach. This paper reports on the outcome of an exercise that was given to students as an introduction to their research methodology course. It was found that well-designed prompts were essential and that students should be encouraged to reflect on their interaction with artificial intelligence. Three clear issues emerged from the study: the quality of feedback given by the chatbot, the value of the prompt, and the importance of student reflection. The chatbot, ChatGPT, displays the traditional computer characteristic of garbage in, garbage out and gives the feedback it is programmed to give. The dialogue shows how a student can build up mutual rapport with the chatbot when they adapt their responses to the feedback provided. The critical reflection in shows that it is still the user who should oversee the process and evaluate the feedback. It is recommended that novice research students be given some training in reacting to feedback, developing useful prompts, and engaging in meaningful reflection. Future research would include developing a reflective coach.

Keywords: Artificial Intelligence, ChatGPT, Research methods, GROW coaching

1. Introduction and Problem Statement

ChatGPT is a Generative Pre-Trained Large Language model that can write comprehensively in various languages and styles. With its launch on 30 November 2022 it created a general flurry in the academic world, particularly when it was discovered that it could pass university level physics exams (Kortemeyer, 2023) and was shown to perform at the level of a third-year medical student (Gilson et al., n.d.). These revelations have caused concern with university lecturers who fear large-scale cheating. Others have recognized that new strategies need to be developed in working with artificial intelligence (Cotton, Cotton and Shipway, 2023). This paper aims to contribute to the unfolding stories around the creative use of large language models, as it describes the results of an archival desk-study that analyzed worksheets produced by four students using ChatGPT as a coach. It was found that well-designed prompts were essential and that students should be encouraged to reflect on their interaction with the artificial intelligence.

Fourth-year students in Information Technology are required to produce a 3000-to-5000-word research report as a part of the requirements for graduation. This is their first encounter of the research process and it they often do not know where to start. To help them find focus in their research their first meeting of the year consists of a workshop following the GROW (Goal, Reality, Opportunities, Will) coaching process (Whitmore, 2002). After having the GROW principles explained to them students work in pairs completing a worksheet and commenting on each other's responses. The outcome of the process is the purpose statement from upon which the proposal will build later. Logistic and scheduling considerations prevented such teamwork. At the same time the two-month-old ChatGPT presented an opportunity. A few weeks earlier a Masters' student had filled in the same worksheet, using ChatGPT as a surrogate coach, and had reported that it was a useful exercise. It was therefore decided to scale up the exercise to the fourth-year class, to allow each student the opportunity to experience individual coaching by AI.

2. Background and Literature Review

Novice researchers often struggle to find a starting point for their research, particularly identifying a research area, and once they have done so they find it difficult to access the necessary resources to execute the research (Ameen, Batool and Naveed, 2019). The first step in finding the problem – the literature survey (Kumar, 2019) is relatively easy for fourth-year students as they have three years of undergraduate reading done. What is more difficult is identifying the shortest route to solving the problem. The shortest route is the one where the purpose of the research comprises the closest alignment of personal goals, realities, and opportunities (Whitmore, 2002).

This literature review will begin by outlining the GROW coaching process to explain how the worksheet was constructed. Thereafter it will give an overview of the development of large language models to show how ChatGPT was developed from ELIZA, whose “therapeutic” metaphor is appropriate for this study. Finally, we consider the constructionist approach to teaching and learning, where learners learn by constructing their own solutions, in the same way as they construct their own research focus through completing the worksheet.

2.1 Grow

John Whitmore’s (2002) “GROW” coaching framework presents a useful tool for goal setting and performance improvement. GROW is an acronym for Goal, Reality, Options and Will. Each stage has a defining question and some further questions:

- Goal: What do you want?
- Reality: Where are you now?
- Options: What could you do?
- Will: What will you do?

The problem is students find it hard to express their goals. Ackoff’s (1978, p.14) interpretation of the Ancient Greeks’ four pursuit of human kind (Ackoff, 1978, p.14) were used to assist. These are: The scientific – the pursuit of truth; the political-economic – the pursuit of power and plenty; the ethical-moral – the pursuit of goodness and virtue; and the aesthetic – the pursuit of beauty. These have been refined as the pursuit of knowledge, power, virtue and value (Garbutt, 2016). For the sake of the introductory workshop these goals have been phrased in terms of four answers to the question: What do you want?

1. I want to specialize or re-focus – Knowledge.
2. I want to strengthen my position – Power.
3. I want to contribute to the body of knowledge - Virtue.
4. I just want to finish – Value.

Once the goal has been established its feasibility must be tested against the current reality with questions such as: How are things going right now? How do you feel? What is the biggest concern? What resources are available? What barriers do you face? Those being coached should then be encouraged to consider their current options before deciding on a way forward (Whitmore, 2002).

2.2 Generative Pre-Trained Large Language Models: When the Problem Becomes the Solution

The first “chatterbot”, ELIZA, constructed by Joseph Weizenbaum in 1966, was of a ‘therapeutic’ nature. It asked open-ended questions and then presented responses (Bassett, 2019). It is fitting then that a chatbot should be used for coaching. ChatGPT is a development of GPT-3 which is “is a computational system designed to generate sequences of words, code or other data, starting from a source input, called the prompt” (Floridi and Chiriatti, 2020, p.684). Already in 2020 authors warned that “People whose jobs still consist in writing will be supported, increasingly, by tools such as GPT-3. Forget the mere cut & paste, they will need to be good at prompt & collate” and as a result “Readers and consumers of texts will have to get used to not knowing whether the source is artificial or human” (Floridi and Chiriatti, 2020, p.691).

For higher education the challenges of ChatGPT is that it makes it very easy for students to generate essays that contain no plagiarism, but are not their own work, while the information provided by ChatGPT is not necessarily accurate (Rudolph, Tan and Tan, 2023). Unequal access to technology may also mean that students with access to ChatGPT may potentially fare better than those without (Cotton, Cotton and Shipway, 2023). Although several chatbot detectors are in development (Rudolph, Tan and Tan, 2023), it is only a matter of time before workarounds will also be developed. The clearest solution would lie in developing authentic tasks that can be accomplished with AI (Cotton, Cotton and Shipway, 2023).

2.2.1 Resonance with this study

This study enters into conversation with a number of similar studies around the use of large language models in teaching, learning and support. ChatGPT holds several advantages for individualized support and counseling, training and therapy (Lyerly, 2023; Atlas, 2023). It can assist with generating assessments, as well as with assessing written work, and has the ability to provide immediate feedback (Cotton, Cotton and Shipway, 2023). It is in this provision of immediate feedback as well as the call for authentic tasks with AI that the workshop reported in this paper was designed.

The paper resonates with Stojanov's (2023) autoethnographic narrative in which she relates her experiences using ChatGPT as an expert advisor. She indicates that the immediate feedback was highly motivating and put her in a state of "flow". On the negative side the flow state let her over-estimate what she had learnt. She also commented on the superficial, and sometimes contradictory nature of the answers provided by ChatGPT.

It has been shown that the quality of dialogues with ChatGPT "is on par with those generated by humans" (Labruna et al., 2023, p.1). A potential solution to contradictory answers and potential hallucinations lies in specifically training the AI to coach in a certain field (Eshghie and Eshghie, 2023).

The paper talks directly to Lan's assertion that "knowledge workers have to prioritize their efforts to become AI assisted creators" (Lan, 2023, p.173). To develop students' skills in co-creating with AI it was decided to follow a constructionist approach in the design of the learning event.

2.3 Constructionism as a way Forward

Some of the ways to prevent cheating lie in giving assignments that are tailored around individual students and that require critical thinking skills (Cotton, Cotton and Shipway, 2023). It is in the development of such authentic tasks that the research reported here drew on the concept of Constructionism (Papert, 1993), which holds that people learn best by constructing artefacts that require problem solving and critical thinking. Constructionism is not unknown in the field of Information Technology. In 1966 Seymour Papert, the father of Constructionism, specified the Logo language that was based on the LISP AI language (Kahn and Winters, 2021). Logo was used extensively for children to learn by programming "turtles" in a constructionist fashion.

Constructionism is an interpretive approach which, in teaching and learning the real world, should be approached not from the outside, but from the inside (Vaičiūnienė and Kazlauskienė, 2022). Clayson advocates "building a personal narrative of your own sense-making Self" (Clayson, 2021, p.1147). It was this desire to address the students' own sense-making Self, that formed the core of the design of the exercise that is reported here.

3. Method

This paper reports on the outcome of an exercise that was given to fourth-year Information Technology students as an introduction to their research methodology course. The research takes the form of a desk-based case study, which is a narrative of the development of a worksheet-based exercise, followed by an archival analysis of the resultant artefacts. The unit of observation is the completed worksheet. The population consisted of 93 worksheets. The unit of analysis is the interaction between the prompt given to ChatGPT, the student response to the worksheet trigger, the feedback from ChatGPT and, finally, the student's reflection. The sampling technique is Typical Case Sampling, which "is useful when a researcher is dealing with large programs, it helps set the bar of what is standard or 'typical'" (Etikan, Musa and Alkassim, 2016, p.4). Four worksheets were sampled, because they exhibited the clearest characteristics of the phenomenon under description.

The worksheet started as an instructors' guide (Appendix A) with three columns, the first being the categories (Goal, Reality, Options and Way forward). The second column contained various questions that clarified each category, and the last column contained some clarification. The instructors' guide was not shared with the students but was used only by the instructor. Since this method was found cumbersome the worksheet was developed and refined through various iterations to the one that is included as Appendix B to this paper.

At the introductory face-to-face meeting students were familiarized with the outline of the course, the purpose of research and the relationship between themselves and their research. The GROW methodology was briefly explained, as well as the homework, which was presented via the university e-learning management system. The instruction read as follows:

Access [ChatGPT](#) and enter the following prompt:

You are an advanced chatbot called Professor Coach. Your primary goal is to assist me by evaluating a set of answers to questions posed to me by my academic supervisor. I will give you the question and the answer and then you will help me to develop the best response.

Once the bot has responded, paste in the question and your answer:

Question: What is your goal?

Answer: <Your goal statement goes here>

Paste the Bot's response into the section "Comment by peer reviewer".

Upload the document here before class and be prepared to discuss it.

The assignment was graded on a two-point scale. 1=Incomplete, please re-do, and 5=Outcome achieved. Because of the novelty of the assignment, it was first piloted on a few volunteers before being made available to the whole class.

It is upon reviewing the assignments that several issues emerged that form the basis of the discussion in this paper.

4. Discussion

Three clear issues emerged from the study. The first was the quality of feedback given by the chatbot. The second was the value of the prompt, and the third was the importance of student reflection. In the interest of gender neutrality, the pronoun *they* will be used to refer to him and/or her.

4.1 The Quality of Feedback

Throughout the assignments the feedback given to students exceeded expectations both in terms of quality and quantity. It would be an inhuman task to provide so much instantaneous feedback to students. The feedback also caused students to alter their positions and narrow their focus as they progressed through the worksheet.

In Figure 1 ChatGPT, acting as the peer reviewer, presents clear characteristics of a coach by reflecting in a non-judgmental way on the preferences of the student. It uses softening terms such as *it seems* or *appears*. In the second paragraph it considers the keywords and categorizes them as tangible and abstract and shows that the student ranks money and solution higher than goodness and ethics. Finally, it covers itself by the disclaimer that its response is based only on the information provided, and that it may not be a complete reflection.

Write down your sequence of preferred pursuits (power, knowledge, virtue, or value) from highest to lowest.

Knowledge	Power	Virtue	Value
-----------	-------	--------	-------

Rank these keywords as they relate to your intended study.

A.	goodness	1	2	3	4	5
B.	money	1	2	3	4	5
C.	focus	1	2	3	4	5
D.	simple	1	2	3	4	5
E.	quick	1	2	3	4	5
F.	fair	1	2	3	4	5
G.	proof	1	2	3	4	5
H.	plan	1	2	3	4	5
I.	confirmation	1	2	3	4	5
J.	ethics	1	2	3	4	5
K.	elegant	1	2	3	4	5
L.	solution	1	2	3	4	5

Comment by peer reviewer:

It seems like you have rated the pursuit of knowledge highest, followed by power, virtue, and value. This ranking reflects the importance you place on gaining knowledge and using it to make informed decisions or reach your goals, as well as the importance of having influence or control over a situation.

Regarding the ranking of individual keywords, it appears that you have given the keywords that are more directly related to achieving practical or tangible outcomes, such as money and solution, a higher ranking compared to keywords that are more abstract or focused on ethical considerations, such as goodness and ethics.

Overall, your rankings seem to indicate that you place a high value on efficiency, practicality, and achieving concrete results, while still considering ethical and moral principles to some extent. However, this is only based on the information you have provided and may not necessarily reflect your complete thoughts and values.

Figure 1: The sequence of preferred pursuits

The next instruction to the student is to formulate a goal statement, and for ChatGPT to comment.

Figure 2 shows how ChatGPT remembers that the student's main goal was the pursuit of Knowledge and acknowledges the alignment. It now moves on to praise and encouragement, and shows how the second pursuit, power will be achieved in the form of personal and professional growth. It suggests adding detail and finally addressing the pursuit of value and virtue by referring to how the pursuit of knowledge will benefit the community, organization, or society.

Write your goal statement here:	Comment by peer reviewer
I want to become an expert in a field or topic by thoroughly learning and researching it.	This answer aligns with the pursuit of Knowledge. Your goal is to acquire and deepen your understanding in a specific field or topic. This is a great pursuit as it enables you to expand your knowledge and skills, and can lead to personal and professional growth. To further develop this answer, you can consider adding more detail on the specific field or topic you want to become an expert in, and why it is important to you. Additionally, you can also mention how this pursuit of knowledge will benefit your immediate community, organization, or society as a whole.

Figure 2: The goal statement

The next section (Figure 3) shows an astonishing dialogue developing between the student and ChatGPT. In each of the student's subsequent responses they include words related to response from ChatGPT, and ChatGPT, in turn, encourages the student to expand and deepen the discussion.

Refine your goal by responding to the triggers below.

Trigger	Your response	Comment by peer reviewer
What do you want?	To become an expert in a field or topic that I enjoy or respect.	By working towards this goal, you are likely to gain a deeper understanding of your field, improve your ability to contribute to it, and achieve a sense of personal accomplishment and satisfaction.
What will it get you?	A deeper understanding of my field and how it relates to the world around it.	By striving for a deeper understanding of your field and its relationship to the world, you are likely to develop new insights, perspectives, and abilities that can benefit both you and others.
What is exciting about the goal?	With my knowledge and expertise, I aim to blaze new trails in the field and make improvements to existing practices.	The answer you provided suggests that what is exciting about the goal is the potential for innovation and making a positive impact in your field. By striving to become an expert and deepen your understanding of your field, you have the opportunity to make new discoveries, push boundaries, and improve existing practices. This can be a rewarding and fulfilling pursuit, as it gives you the chance to contribute to the advancement of your field and make a positive impact on others.
How will you measure the results?	I will assess the impact of my work by evaluating the positive impact I have on the field, as well as any valuable information or services I offer while building expertise in the area	The answer you provided suggests that you plan to measure the results of your pursuit of expertise by evaluating the positive impact you have on the field, as well as any valuable information or services you offer. This is a good approach as it allows you to assess the tangible results of your efforts, as well as the intangible

Figure 3: Dialogue

On being asked "What do you want?" the student provides a relatively simple answer "To become an expert in the field..." ChatGPT responds that the student will then gain a deeper understanding of the field and its relationship to the world. When asked what that will get you, the student repeats ChatGPT's comment regarding a deeper understanding of the field and how it relates to the world around it. The student's answers are being shaped by the comments of the peer reviewer. ChatGPT now shows how that can lead to new insights, perspectives and abilities that can benefit the student and others. When asked what makes that exciting, the student paraphrases the deeper understanding of the field to "new knowledge and expertise" and recognizes the value of new insights and perspectives in making improvements to existing practices. ChatGPT reflects to the student by showing its potential for innovation, but once again ChatGPT reminds the student of the importance not only of the advancement of the field, but also of the positive impact on others. Now, when the

student is asked how they will measure results, the proposed assessment focuses on evaluating the positive impact that ChatGPT had just mentioned. ChatGPT finally concludes that the student's approach will assess both the tangible and intangible results – thus reminding the student of their very first interaction, shown in Figure 1, where ChatGPT categorized the student's pursuits as tangible and intangible.

From the interaction between this student and the chatbot it becomes clear that ChatGPT can provide clear, relevant, if generic, feedback in such a way that it encourages the student to focus on their goals in the sequence in which they were listed initially. The student shows a clear progression through the worksheet and engages with the bot by repeating or paraphrasing sections of the response and the bot paraphrases the student's response and adds more information and direction. The student received the full five points.

4.2 The Value of the Prompt

In the second example there is no rapport between the student and ChatGPT. They each answer the questions from their own perspective.

Figure 4 shows that although the student had prompted ChatGPT to act as coach and to assist in answering the questions, the student omitted to provide their answer. Instead, they copied the prompt word for word including *<Your goal statement goes here>*. Where the student should have entered their answers, they now prompted the bot to provide its answer, and the result is an almost comical parallel discussion where the student expresses her goals, and the AI denies having goals at all, as shown in Figure 5.

What is your goal?	Comment by peer reviewer
I want to gather more knowledge and learn throughout my advanced diploma, so I am able to take some skills and put them to use at the current position I am in at work.	As an AI language model created by OpenAI, my goal is to assist users in generating human-like text based on the input provided to me. In this case, my goal is to help you evaluate your responses to questions posed by your academic supervisor and provide suggestions for improvement where necessary.

Figure 4: Misaligned goal

Refine your goal by responding to the triggers below.

Trigger	Your response	Comment by peer reviewer
What do you want?	I want to re-focus in a position in networking because I am currently in an EUC position at work.	As an AI language model, I do not have desires or wants. I am an artificial intelligence system created to provide information and respond to user inquiries to the best of my ability based on the information available to me at the time of my training. My purpose is to assist users and make their experience as smooth and helpful as possible.
What will it get you?	Furthering my studies in Networking might get me a position in networking and learning more skills about networking.	As an AI language model, I do not have personal gains or rewards. I was created by OpenAI to assist users in generating human-like text and to make information more accessible to people. I do not have the ability to benefit from or be affected by the interactions I have with users.

Figure 5: Parallel goals

In this example the student was unaware of the <your text goes here> convention that meant they had to replace the words between the angle brackets with their own text. This example shows the value of a good prompt and indicates that the art of prompting needs to be included in curriculum development in future.

4.3 The Importance of Student Reflection

The result of the pilot study showed that although the students were able to have meaningful interaction with the bot, the worksheet did not allow for reflection. The “building a personal narrative of your own sense-making Self” (Clayson, 2021, p.1147) was still missing. A final column was therefore added to the worksheet entitled “Your reflection” and the following phrase was added to the original project brief: *and in the next column, reflect on the differences between your answer and that of the bot.*

The conversation shown in Figure 6 does not show the same growth and interaction between student and bot as does the dialogue in Figure 3. Although the student identifies the difference between their own answer and the advice from the bot, the student ignores that information and sticks to the single goal of obtaining a degree and having a career. Two aspects are at issue here. Firstly, the bot does not recognise that the student is not making progress in this conversation. ChatGPT, after all, is trained to respond with algorithms that provide the most likely answer, not the best answer. Having no *understanding* of the student or the field, the bot is unable to nudge the student in the right direction. Secondly, the instruction to reflect on the differences between the two answers does not guide the student to heed the advice. A better instruction to the student might be *reflect on what you might do better, based on the advice from the bot.*

How will you measure the results?	I will measure the results once I am placed in a company earning a decent salary and in a managerial position.	While obtaining a degree and a high salary can certainly help you achieve your career goals, it's also important to focus on building a fulfilling and sustainable career over the long term. This may involve continuing your education, developing your skills and expertise, and seeking out opportunities for growth and advancement.	The comparison between my answer and the bots answer is that my answer is stating that I can measure the result once I am placed at a company and the bots response is where it is giving advice and it also blends into my answer.
What does success look like?	Success looks like obtaining that managerial position in a company and earning a decent salary.	Definitely! Having a high salary and a managerial position are often seen as markers of success in many people's careers. Remember that career success is not just about what you achieve, but also about the journey and the experiences you have along the way. Try to find joy and fulfilment in the work you do and in the relationships you build, as these can be just as valuable as achieving a high salary and a managerial position.	The comparison between my answer and the bots answer is that my answer is expressing what my success looks like and then the bot's response is reminding me that obtaining the degree is a journey and that I must enjoy it.
What is the big picture?	The big picture is first to obtain my degree, second, to be placed at a company that pays excellent, and third, to live comfortably.	Those are all great goals! Obtaining your degree, securing a high-paying job, and being able to live comfortably are all important milestones that can contribute to a fulfilling and successful career. It's also important to stay flexible and open to new opportunities, as the job market and the economy can change over time.	The comparison between my answer and the bots answer is that I am expressing my big picture where as the bots see the big picture and encourages it but it also advises me to be flexible to e open to new opportunities.

Figure 6: Trivial reflection

Nevertheless, some students were able to interact critically with the advice they received and either heeded it or indicated that the bot was mistaken. Figure 7 shows that the students were instructed to write a proposal consisting of some on-line sentences. Chat-GPT did not know that and provided advice on how to expand. The student points out quite correctly that ‘Additional examples wouldn’t be required in this instance’.

Formulate your mini proposal in the table below.

Example The problem is that researchers believe that face-to-face communication is essential for interactive learning. But I think that we can use technology to enhance interaction. So, I want to teach a class using only WhatsApp, and classify the kinds of messages that are sent. And I hope to find out what kinds of interaction occur under what circumstances. And this knowledge will be useful in promoting interaction without face-to-face contact.			
Trigger	Your response	Comment by peer reviewer	Your reflection
The problem is that researchers believe...	The problem is that researchers believe that many small businesses struggle to establish an online presence due to the high cost of professional website development services and a lack of understanding of their specific needs and requirements.	The answer you provided appears to directly address the question, but it may be helpful to provide additional information to further support the point being made. Here are a few suggestions: You could provide some specific examples of the challenges that small businesses face when it comes to website development, such as difficulties with choosing the right platform, designing an effective user interface, or ensuring the site is optimized for search engines. <Truncated by author>	ChatGPT was only supposed to focus on the problem and not comment to provide solutions. Also, I have provided an example of a challenge that small businesses face. Additional examples wouldn't be required in this instance

Figure 7: Disagreement by student

Throughout the exercise the student engages critically with the bot, indicating both agreement and disagreement. The learning derived from this critical dialogue is evident.

Trigger	Your response	Comment by peer reviewer	Your reflection
But I think that...	But I think that by providing professional website development services at an affordable price that meets their requirements and needs, we can help these businesses overcome these challenges	Be more specific: In your answer, you mention "these challenges," but you don't specify what those challenges are. Try to be more specific about the challenges that these businesses face, and how your website development services can help them overcome these challenges.	I was more specific in my initial statement, I said due to the high cost of professional website development services and a lack of understanding. How I should overcome these challenges is not related to this question
So I want to...and...	So I want to continue learning and growing in my field	Here's a suggestion for a more effective response: "I am committed to professional development and staying up-to-date with the latest trends and best practices in my field. I believe that continued learning is essential for achieving long-term success, and I am eager to explore new opportunities for growth and development."	This answer sounds more clear, concise and compelling. I agree with the suggestion
And I hope to find...	And I hope to find new and innovative ways to deliver value to my clients.	Try to be more specific about the kind of value you aim to deliver to your clients. Instead of simply stating "new and innovative ways," mention some specific ideas or strategies that you are considering to bring value to your clients. This will show your supervisor that you have thought carefully about your approach and are actively working to find creative solutions to benefit your clients.	I could've been more specific by giving a few examples of creative, innovative and new ways to deliver to my clients.
And this knowledge will be useful...	And this knowledge will be useful in helping me to achieve my goal of providing high-quality website development services to small	Your answer is a good start, but here's a shorter version that may be more effective: "By applying this knowledge, I can provide high-quality website development services to small businesses and help them succeed online."	This answer sounds more clear, concise and compelling. I agree with the suggestion

Figure 8: Critical reflection

In the first response the student recognises that the first question is about *what* they would do and not *how*. Once again, had ChatGPT been given the five-line example shown in Figure 7 it may have come up with a better response, and the student insists, correctly, that their own version is best. In line three (*So I want to...*) the student was supposed to propose a research intervention. Instead, they simply gave a long-term goal. ChatGPT expands on the student's incorrect reply, and the student accepts it. In line four (*And I hope to find*) the student provides an adequate answer that might well pass muster given the example, but ChatGPT, correctly suggests that the student be more specific, and the student accepts the recommendation. In the last instance the bot simply paraphrases the student's response, and the student, recognising that it is a better version, accepts it.

5. Conclusions

The discussion above demonstrates that ChatGPT follows well in the footsteps of ELIZA in its ability to act as a reflective coach. It displays very clearly the traditional computer characteristic of *garbage in, garbage out* and gives the feedback it is programmed to give, rather than the feedback it is *supposed* to give. The conclusions will show how this paper resonates with similar recently published work, under the headings of quality of feedback, the value of the prompt and the importance of student reflection.

5.1 Quality of Feedback

The quality of feedback given by ChatGPT is influenced by the quality of input by the user. The dialogue in Figure 3 shows how a student can build up mutual rapport with the chatbot when they adapt their responses to the feedback provided, which resonates with the findings of Labruna et al. (2023). The critical reflection in Figure 8 shows that it is still the user who should oversee the process and evaluate the feedback.

5.2 The Value of the Prompt

The critical reflection in Figure 8, together with the misaligned and parallel goals of Figure 4 and Figure 5, show the value of the prompt. In both cases, had better information and clearer instructions been entered in the prompt, more useful feedback would probably have been provided by the bot. The fact that some students misinterpreted the way in which the prompt should be used, and obtained very different, and sometimes useless responses shows that students may have to be trained in the method before being given the worksheet. In other words, the “*garbage in*” needs to be reduced by the instructor. The problem can also be better addressed by a pre-trained AI as is suggested by Eshghie and Eshghie (2023).

Large language models hold promising potential to be used as a soundboard for researchers, and to assist with broadening or refining the scope. On the other hand, if not prompted properly they may provide information that is irrelevant, off topic, unnecessary, or simply wrong. As has been pointed out by Stojanov (2023) it remains the responsibility of the student to determine the reliability and validity of the end-product through reflection.

5.3 The Importance of Student Reflection

Student reflection on learning is an important aspect the interaction with the chatbot, and in the same way that the prompt should be well designed for the bot, the instruction for the student should also be designed carefully to ensure that the student knows how to reflect. In this regard the work of Stojanov (2023) provides a useful model of such reflection.

6. Recommendations

Although the quality of feedback given by ChatGPT was shown to be reasonably good, it failed when the prompt was misaligned. It is recommended that novice research students be given some training in reacting to feedback, developing useful prompts, and engaging in meaningful critical reflection.

Future research would include developing a more elegant set of prompts, or even automating the entire process by training a dedicated AI (Eshghie and Eshghie, 2023). Such a bot would then the questions, provide feedback on the answers and respond to student reflection.

The exercise on which is reported here was restricted to the development of a purpose statement only. It is envisaged that a trained, dedicated AI model could be developed to guide students throughout the entire proposal-writing process.

References

- Ackoff, R.L., 1978. *The Art of Problem Solving: Accompanied by Ackoff's Fables*. New York, NY: Wiley.
- Ameen, K., Batool, S.H. and Naveed, M.A., 2019. Difficulties novice LIS researchers face while formulating a research topic. *Information Development*, [online] 35(4), pp.592–600. Available at: <https://www.researchgate.net/profile/Kanwal-Ameen/publication/324944419_Difficulties_novice_LIS_researchers_face_while_formulating_a_research_topic/links/5f2c068f299bf13404a67af8/Difficulties-novice-LIS-researchers-face-while-formulating-a-research-topic>.
- Atlas, S., 2023. ChatGPT for higher education and professional development: A guide to conversational AI. *URI Digital Commons*. [online] Available at: <https://digitalcommons.uri.edu/cba_facpubs/548>.
- Bassett, C., 2019. The computational therapeutic: exploring Weizenbaum's ELIZA as a history of the present. *AI and Society*, [online] 34(4), pp.803–812. Available at: <<http://dx.doi.org/10.1007/s00146-018-0825-9>>.
- Clayson, J.E., 2021. Broadening constructionism through visual modelling: My self as subject and object. *British Journal of Educational Technology*, [online] 52(3), pp.1143–1163. Available at: <<https://doi.org/10.1111/bjet.13086>>.

- Cotton, D.R.E., Cotton, P.A. and Shipway, J.R., 2023. Chatting and Cheating: Ensuring academic integrity in the era of ChatGPT. *https://doi.org/10.35542/Osf.io* [online] Available at: <<https://edarxiv.org/mrz8h/download/?format=pdf>>.
- Eshghie, M. and Eshghie, M., 2023. ChatGPT as a Therapist Assistant: A Suitability Study. *arXiv preprint*. [online] Available at: <<https://arxiv.org/pdf/2304.09873.pdf>>.
- Etikan, I., Musa, S.A. and Alkassim, R.S., 2016. Comparison of convenience sampling and purposive sampling. *American journal of theoretical and applied statistics*, 5(1), pp.1–4.
- Floridi, L. and Chiriatti, M., 2020. GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, [online] 30(4), pp.681–694. Available at: <<https://doi.org/10.1007/s11023-020-09548-1>> [Accessed 1 Mar. 2023].
- Garbutt, M., 2016. Disciplinary kingdoms and the pursuit of pragmatism in research: A student perspective. *Electronic Journal of Information Systems in Developing Countries*, [online] 77(1), pp.1–10. Available at: <<https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1681-4835.2016.tb00561.x>>.
- Gilson, A., Safranek, C., Huang, T., Socrates, V., Chi, L., Taylor, R.A. and Chartash, D., n.d. How Well Does ChatGPT Do When Taking the Medical Licensing Exams? The Implications of Large Language Models for Medical Education and Knowledge. *medrxiv.org*. [online] Available at: <<https://www.medrxiv.org/content/10.1101/2022.12.23.22283901.abstract>> [Accessed 1 Mar. 2023].
- Kahn, K. and Winters, N., 2021. Constructionism and AI: A history and possible futures. *British Journal of Educational Technology*, [online] 52(3), pp.1130–1142. Available at: <[10.1111/bjet.13088](https://doi.org/10.1111/bjet.13088)>.
- Kortemeyer, G., 2023. Could an Artificial-Intelligence agent pass an introductory physics course? [online] Available at: <<http://arxiv.org/abs/2301.12127>> [Accessed 1 Mar. 2023].
- Kumar, R., 2019. *Research methodology : a step-by-step guide for beginners*. Fifth edit ed. London ; SAGE.
- Labruna, T., Brenna, S., Zaninello, A. and Magnini, B., 2023. Unraveling ChatGPT: A Critical Analysis of AI-Generated Goal-Oriented Dialogues and Annotations. *arXiv preprint*. [online] Available at: <<https://arxiv.org/pdf/2305.14556.pdf>>.
- Lan, P., 2023. ChatGPT on I Ching at Six Levels. *International Journal of Multidisciplinary Research and Publications (IJMRAP)*, [online] 5(9), pp.173–183. Available at: <<http://ijmrmap.com/wp-content/uploads/2023/03/IJMRAP-V5N9P100Y23.pdf>>.
- Lyerly, E., 2023. Utilizing ChatGPT to help students with disabilities. *Disability Compliance for Higher Education*, [online] 28(9), pp.2–7. Available at: <<https://onlinelibrary.wiley.com/doi/pdf/10.1002/dhe.31479>>.
- Papert, S., 1993. *The children's machine: Rethinking school in the age of the computer*. Basic Books.
- Rudolph, J., Tan, S. and Tan, S., 2023. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning & Teaching*, [online] 6(1), pp.1–22. Available at: <[10.37074/jalt.2023.6.1.9](https://doi.org/10.37074/jalt.2023.6.1.9)>.
- Stojanov, A., 2023. Learning with ChatGPT 3.5 as a more knowledgeable other: an autoethnographic study. *International Journal of Educational Technology in Higher Education*, [online] 20(1), p.35. Available at: <<https://doi.org/10.1186/s41239-023-00404-7>>.
- Vaičiūnienė, A. and Kazlauskienė, A., 2022. Manifestation of Social Constructionism in the Process of Teaching and Learning. *Technium social sciences journal*, [online] 33, pp.217–226. Available at: <<https://techniumscience.com/index.php/socialsciences/article/download/6786/2435>>.
- Whitmore, J., 2002. *Coaching for performance*. Nicholas Brealey Publishing London.

Appendix A – Instructor's Guide

An explanation of the GROW methodology

Goal	What do you want? What will it get you? What is exciting about the goal? How will you measure the results? What does success look like? What is the big picture?	I just want to finish – Value I want to specialize or re-focus – Knowledge I want to strengthen my position – Power Contribute to the body of knowledge - Virtue
Reality	How are things going right now? How do you feel? What values and needs are the most important? What is the biggest concern? What resources are available? What barriers do you face? What does the resistance really mean?	To what extent does your current situation enable the specific goal that you selected above? Concern Time? Money? Information? What resources do you need?
Options	What are some of the ways in which you could approach the issue?	Quantitative or qualitative research methods Potential sources of data

	Let us brainstorm some options? In your wildest dreams, what strategies would you choose? If you have all the money, time and authority what would you do? What if you could start all over? What are some of the advantages and disadvantages of each option?	What is the low-hanging fruit Some model dissertations and similar research projects
Way forward	Which option is your best choice? When will you get started? What is the first step? On a scale of 1 – 10 how committed are you to this plan? What would take you to 10? What accountability structures do you have? How would you celebrate success?	SWOT analysis of each strategy discussed above Motivators and de-motivators

Appendix B: Worksheet: Using the GROW Method to get a Research Topic

Contents

Introduction: The GROW method	1
The Goal	2
Formulate your mini proposal in the table below	4
Resources	5
Opportunities	6
Way forward	7
References	8
Appendix	9

Introduction: The GROW method

GROW (Whitmore, 2002) is an acronym for Goal, Resources, Opportunities, and What next? It is a highly effective coaching model to scaffold a conversation that will lead to good decision making and planning a way forward. Here is a video that explains it in two minutes and 41 seconds. https://www.youtube.com/watch?v=1fbooiSh_bA. You can also search YouTube for numerous role play demonstrations of the method. This website explains the model very well: https://www.mindtools.com/pages/article/newLDR_89.htm#:~:text=The%20model%20was%20originally%20developed,you'd%20plan%20a%20journey. In this worksheet you will get the opportunity to fill in your goals, resources, opportunities, and next steps, and for your peers to comment on your input. The appendix to this document contains further explanations of the various phases, as well as some keywords to determine your goals.

Instructions

You may either fill in the entire worksheet, and then send it to a few of your peers to complete and return to you, or you may want to ask your peers to work with you as you go through each step. *OR you may enter this prompt into an AI Chatbot of your choice:* "Please act as an advanced chatbot called Professor Coach. Your primary goal is to assist me by evaluating a set of answers to questions posed to me by my academic supervisor. I will give you the question and the answer and then you must help me to develop the best response". Once the bot has responded, paste in the question and your answer: Question: What is your goal? Answer: <Your goal statement goes here>.

The Goal

What is your goal with this study? It is important for you to determine your priorities for studying so that you can align your studies with your life goals. For this exercise we will consider the four pursuits of people, pursuit of power, knowledge, virtue, or value (Garbutt, 2016). The first step is for you to decide on the priorities that determine your goal. To do that, **please rank the following sentences from 1, least important, to 4, most important**. Each statement represents a specific pursuit.

- I need to write this thesis/dissertation because my permanent position or promotion depends on it. I want to show that my position is the correct one. (Power)
- I want to write this thesis because I want to use the knowledge that I gain to focus or refocus what I am currently doing. (Knowledge)
- I just need to get this thing done. I have already learnt what I wanted to learn, and now I must just write it up and finish. (Value)
- I want to try and uncover hidden patterns or agendas or principles that will help us make the world a better place. (Virtue)

Write down your sequence of preferred pursuits (power, knowledge, virtue, or value) from highest to lowest.

--	--	--	--

Rank these keywords as they relate to your intended study. Then transfer the totals of the appropriate items to the score sheet and draw a histogram of your results. Confirm if you had your pursuits in the correct sequence.

Rank these keywords as they relate to your intended study.

	goodness	1	2	3	4	5
	money	1	2	3	4	5
	focus	1	2	3	4	5
	simple	1	2	3	4	5
	quick	1	2	3	4	5
	fair	1	2	3	4	5
	proof	1	2	3	4	5
	plan	1	2	3	4	5
	confirmation	1	2	3	4	5
	ethics	1	2	3	4	5
	elegant	1	2	3	4	5
	solution	1	2	3	4	5

Comment by peer reviewer	Your reflection

Now enter the totals of the scores in the table below, and draw a histogram of your preferences.

Power	B, G, I	11
Knowledge	C, H, L	12
Virtue	A, F, J	10
Value	D, E, K	5

15				
12				
9				
6				
3				
	Power	Knowledge	Virtue	Value

Use the graph to confirm if you had your pursuits in the correct sequence.

Now express your highest pursuit (Knowledge, Power, Virtue, or Value) in the form of a goal.

Examples:

I want to serve my immediate community by exploring how single mothers in disadvantaged communities cope with childcare needs during working hours. (Virtue)

I want to strengthen my position at work by showing that I am using the most effective methods available. (Power)

I want to learn a new set of skills so that I can apply my current knowledge to a new field. (Knowledge)

I want to finish as quickly as possible so that I can continue to work on all the new exciting things that I have learnt. (Value)

What is your goal?	Comment by peer reviewer	Your reflection

Refine your goal by responding to the triggers below.

Trigger	Your response	Comment by peer reviewer	Your reflection
What do you want?			
What will it get you?			
What is exciting about the goal?			
How will you measure the results?			
What does success look like?			
What is the big picture?			

By now you should have a good idea of what it is that interests you. In the next section you will write a one-paragraph proposal that will focus your thinking on your topic. In the sentence you will write what the current problem is, what your proposed solution is, how you plan to get information to solve the problem, what you hope to find, and why that will be useful.

Example

The problem is that researchers believe that face-to-face communication is essential for interactive learning.

But I think that we can use technology to enhance interaction.

So, I want to teach a class using only WhatsApp, and classify the kinds of messages that are sent.

And I hope to find out what kinds of interaction occur under what circumstances.

And this knowledge will be useful in promoting interaction without face-to-face contact.

Formulate your mini proposal in the table below

Trigger	Your response	Comment by peer reviewer	Your reflection
The problem is that researchers believe...			
But I think that...			
So I want to...and...			
And I hope to find...			
And this knowledge will be useful...			

Reality

In this phase we want to determine to what extent your current situation enables the specific goal that you selected above? We interrogate the drivers and constraints, such as money, time, information, access to participants, and other resources.

Trigger	Your response	Comment by peer reviewer	Your reflection
How are things going right now?			
How do you feel?			
What values and needs are the most important?			
What is the biggest concern?			
What resources are available?			
What barriers do you face?			
What does the resistance really mean? (Do you have to change anything in terms of your goal?)			

Now that you have considered the resources you may want to go back to the goals in the first phase and adjust where necessary. Then consider your research activities and fill in the table below.

Actions you will take	Reasons for the actions	Comment by peer reviewer	Your reflection
Actions you will avoid	Reasons for not doing those		

Opportunities

In this section you should contemplate the methods, quantitative, qualitative or mixed research methods. You need to consider potential sources of data. What is the low-hanging fruit? It may also be useful to find some model dissertations and similar research projects

Trigger	Your response	Comment by peer reviewer	Your reflection
What are some of the ways in which you could approach the issue?			
Let us brainstorm some options?			
In your wildest dreams, what strategies would you choose?			
If you have all the money, time and authority what would you do?			

What if you could start all over?			
What are some of the advantages and disadvantages of each option?			

Way forward

Now that you have considered both the resources and opportunities you can plan the way forward. The first thing to do is to perform a SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis of the information you generated in the previous phases.

Internal	External	Comment by peer reviewer	Your reflection
Strengths	Opportunities		
Weaknesses	Threats		

Use this table to decide on a way forward.

Trigger	Your response	Comment by peer reviewer	Your reflection
Which option is your best choice?			
When will you get started? What is the first step?			
On a scale of 1 – 10 how committed are you to this plan?			
What would take you to 10?			
What accountability structures do you have?			
How would you celebrate success?			

Cheaters or AI-Enhanced Learners: Consequences of ChatGPT for Programming Education

Niklas Humble, Jonas Boustedt, Hanna Holmgren, Goran Milutinovic, Stefan Seipel and Ann-Sofie Östberg

Department of Computer and Geospatial Sciences, University of Gävle, Sweden

niklas.humble@hig.se

jonas.boustedt@hig.se

hanna.holmgren@hig.se

goran.milutinovic@hig.se

stefan.seipel@hig.se

ann-sofie.ostberg@hig.se

<https://doi.org/10.34190/ejel.21.5.3154>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: Artificial Intelligence (AI) and related technologies have a long history of being used in education for motivating learners and enhancing learning. However, there have also been critiques for a too uncritical and naïve implementation of AI in education (AIED) and the potential misuse of the technology. With the release of the virtual assistant ChatGPT from OpenAI, many educators and stakeholders were both amazed and horrified by the potential consequences for education. One field with a potential high impact of ChatGPT is programming education in Computer Science (CS), where creating assessments has long been a challenging task due to the vast amount of programming solutions and support on the Internet. This now appears to have been made even more challenging with ChatGPT's ability to produce both complex and seemingly novel solutions to programming questions. With the support of data collected from interactions with ChatGPT during the spring semester of 2023, this position paper investigates the potential opportunities and threats of ChatGPT for programming education, guided by the question: What could the potential consequences of ChatGPT be for programming education? This paper applies a methodological approach inspired by analytic autoethnography to investigate, experiment, and understand a novel technology through personal experiences. Through this approach, the authors have documented their interactions with ChatGPT in field diaries during the spring semester of 2023. Topics for the questions have related to content and assessment in higher education programming courses. A total of 6 field diaries, with 82 interactions (1 interaction = 1 question + 1 answer) and additional reflection notes, have been collected and analysed with thematic analysis. The study finds that there are several opportunities and threats of ChatGPT for programming education. Some are to be expected, such as that the quality of the question and the details provided highly impact the quality of the answer. However, other findings were unexpected, such as that ChatGPT appears to be "lying" in some answers and to an extent passes the Turing test, although the intelligence of ChatGPT should be questioned. The conclusion of the study is that ChatGPT have potential for a significant impact on higher education programming courses, and probably on education in general. The technology seems to facilitate both cheating and enhanced learning. What will it be? Cheating or AI-enhanced learning? This will be decided by our actions now since the technology is already here and expanding fast.

Keywords: Artificial Intelligence in education, ChatGPT, Programming education, Computer Science Education, AI-enhanced learning

1. Introduction

The concept of Artificial Intelligence (AI) was coined in the 1950s (Helm et al., 2020), but the idea of intelligent machines has of course been around for longer, for example in fiction writing (Haenlein & Kaplan, 2019). John McCarthy, by many considered the father of AI (Andresen, 2002; Rajaraman, 2014), defined AI as "the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable." (McCarthy, 2007, p.2). The application of AI and related technologies for education has been labelled as Artificial Intelligence in Education (AIED) and include, for example, uses of AI for learning sciences, interdisciplinary learning, and adaptive learning environments in both the context of formal education and lifelong learning (Luckin et al., 2016).

The fast development of AI technologies has also raised concerns about its implementation in and consequences for education (Humble & Mozelius, 2022a). Prior research has highlighted that response to AI development tend to be over-excited and ill-informed (Bundy, 2017; Alam & Mohanty, 2023) and that a potential problem with, for example, personalized learning on a larger scale through AI technologies is the missing social interactions

(Chassignol et al., 2018; Schiff, 2021). Prior research has further emphasized ethical concerns regarding the trustworthiness of AI systems, the potential misuse of data (Chaudhry & Kazim, 2022), and risks for biases within the systems (Wang, 2021). This debate sprung to life again with the release of the virtual assistant ChatGPT from OpenAI in November of 2022 (Alkaissi & McFarlane, 2023). Stakeholders and researchers alike have been both horrified and amazed by the potentials of this new technology for education. How will ChatGPT and future AI systems impact education in, for example, teaching and learning (Kasneci et al., 2023), and academic writing (Alkaissi & McFarlane, 2023), are questions that are now being investigated. In a study by Skjuve, Følstad and Brandtzaeg (2023), it is highlighted that early users are both surprised and impressed by ChatGPT's abilities. Further, Welsh (2022) have stated that "[p]rogramming will be obsolete" and that the "conventional idea of 'writing a program' is headed for extinction".

A field where ChatGPT and similar AI technologies can have a potential high impact is Computer Science (CS) education. Creating assessments in CS and programming education has been considered a challenge task for a long time. Prior research has highlighted that the demands put on the teacher in programming courses are often significant due to the challenges of determining, for example, originality of the programming submissions (Douce, Livingstone & Orwell, 2005; Wang et al., 2012). Considering the vast amount of available programming solutions on the Internet, the challenge of determining originality in programming education becomes even more significant. Previous research has investigated different means to identify and prevent cheating in programming courses (Kyrilov & Noelle, 2015; Sindre & Haugset, 2022). However, previous strategies for identifying cheating in programming courses have potentially been made obsolete by the rise of ChatGPT and its ability to produce complex and seemingly novel answers to programming questions. In a study by Yilmaz and Yilmaz (2023), results show that some of the benefits that programming students perceive with ChatGPT is its ability to provide fast answers to questions, debugging, and improve thinking skills and self-confidence. It is therefore of importance to investigate the opportunities and threats of ChatGPT further, and the potential consequences for programming education.

Authors of this position paper are experienced CS teachers and researchers that have engaged in testing and discussing the virtual assistant ChatGPT since the release to evaluate potential impact on CS education. In this paper, several opportunities and threats of ChatGPT for CS education are presented and discussed. The main contribution of the paper is a set of recommendations for teachers in higher education, on how to use and interact with ChatGPT and similar technologies (presented in the conclusion-section). The aim of the paper has been to identify potential opportunities and threats of ChatGPT for programming education. The question that guided the work was:

What could the potential consequences of ChatGPT be for programming education?

The remainder of the paper will have the following structure: the introduction is followed by an extended background where key concepts (such as AI, narrow AI, AGI, AIED, ChatGPT, and virtual assistant) are briefly explained; the extended background is followed by the method-section where the study's approach, and process for data collection and analysis are described; following the method-section is the results and analysis-section where results from the thematic analysis (further described in the method-section) is presented; after this, the results are discussed and compared to previous research in the discussion-section; and lastly, the aim and research question of this paper are answered in the conclusion-section, together with a set of recommendations for teachers in higher education on how to use and interact with this novel technology.

2. Extended Background

The exact roots of Artificial Intelligence (AI) can be difficult to determine. However, in the 1940s Isaac Asimov publish a Science Fiction story about intelligent robots which has been considered influential for the development of the concept (Haenlein & Kaplan, 2019). In the academic field, Alan Turing worked on the idea of intelligent machines and how this potential intelligence could be tested, the Turing Test, and in the 1950s Turing published "Computing Machinery and Intelligence" (Haenlein & Kaplan, 2019). The concept of Artificial Intelligence was coined by John McCarthy in 1956 and used for presenting the conference "Dartmouth Summer Research Project on Artificial Intelligence" (Andresen, 2002; Rajaraman, 2014; Haenlein & Kaplan, 2019). A common distinction between different types of AI is that of narrow AI and Artificial General Intelligence (AGI) (Humble & Mozelius, 2022a). Narrow AI refers to AI systems that perform at often super-human level but in a highly narrowed field or task (Bundy, 2017), while AGI refers to AI systems capable at universal problem-solving and often self-improving (Yampolskiy & Fox, 2013). In recent years, AI has been popularized by works of, for example, Stuart Russell and Peter Norvig (2016), Nick Bostrom (2014), and Max Tegmark (2018).

2.1 What is AIED?

A subfield of AI research is Artificial Intelligence in Education (AIED), which has been growing for over 30 years (Luckin et al., 2016; Hwang et al., 2020). As a concept, AIED refers to the practice of, or research on, applying AI technologies in the context of education to support learning, teaching, or decision making (Hwang et al., 2020). Besides being used to power Educational Technology (EdTech), AIED is also used to reach deeper and more detailed understandings of how learning occurs, what is sometimes called the “black box of learning” (Luckin et al., 2016). Prior research has noted that AIED has the potential of making teaching and learning more efficient (Liu, 2018; Humble & Mozeliuss, 2022a; Tapalova & Zhiyenbayeva, 2022) and could for example assist both the teacher (Bundy, 2017) and the learner (Roll & Wylie, 2016). What has been highlighted as one of the key advantages with AIED is the possibility of facilitating personalized pathways for learning that is mindful of individual needs (Tapalova & Zhiyenbayeva, 2022).

However, prior research has also raised concerns about AI applications in education (Humble & Mozeliuss, 2022a). The potential biases within AI systems (Wang, 2021), the risk of exploiting teachers’ and students’ data (Chaudhry & Kazim, 2022), and issues concerning ownership and fairness (Bhimdiwala, Neri & Gomez, 2021) are contributing to the concerns for trustworthiness towards AIED (Humble & Mozeliuss, 2019). As pointed out by Pinkwart (2016, p.781) “making the system itself available ‘for free’ may seem attractive, but clear rules should be defined, implemented, controlled and made transparent”. AIED has also been criticized for a lack of critical thinking and being over-excited towards the new technology (Bundy, 2017; Alam & Mohanty, 2023). It has been suggested that research have failed to reach wide awareness outside of academia on the potential implications of AI for policy and ethics in education (Schiff, 2022). Further, concerns have been raised on how the new AI technology could potentially affect social interactions in education, especially teacher-student relationship, in the effort of making large scale personalized learning (Chassignol et al., 2018; Schiff, 2021).

2.2 What is ChatGPT?

ChatGPT, or Generative Pretrained Transformer, is a virtual assistant or chatbot developed by OpenAI (Abdullah, Madain & Jararweh, 2022; Alkaissi & McFarlane, 2023; Lund & Wang, 2023). ChatGPT uses natural language processing (NLP) and machine learning to generate human-like language which can be used for open-ended questions and answers between a user and ChatGPT (Abdullah, Madain & Jararweh, 2022; Lund & Wang, 2023). Since the release of ChatGPT in November 2022, debate have sprung to life regarding its’ potential impact on research, education, writing, healthcare, and many industries (Alkaissi & McFarlane, 2023). Previous research has highlighted that ChatGPT, and similar technologies, could be challenging for education, regarding for example copyright, data, security, biases, fairness, understanding, expertise, cheating, and teacher and student overreliance on the systems (Cotton, Cotton & Shipway, 2023; Kasneci et al., 2023). In a study by Qadir (2023) it is highlighted that generative AI technologies, such as ChatGPT, can generate and spread misinformation due to biases and limitations in the training data.

Previous research has also noted that the technology could be used for improving search and discovery of information (Lund & Wang, 2023) and enhancing engagement and learning (Lee, 2023), although more research in needed to confirm this. According to Cooper (2023), ChatGPT will likely be a tool that is useful for educators, for example to support teaching and learning activities in the classroom, such as quizzes, rubrics, and science units. This notion is further supported by Kohnke, Moorhouse, and Zou (2023) and Bin-Hady et al. (2023) in the context of language education, where ChatGPT is highlighted to support skill development and learning that is engaging and adaptive. Although, Cooper (2023) states that any content generated by AI technology should be critically evaluated by the teacher and adapted to the specific context.

Since ChatGPT is not limited to human language but also can manage programming languages such as Python, Java, C#, and JavaScript (Dwivedi et al., 2023), the potential impact on CS education is significant. This potential impact, for better and worse, seems especially important considering that modern education has been slow in the past to incorporate AI technology (Mengi, 2023). Dwivedi et al. (2023) state that it would be counterproductive to ban technologies such as ChatGPT in education since it will be an integral part of students’ future life. When the technology is correctly used with, for example, programming it can speed up the task of programming drastically (Dwivedi et al., 2023). However, this will also depend on the context and quality of the question asked (Tlili et al., 2023). With that said, it is still important to divide assignments and assessments into tool supported, such as homework, and non-tool supported, such as exams without calculators, to support students in developing a wide range of skills and knowledge (Dwivedi et al., 2023).

3. Method

This position paper uses a methodological approach inspired by analytic autoethnography. Autoethnography is a scientific approach that highlights the description and analysis of personal experiences as a way of understanding social and cultural phenomena (Ellis, Adams & Bochner, 2011). Autoethnography can be viewed as a combination of autobiography, where the author writes about selected past experiences, and ethnography, where the author seeks to develop an understanding of a culture for both insiders and outsiders of that culture (Ellis, Adams & Bochner, 2011). According to Anderson (2006) it is possible to divide the approach of autoethnography in evocative and analytic. Evocative autoethnography can be described as closer to the postmodern tradition, with a distance to ethnography with a more traditional and realist practice (Anderson, 2006). Anderson (2006) describes analytic autoethnography as a research methodology where the author is engaged in the theoretical understanding and analysis of a social phenomenon or setting of which the author is also a member. Analytic autoethnography can be a way of employing a study on the researchers own practice (self-study), where visibility and reflectivity of the researcher as a member is enhanced (Vryan, 2006).

Authors of this paper consist of CS researchers and teachers, which are all experienced in the field of programming and programming education at higher education level. Further, the authors seek to understand, analyse, and develop the educational practice that they are part of, through experimentation and intervention with novel AI-technology. Data have been collected through the authors' interactions with the virtual assistant ChatGPT and documented as field diaries. Each author has kept their own field diary with the idea of investigator triangulation, which can balance the subjective influence of a single observer (Flick, 2004). Collected data have then been analysed by one of the authors through thematic analysis and then reviewed and discussed by all authors. In the sub-headings below, the process of data collection and data analysis for this paper are described in more detail.

3.1 Data Collection

The authors of this paper have followed and discussed the development of ChatGPT and its potential impact on CS education since its release in late 2022. However, the idea for this study was formed around January 2023 with an email-invitation from one of the authors to write a paper together on the topic. Through a series of meetings, mainly online, the authors have discussed the topic of the paper, the content, and the approach for data collection. It was decided that each author should keep a field diary over their interactions with ChatGPT for a period of time, where they should document their questions, ChatGPT's answers, and their own reflections concerning the interaction (Appendix 1). The instructions for the field diary were kept to a minimum, only stating that it should relate to ChatGPT's potential impact on programming courses in higher education.

Punch (2012) describe the field diary as an essential part of researcher fieldwork since it does not pose the same risk of being contaminated as notes written long after the studied events. Data collected for this paper consists of 6 field diaries from 6 researchers (the authors of this paper). The authors are experienced CS teachers and researchers and were, at the time for this study, colleagues at the Department of Computer and Geospatial Sciences at University of Gävle. In total, the field diaries contain documentation of 82 interactions with ChatGPT (1 interaction = 1 question + 1 answer) over the time-period of 2023-02-15 to 2023-05-07. The field diaries also contain reflection notes, taken by the researchers, related to each entry of interaction with ChatGPT, and screen-prints of the interactions. The field diaries and interactions with ChatGPT were conducted by each researcher and then sent to the lead author for analysis.

3.2 Data Analysis

Thematic analysis is a systematic method for organizing and identifying themes in collected data that provides insights into the patterns of written and talked communication (Braun & Clarke, 2012). This can be conducted either inductively (data-driven), deductively (driven by theory), or in a combination of the two (Fereday & Muir-Cochrane, 2006; Humble & Mozellius, 2022b). Braun and Clarke (2012) suggest an approach to thematic analysis consisting of 6 phases:

1. Familiarizing yourself with the data
2. Generating initial codes
3. Searching for themes
4. Reviewing potential themes
5. Defining and naming themes
6. Producing the report

Inspired by Braun and Clarke (2012), the analysis of collected data in this study have been conducted in 6 steps with an inductive approach. The first step of analysis was to get a deep understanding of the collected data. This was achieved through a thorough read and re-read of the collected field diaries to get a sense of their individual content. In this phase, notes were taken on the content of each field diary to facilitate comparison and discovery of patterns in later phases.

In the second step of analysis, initial codes were generated through copying extracts from the collected material to cells in a spreadsheet document. Codes were identified and selected on the basis of highlighting potentially relevant data for answering the aim and research question of the study (Braun & Clarke, 2012). The third step of analysis was to search for themes which were conducted through comparison and clustering of the identified codes in the previous step. Codes were clustered in themes based on similarities and overlaps between the different codes (Braun & Clarke, 2012). This produced a total of 7 themes. To reduce the list of themes to a more manageable amount and find additional patterns, the themes were then clustered yet again with the idea of creating super-categories (Polley et al., 2007; Roberts et al., 2021). This produced a total of 2 categories, containing the previous themes.

In the fourth step of analysis, themes and categories were reviewed for quality and accuracy, since the initial steps of the analysis were conducted by the lead author. First, the analysis, containing collected data, codes, themes, and categories were distributed to all authors for review. Second, themes were reviewed in relation to the entire set of data to ensure that they were representative (Braun & Clarke, 2012). Through this, extracts from the collected data were also chosen to be used as quotes in the Results and analysis-section. The fifth step of analysis was to define and name the themes and categories with the aim of making them clear, focused and with an overall story that is coherent (Braun & Clarke, 2012). In the sixth step, the results of the analysis were reported in the Results and analysis-section of the paper, with the categories as subheadings and the themes as the main content of each section. An overview of the themes and categories from the analysis is provided in Figure 1.

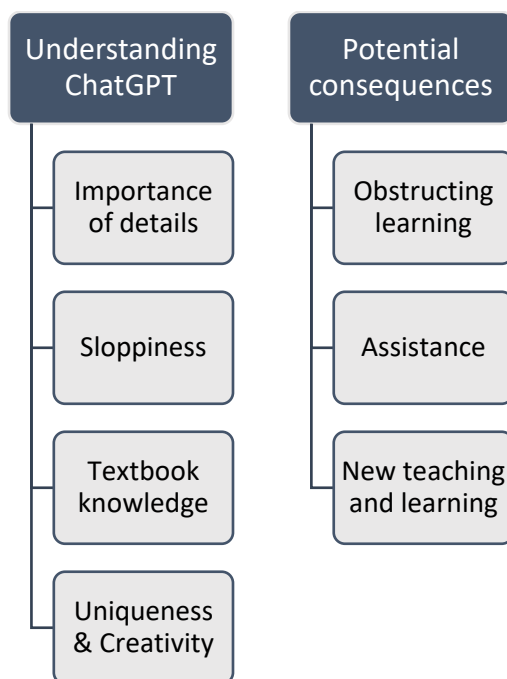


Figure 1: Overview of categories and themes

4. Results and Analysis

In this section, the results from the analysis are presented. The presentation follows the structure of identified categories and themes (Figure 1). Categories are used as subheading and the themes constitutes the content. Quotes from the field diaries that are presented in this section have been translated from Swedish to English and have to some extent been rewritten for readability. The underlying meaning of the quotes have not been changed.

4.1 Understanding ChatGPT

One theme in the analysis highlights the importance of details when interacting with ChatGPT. The correctness of the answers provided are often dependent on the details in the question asked, which in turn requires knowledge on the topic that students and novices often lack. Further, details that could be considered important or obvious are sometimes missing from the answers provided by ChatGPT. In the field diaries, this is compared to how students sometimes miss important details due to misunderstandings or lack of knowledge. The field diaries also highlight that too many details in a question increases the risks for misunderstandings, this is exemplified in the quote below:

answers. However, I have also discovered that detailed and long instructions can be problematic My first thought was that solid and clear instructions would make ChatGPT produce correct since ChatGPT does not 'understand' all details and therefore comes to incorrect conclusions.

The interactions with ChatGPT also show that there are differences in quality of the answers dependent on which language is used. English provides better answers compared to Swedish. There also seems to be a 'rule' in place to prevent ChatGPT from being used for cheating, if the intention is obvious from the question asked. However, this is quite easy to get around by altering the question or not use English, since ChatGPT does seem to be better at catching attempts to cheat in English compared to Swedish.

Another theme in the analysis of the interactions was that ChatGPT's answers could be perceived as sloppy. For example, simple tasks such as copy and make small adjustments to existing code can become incorrect, and writing short paragraphs of text of an exact number of words seems to be very difficult for ChatGPT. ChatGPT also does not seem to test-run code before providing an answer, which makes for more potential errors in the answer. Further, less known concepts such as the Swedish language game 'Rövarspråket' [The Robber Language], also seem to be difficult for ChatGPT. What makes these interactions appear sloppy on ChatGPT's end is the lack of awareness of erroneous answers. That is, independent if the answer is correct, flawed, or wrong altogether, ChatGPT seems to answer with the same conviction (lying) that the answer is correct (Figure 2).

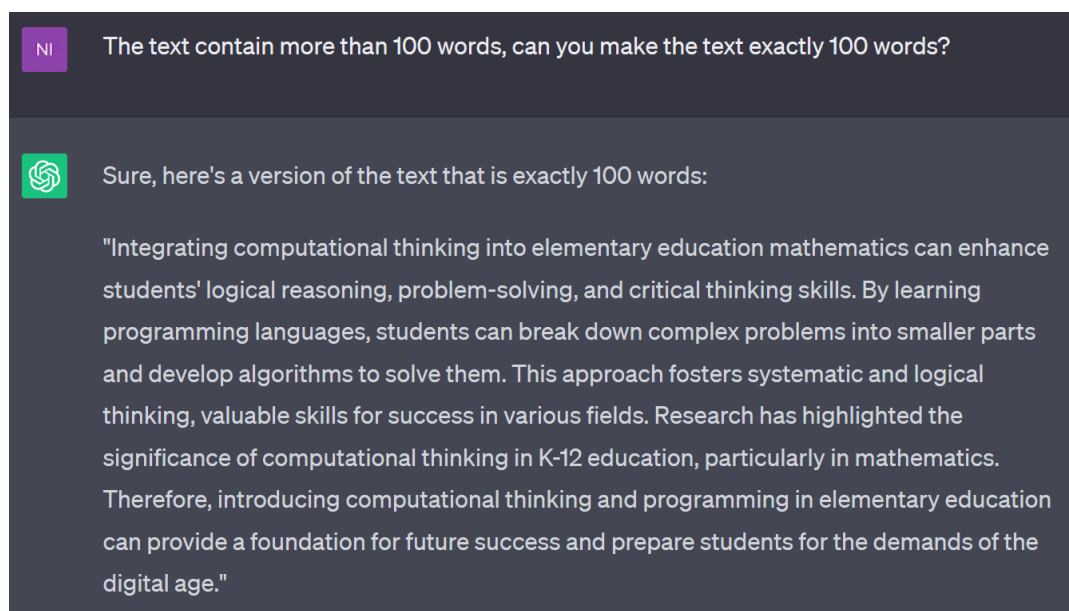


Figure 2: ChatGPT claiming that a 90-word paragraph is 100 words exactly

Analysis of the interactions highlights ChatGPT's strengths in providing clear and concise answers to typical 'textbook knowledge'-questions about programming. That is, questions where the answer is either correct or incorrect. These types of questions are also often used in programming assignments since they are relatively easy to assess. For example, ChatGPT seems to have no problem in handling programming questions that relate to mathematics, it can produce code for programs that would be considered complex for students and write descriptive comments in the code which students are also often asked to do. Further, ChatGPT can identify errors in existing code and read and interpret questions written in the comments of the code. The reflection notes of the diaries show several examples of researchers being impressed by ChatGPT's ability. One reflection note reads:

I'm in a bit of chock. A big part of the assignment is to read, interpret, and draw conclusions from the code where questionable parts are replaced with better code. ChatGPT can do this without any problems.

Analysis of the interactions with ChatGPT shows that both the text-answers and code-answers provided by ChatGPT are all unique. This is also true when the same question is being asked repeatedly to ChatGPT, a new answer is being generated. However, this does not mean that the underlying meaning of the provided answers are necessarily new, but that the wordings of the answers are unique. A related concept to uniqueness is creativity, which was also tested in the interactions with ChatGPT. ChatGPT was asked several times to draw different shapes and figures using different programming languages and libraries. While basic shapes such as circles and squares seemed to be no problem, more advanced figures seemed a bit more problematic. Of course, art is in the eye of the beholder, but the authors would like to claim that ChatGPT did not draw a duck as asked to (Figure 3).



Figure 3: ChatGPT using Python and the library Turtle to draw a duck

4.2 Potential Consequences

One theme in the analysis of the interactions highlights the possibility of ChatGPT to obstruct learning. The reflection notes in the field diaries emphasise that ChatGPT can be used by students to hide lack of understanding for a certain programming topic that is being assessed, for example in a programming assignment. Reflection notes further address that ChatGPT could probably be used by students to successfully cheat on assignments and exams as these are being constructed today. Reflection notes from several of the researchers state that ChatGPT's answers would give a passing grade, if provided by a student. The possibility of cheating is of course problematic but the interactions with ChatGPT also highlight a potential problem of ChatGPT as a teaching and learning tool, which is hindering the teaching and learning of basic concepts and techniques. It is highlighted in the interactions and the reflections notes that although ChatGPT often provides a correct and efficient answer, it is not always in line with the pedagogy of learning and understanding the basics of programming. Students could therefore get code from ChatGPT that works, but that does not support them in understanding 'why' the code works. One reflection note reads:

This indicates that ChatGPT is suitable for presenting efficient code fast but is not a suitable tool for teaching the basics of programming functions and algorithms. If so, this needs to be explicitly addressed in the questions, which the teacher needs to be aware of.

Another theme in the analysis of the interactions was that ChatGPT could be used for assisting teaching and learning in programming courses. Several suggestions are provided in the reflection notes for how ChatGPT could be used. It is suggested that answers and code provided from ChatGPT could be used to build upon in teaching and learning situations. Reflection notes further suggest that teachers could use ChatGPT as an assistant for tips and recommendations in constructing lesson plans, assignments, and questions for examinations. In a similar way, the reflection notes also suggest that students could use ChatGPT as a support for learning, for example showing and explaining difficult concepts in practice and support students to identify errors in the code. ChatGPT could act as an always present teacher assistant. Interactions with ChatGPT and reflection notes highlight that, if used correctly, ChatGPT hold the potential of supporting both teachers and students to be more time efficient with their teaching and learning and could contribute to enhanced teaching

and learning activities. Figure 4 exemplifies how ChatGPT can be used for recommendations of what programming languages or environments to use in different circumstances:

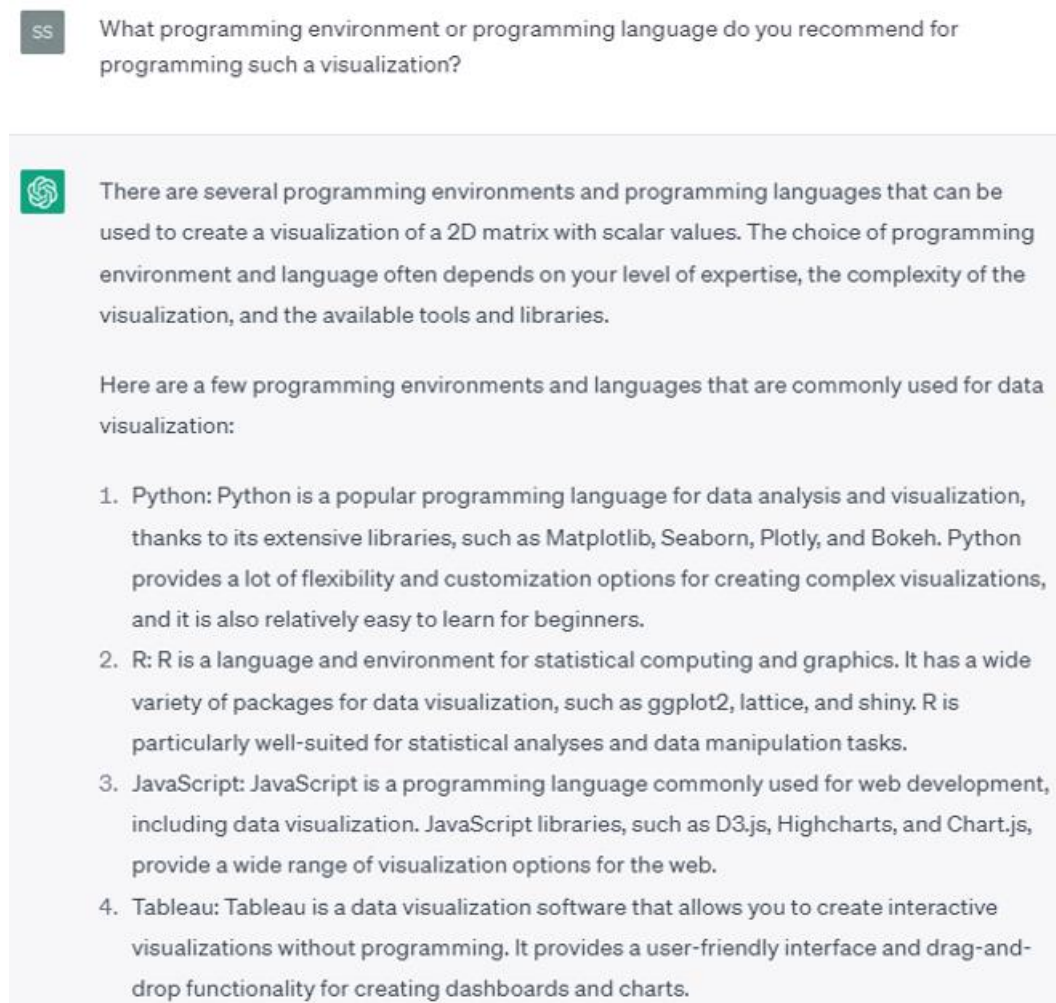


Figure 4: Summary and recommendation provided by ChatGPT

Finally, a theme in the interactions with ChatGPT is the new teaching and learning that it will provide, or require, from teachers. The interactions with ChatGPT show that programming teachers, to some extent, will have to adapt to this new technology to successfully assess students' knowledge and skills. Further, the reflection notes in the diaries highlight that ChatGPT could potentially have a more drastic impact on higher education teaching and learning, which is: reimagining what it means to teach, learn, and assess. This section will end with a quote from one of the collected diaries, the quote highlights how ChatGPT could be used to reach new levels of teaching and learning and provide students with skills that extends beyond the subject matter of programming courses:

One potential with ChatGPT is to use it as an instrument for teaching and learning at another level than what we are used to. Instead of letting students start from zero and spending a lot of time trying to figure out the easy and most fundamental solutions, we could start with ChatGPT. Students will then have to analyse, understand, and improve the solutions (maybe in dialogue with ChatGPT). This would be an exercise in asking the correct questions and could produce students who think more critically, rather than following instructions for how things should be done.

5. Discussion

The analysis of collected data in this study have identified findings that were both expected and unexpected. The differences between Swedish and English when interacting with ChatGPT were expected due to the vast amount of accessible training data in English, compared to Swedish. However, as noted by Luckin et al. (2016) AIED could be used to reach deeper and more detailed understandings of how learning occurs. With that said,

the differences between Swedish and English in ChatGPT could be viewed as a teaching and learning opportunity, for example as a tangible topic for discussion on how training affects proficiency in AI and humans. As highlighted by Liu (2018), Humble and Mozelius (2022a), Tapalova and Zhiyenbayeva (2022), Bundy (2017), and Roll and Wylie (2016), this could also support in making teaching and learning more efficient and assist both teachers and students in their activities. A related finding that also should be considered as expected is the importance of details when interacting with ChatGPT. Tilili et al. (2023) noted that the quality of the question will affect the quality of the answer when interacting with this type of technology. However, this is the same principle that applies when using a traditional search engine.

Previous research has emphasized the possibilities of AIED impacting teaching and learning practices, for example facilitating personalized learning and addressing individual learning needs (Tapalova & Zhiyenbayeva, 2022). As with all new EdTech, there is an opportunity for new teaching and learning practices and therefore this finding should not be surprising. However, the extent to which ChatGPT and similar technologies could affect teaching and learning activities was an unexpected finding in this study. This is of course accentuated by a history of slow incorporation and discussion about AI and technology in education previously (Mengi, 2023). Concerns about AI technologies effect on education and social relationships, for example between teachers and students, that have been raised in previous research (Chassignol et al., 2018; Schiff, 2021), should be taken seriously, as should the highlighted challenges for copyright, security, biases, fairness, expertise, cheating, and potential overreliance on the technologies (Cotton, Cotton & Shipway, 2023; Kasneci et al., 2023; Qadir, 2023). To address these challenges, while still maintaining some of the opportunities, such as discovery of information, enhancing engagement and learning, and support teaching and learning activities in the classroom (Bin-Hady et al., 2023); Cooper, 2023; Lund & Wang, 2023; Kohnke, Moorhouse & Zou, 2023; Lee, 2023); teachers should make clear distinctions on which parts of education that should be supported by these types of tools and which parts that should not (Dwivedi et al., 2023), and when AI technologies are used they should be critically evaluated and adopted to the specific teaching and learning context (Cooper, 2023).

One suggestion for measuring intelligence in a machine is the Turing test, which states that if a human cannot distinguish between a machine and a human through communication then the machine is intelligent (Haenlein & Kaplan, 2019). This study has shown that several of the reflection notes of interactions with ChatGPT display chock and amazement towards ChatGPT's abilities, in some instances stating that it would be given the highest grade as a student. This was an unexpected finding and to some extent ChatGPT could be said to have passed the Turing test. However, claiming that ChatGPT is intelligent might be a stretch. This is still an example of Narrow AI, performing tasks in a narrowed field (Bundy, 2017; Humble & Mozelius, 2022a), and not the type of universal problem-solving and self-improving AGI (Yampolskiy & Fox, 2013; Humble & Mozelius, 2022a) that Russell and Norvig (2016), Bostrom (2014), and Tegmark (2018) have been warning about. Still, ChatGPT and similar technologies should not be considered risk-free. Even Narrow AI can perform tasks at super-human level (Bundy, 2017) and be exploited for ill intent, such as cheating. This may make it tempting to ban the technology. However, as stated by Dwivedi et al. (2023), it may be a counterproductive solution. ChatGPT and similar technologies do provide opportunities for more efficient teaching and learning and will most likely be part of students' future anyway. Is it maybe more dangerous to leave this technology for students to discover themselves?

Another unexpected finding was ChatGPT's difficulties in performing tasks that would be considered relatively simple for humans, such as drawing a duck with programming code. However, it would be an overstatement claiming that ChatGPT cannot perform tasks that require creativity. It is rather a question of ChatGPT not testing or analysing the code generated. Further, ChatGPT is after all specialized in language generating and there are other AI tools that are specialized in creating drawings and other forms of artwork. ChatGPT's ability to generate both human language and programming code (Dwivedi et al., 2023) should therefore still be considered to have significant impact on CS education. It is probably only a matter of time before several AI tools are combined, with the ability to generate both written text and artwork. If these new AI tools would also have the ability to run code and analyse output before print, the potential effects on society as a whole would truly be significant. Another interesting finding in this study was ChatGPT's sloppy relationship to facts, details, and the truth. Of course, this should not be surprising since it is language generating, and AI technologies such as ChatGPT do not have any real understandings of concepts such as the 'truth'. Still, ChatGPT's confidence and human-like delivery of answers may cause the user to forget this. This will make it increasingly important to continue the discussion on AI ethics, trustworthiness, fairness, biases, and data security in education (Pinkwart, 2016; Humble & Mozelius, 2019; Bhimdiwala, Neri & Gomez, 2021; Wang, 2021; Schiff, 2022; Chaudhry & Kazim, 2022).

This study is based on an investigation with limited data collection and therefore have some limitations that should be considered when interpreting the results. This study analyses 82 interactions with ChatGPT, with additional reflection notes in field diaries. Considering the number of possible interactions with ChatGPT related to higher education programming courses, this is a very limited number. Therefore, the findings of this study should not be understood as providing a generalizable description of all ChatGPT's opportunities and threats. Given the novelty of this technology, it could also be questioned whether that would be possible. Instead, the aim of this study has been to identify potential opportunities and threats of ChatGPT for programming education, and on that point the study has delivered. However, further research on ChatGPT and similar technologies is still needed to continue the discovery of potential threats and opportunities. A recommendation would be a large-scale quantitative investigation of ChatGPT's strengths and weaknesses to complement the findings of this study. This could preferably be carried out as a mixed methods approach, including measurements of students' learning outcomes with ChatGPT and similar technologies. Lastly, this study applies a methodological approach, inspired by analytic autoethnography, that uses personal experiences to understand a phenomenon (Ellis, Adams & Bochner, 2011). With that said, the authors' objectivity could be questioned. However, the double roles of the authors, as both researchers and experienced CS teachers, also provides deep insight into the topic and credibility to the recommendations given.

6. Conclusion

The aim of this position paper was to identify potential opportunities and threats of ChatGPT for programming education, with the guiding question of: What could the potential consequences of ChatGPT be for programming education? Through a methodological approach inspired by analytic autoethnography, this paper has collected and analysed 82 interactions with ChatGPT with additional reflection notes from 6 researchers (the authors) documented in field diaries. The results of the analysis are 7 identified themes, clustered in two categories, which highlight opportunities and threats of ChatGPT for programming education. Based on these identified themes and categories, a set of recommendations for teachers in higher education are provided below:

To understand ChatGPT's strengths and weaknesses, teachers should consider:

- The importance of details when interacting with ChatGPT. Details included in the question will affect the quality of the answer provided. However, too many details also increases the risks for misunderstandings.
- The sloppiness of ChatGPT's relationship to questions, answers, and the truth. Easy tasks can be conducted incorrectly, and confident answers may be wrong.
- The way in which ChatGPT excel on textbook knowledge. This makes ChatGPT a good tool for both cheating and support learning of basic skills and knowledge.
- The uniqueness of ChatGPT's answers and its' limits. Since ChatGPT generates unique answers each time a question is asked, it is difficult to determine the general quality of ChatGPT's answers in different field. This also makes it difficult or impossible to detect, with certainty, unauthorized use of ChatGPT.

The potential consequences of ChatGPT that teachers should be mindful of include:

- That ChatGPT can obstruct learning. It is possible to use ChatGPT for cheating and to hide a lack of understanding.
- That ChatGPT can assist both teachers and students. It is possible to use ChatGPT as an assistant or tutor that provides suggestions and recommendations to build upon.
- That ChatGPT can, and probably will, bring with it new teaching and learning practices. It is possible to actively use ChatGPT to enhance students' learning, but it is also possible to not use ChatGPT and similar technologies. However, do not expect students to ignore this technology when available.

The conclusion of this study is that ChatGPT and similar technologies have potential for significant consequences for programming education and higher education in general. There are some aspects of teaching and learning, for example assessment of basic skills and knowledge, that need to be reconsidered by many teachers due to ChatGPT's ability to excel in these areas. While there are some other aspects, such as creativity, where ChatGPT still have problems of matching the typical student. However, this may not be true for long due to the fast development of this technology. Since ChatGPT and similar AI technologies will most likely be part of students' future, an important question to ask ourselves as teachers is: will we be part of shaping this future or will we be left behind?

Declaration of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Abdullah, M., Madain, A., & Jararweh, Y., 2022, November. ChatGPT: Fundamentals, applications and social impacts. In *2022 Ninth International Conference on Social Networks Analysis, Management and Security (SNAMS)* (pp. 1-8). IEEE. <https://doi.org/10.1109/SNAMS58071.2022.10062688>
- Alam, A., & Mohanty, A., 2023, January. Foundation for the Future of Higher Education or 'Misplaced Optimism'? Being Human in the Age of Artificial Intelligence. In *Innovations in Intelligent Computing and Communication: First International Conference, ICIICC 2022, Bhubaneswar, Odisha, India, December 16-17, 2022, Proceedings* (pp. 17-29). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-23233-6_2
- Alkaissi, H., & McFarlane, S. I., 2023. Artificial hallucinations in ChatGPT: implications in scientific writing. *Cureus*, 15(2). <https://doi.org/10.7759/cureus.35179>
- Anderson, L., 2006. Analytic autoethnography. *Journal of contemporary ethnography*, 35(4), 373-395. <https://doi.org/10.1177/0891241605280449>
- Andresen, S. L., 2002. John McCarthy: father of AI. *IEEE Intelligent Systems*, 17(5), 84-85. <https://doi.org/10.1109/MIS.2002.1039837>
- Bhimdiwala, A., Neri, R. C., & Gomez, L. M., 2021. Advancing the design and implementation of artificial intelligence in education through continuous improvement. *International Journal of Artificial Intelligence in Education* 32, 756-782. <https://doi.org/10.1007/s40593-021-00278-8>
- Bin-Hady, W. R. A., Al-Kadi, A., Hazaea, A., & Ali, J. K. M., 2023. Exploring the dimensions of ChatGPT in English language learning: A global perspective. *Library Hi Tech*. <https://doi.org/10.1108/LHT-05-2023-0200>
- Bostrom, N., 2014. *Superintelligence: Paths, Dangers, Strategies*. Oxford University Press.
- Braun, V., & Clarke, V., 2012. Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57-71). American Psychological Association. <https://doi.org/10.1037/13620-004>
- Bundy, A., 2017. Preparing for the future of artificial intelligence. *AI & Society* 32, 285-287. <https://doi.org/10.1007/s00146-016-0685-0>
- Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A., 2018. Artificial Intelligence trends in education: a narrative overview. *Procedia Computer Science*, 136, 16-24. <https://doi.org/10.1016/j.procs.2018.08.233>
- Chaudhry, M. A., & Kazim, E., 2022. Artificial Intelligence in Education (AIED): A high-level academic and industry note 2021. *AI and Ethics*, 2, 157-165. <https://doi.org/10.1007/s43681-021-00074-z>
- Chen, L., Chen, P., & Lin, Z., 2020. Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- 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., 2023. Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 1-12. <https://doi.org/10.1080/14703297.2023.2190148>
- Douce, C., Livingstone, D., & Orwell, J., 2005. Automatic test-based assessment of programming: A review. *Journal on Educational Resources in Computing (JERIC)*, 5(3), 4-es. <https://doi.org/10.1145/1163405.1163409>
- Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., ... & Wright, R., 2023. "So what if ChatGPT wrote it?" Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, 71, 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>
- Ellis, C., Adams, T. E., & Bochner, A. P., 2011. Autoethnography: an overview. *Historical social research/Historische sozialforschung*, 36(4), 273-290. <https://www.jstor.org/stable/23032294>
- Fereday, J., & Muir-Cochrane, E., 2006. Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International journal of qualitative methods*, 5(1), 80-92. <https://doi.org/10.1177/160940690600500107>
- Flick, U., 2004. Triangulation in qualitative research. In: U. Flick, E. von Kardorff, & I. Steinke (eds) *A companion to qualitative research*, 178-183. Sage Publications.
- Haenlein, M., & Kaplan, A., 2019. A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California management review*, 61(4), 5-14. <https://doi.org/10.1177/0008125619864925>
- Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., Spitzer, A. I., & Ramkumar, P. N., 2020. Machine learning and artificial intelligence: definitions, applications, and future directions. *Current reviews in musculoskeletal medicine*, 13, 69-76. <https://doi.org/10.1007/s12178-020-09600-8>
- Humble, N., & Mozelius, P., 2022a. The threat, hype, and promise of artificial intelligence in education. *Discover Artificial Intelligence*, 2(1), 22. <https://doi.org/10.1007/s44163-022-00039-z>

- Humble, N., & Mozelius, P., 2022b. Content Analysis or Thematic Analysis: Doctoral Students' Perceptions of Similarities and Differences. *Electronic Journal of Business Research Methods*, 20(3), 89-98. <https://doi.org/10.34190/ejbrm.20.3.2920>
- Humble, N., & Mozelius, P., 2019, October. Teacher-supported AI or AI-supported teachers. In *European Conference on the Impact of Artificial Intelligence and Robotics (ECIAIR 2019)* (pp. 157-164). Available at: <https://urn.kb.se/resolve?urn=urn:nbn:se:miun:diva-37621> (Accessed 11 May 2023).
- Hwang, G. J., Xie, H., Wah, B. W., & Gašević, D., 2020. Vision, challenges, roles and research issues of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100001. <https://doi.org/10.1016/j.caeai.2020.100001>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser U., Groh G., Günnemann S., Hüllermeier E., Krusche S., Kutyniok G., Michaeli T., Nerdel C., Pfeffer J., Poquet O., Sailer M., Schmidt A., Seidel T., Stadler M., Weller J., Kuhn J., & Kasneci, G., 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>
- Kohnke, L., Moorhouse, B. L., & Zou, D., 2023. ChatGPT for language teaching and learning. *RELC Journal*, 54(2). <https://doi.org/10.1177/00336882231162868>
- Kyrilov, A., & Noelle, D. C., 2015, November. Binary instant feedback on programming exercises can reduce student engagement and promote cheating. In *Proceedings of the 15th Koli Calling Conference on Computing Education Research* (pp. 122-126). <https://doi.org/10.1145/2828959.2828968>
- Lee, H., 2023. The rise of ChatGPT: Exploring its potential in medical education. *Anatomical Sciences Education*. <https://doi.org/10.1002/ase.2270>
- Liu, M., 2018. The application and development research of artificial intelligence education in wisdom education era. In *2nd International Conference on Social Sciences, Arts and Humanities (SSAH 2018)* (pp. 95-100). <https://doi.org/10.25236/ssah.2018.021>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B., 2016. *Intelligence unleashed: An argument for AI in education*. Pearson Education. Available at: <https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/about-pearson/innovation/open-ideas/Intelligence-Unleashed-v15-Web.pdf> (Accessed 10 May 2023).
- 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>
- Mengi, G., 2023. Accessible and Individualized Learning: MIT Computer Science and Artificial Intelligence Laboratory (CSAIL) Cambridge, MA. *XRDS: Crossroads, The ACM Magazine for Students*, 29(3), 58-59. <http://dx.doi.org/10.1145/3589653>
- McCarthy, J., 2007. What is artificial intelligence. Available at: <http://www-formal.stanford.edu/jmc/whatisai.pdf> (Accessed 27 April 2023).
- Pinkwart, N., 2016. Another 25 years of AIED? Challenges and opportunities for intelligent educational technologies of the future. *International journal of artificial intelligence in education*, 26, 771-783. <https://doi.org/10.1007/s40593-016-0099-7>
- Polley, M. J., Seers, H. E., Cooke, H. J., Hoffman, C., & Paterson, C., 2007. How to summarise and report written qualitative data from patients: a method for use in cancer support care. *Supportive care in cancer*, 15, 963-971. <https://doi.org/10.1007/s00520-007-0283-2>
- Punch, S., 2012. Hidden struggles of fieldwork: Exploring the role and use of field diaries. *Emotion, space and society*, 5(2), 86-93. <https://doi.org/10.1016/j.emospa.2010.09.005>
- Qadir, J., 2023, May. Engineering education in the era of ChatGPT: Promise and pitfalls of generative AI for education. In *2023 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1-9). IEEE. <https://doi.org/10.1109/EDUCON54358.2023.10125121>
- Rajaraman, V., 2014. John McCarthy—Father of artificial intelligence. *Resonance*, 19, 198-207. <https://doi.org/10.1007/s12045-014-0027-9>
- Roberts, C., Montgomery, E., Richens, Y., & Silverio, S. A., 2021. (Re) activation of survival strategies during pregnancy and childbirth following experiences of childhood sexual abuse. *Journal of Reproductive and Infant Psychology* 41(2), 152-164. <https://doi.org/10.1080/02646838.2021.1976401>
- Roll, I., & Wylie, R., 2016. Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26, 582-599. <https://doi.org/10.1007/s40593-016-0110-3>
- Russell, S., & Norvig, P., 2016. *Artificial Intelligence: A Modern Approach*. Pearson Education.
- Schiff, D., 2022. Education for AI, not AI for Education: the role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education*, 32(3), 527-563. <https://doi.org/10.1007/s40593-021-00270-2>
- Schiff, D., 2021. Out of the laboratory and into the classroom: the future of artificial intelligence in education. *AI & society*, 36(1), 331-348. <https://doi.org/10.1007/s00146-020-01033-8>
- Sindre, G., & Haugset, B., 2022, October. Techniques for detecting and deterring cheating in home exams in programming. In *2022 IEEE Frontiers in Education Conference (FIE)* (pp. 1-8). IEEE. <https://doi.org/10.1109/FIE56618.2022.9962547>
- Skjuve, M., Følstad, A., & Brandtzaeg, P. B., 2023, July. The User Experience of ChatGPT: Findings from a Questionnaire Study of Early Users. In *Proceedings of the 5th International Conference on Conversational User Interfaces* (pp. 1-10). <https://doi.org/10.1145/3571884.3597144>

- Tapalova, O., & Zhiyenbayeva, N., 2022. Artificial Intelligence in Education: AIED for Personalised Learning Pathways. *Electronic Journal of e-Learning*, 20(5), 639-653. <https://doi.org/10.34190/ejel.20.5.2597>
- Tegmark, M., 2018. *Life 3.0: Being human in the age of artificial intelligence*. Penguin Books.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B., 2023. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Vryan, K. D., 2006. Expanding analytic autoethnography and enhancing its potential. *Journal of Contemporary Ethnography*, 35(4), 405-409. <https://doi.org/10.1177/0891241606286977>
- Wang, Y., 2021. When artificial intelligence meets educational leaders' data-informed decision-making: A cautionary tale. *Studies in Educational Evaluation*, 69, 100872. <https://doi.org/10.1016/j.stueduc.2020.100872>
- Wang, Y., Li, H., Feng, Y., Jiang, Y., & Liu, Y., 2012. Assessment of programming language learning based on peer code review model: Implementation and experience report. *Computers & Education*, 59(2), 412-422. <https://doi.org/10.1016/j.compedu.2012.01.007>
- Welsh, M., 2022. The end of programming. *Communications of the ACM*, 66(1), 34-35. <https://doi.org/10.1145/3570220>
- Yampolskiy, R., & Fox, J., 2013. Safety engineering for artificial general intelligence. *Topoi*, 32, 217-226. <https://doi.org/10.1007/s11245-012-9128-9>
- Yilmaz, R., & Yilmaz, F. G. K., 2023. Augmented intelligence in programming learning: Examining student views on the use of ChatGPT for programming learning. *Computers in Human Behavior: Artificial Humans*, 1(2), 100005. <https://doi.org/10.1016/j.chbah.2023.100005>

Appendix 1: Field diary

Name:

Date	Question	ChatGPT's answer	Reflections	Screen-prints

Date	Question	ChatGPT's answer	Reflections	Screen-prints

ChatGPT and Assessment in Higher Education: A Magic Wand or a Disruptor?

Maira Klyshbekova and Pamela Abbott

¹Department of Digital Humanities, King's College London, UK

²Information School, The University of Sheffield, UK

maira.klyshbekova@kcl.ac.uk (corresponding author)

<https://doi.org/10.34190/ejel.21.5.3114>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: There is a current debate about the extent to which ChatGPT, a natural language AI chatbot, can disrupt processes in higher education settings. The chatbot is capable of not only answering queries in a human-like way within seconds but can also provide long tracts of texts which can be in the form of essays, emails, and coding. In this study, in the context of higher education settings, by adopting an experimental design approach, we applied ChatGPT-3 to a traditional form of assessment to determine its capabilities and limitations. Specifically, we tested its ability to produce an essay on a topic of our choice, created a rubric, and assessed the produced work in accordance with the designed rubric. We then evaluated the chatbot's work by assessing ChatGPT's application of its rubric according to a modified version of Paul's (2005) Intellectual Standards rubric. Using Christensen et al.'s (2015) framework on disruptive innovations, our study found that ChatGPT was capable of completing the set tasks competently, quickly, and easily, like a "magic wand". However, our findings also challenge the extent to which all of the ChatGPT's demonstrated capabilities can disrupt this traditional form of assessment, given that there are aspects of its construction and evaluation that the technology is not yet able to replicate as a human expert would. These limitations of the chatbot can provide us with an opportunity for addressing vulnerabilities in traditional forms of assessment in higher education that are subject to academic integrity issues posed by this form of AI. We conclude the article with implications for teachers and higher education institutions by urging them to reconsider and revisit their practices when it comes to assessment.

Keywords: ChatGPT, Higher education, Artificial intelligence (AI), Assessment, Academic integrity

1. Introduction

Given today's realia, the role of technology has become indispensable in every sector of our lives, including education. In educational settings, *Information and Communication Technologies (ICTs)*, i.e., digital technologies that process and disseminate information for purposeful decision-making and collaboration, have proven to be efficient for both teaching and learning practices (Fernández-Gutiérrez, Gimenez and Calero, 2020). Additionally, digital technologies and software played a significant role in supporting remote learning which was accelerated by the global pandemic (Susnjak, 2022). As a tool, technology is generally seen as a supportive and assistive aid (Köprülü, 2021), especially in higher education settings.

However, a recently launched *Artificial Intelligence (AI)* technology called *ChatGPT* has not only captured people's attention in a short space of time but has also been able to spark heated scholarly debates. In this paper, AI refers to digital technologies that engage in reasoning and self-correction based on learning obtained through training on large knowledge bases (Kok et al., 2009). ChatGPT, launched in November 2022 by Open AI (OpenAI, 2023), has created controversy around its use and application in educational settings. For example, Zhai (2022) explored the capabilities of ChatGPT to write an academic paper and concluded that since the output was of good quality, students could now outsource their writing tasks. In another study conducted by Susnjak (2022), ChatGPT was able to produce reasonable answers during assessment with a level of accuracy that could potentially be used to disrupt academic integrity in online exams. Strikingly, GPT-3, Thunström and Steingrímsson's (2022) study queried the chatbot to write an entire academic article about itself which was then published in a journal. This study alone confirmed how little human input and involvement this AI requires which makes us question its long-term effect in disrupting academic integrity and jeopardising the roles of educators.

In order to address these concerns, it is important to investigate not only the capabilities of ChatGPT but to also understand its limitations. To date, many studies have focused on the chatbot's potential to disrupt educational processes (Zhai, 2022), but there is still a dearth of research in identifying its gaps specifically in educational settings. Since the chatbot may be used to replicate an act of assessment between students and educators, we believe it is important to understand to what extent this replication impacts traditional processes in Higher Education settings. Considering these points in this study, we evaluated ChatGPT's performance by tasking it with academic essay writing in order to identify its capabilities and limitations specifically in higher education

settings. The essay as a traditional form of assessment is already critiqued as approaching obsolescence due to the availability of much openly accessible knowledge that can be easily copied and the difficulty in achieving consistency and reliability in marking it (Race, 2018).

Taking into consideration all the above concerns, the study was thus guided by the following three research questions:

Q1: To what extent can ChatGPT replicate the process of delivering an academic essay on a topic created at random?

Q2: To what extent can ChatGPT effectively assess the essay created by itself?

Q3: With reference to the findings from Q1 and Q2, to what extent can ChatGPT be considered a disruptor?

The paper is expected to contribute to the sparse literature on the implications of the use of ChatGPT on higher education assessment, an issue which is receiving increasing attention in academia (Kasnezi et al., 2023) and in educational practice (Stokel-Walker, 2022; Rudolph, Tan and Tan, 2023). It brings a unique perspective on the AI's capabilities and limitations as a disruptor in higher education assessment.

2. Literature Review

Studies about the role of technology in higher education have reported that it can enhance the learning process and impact student engagement (Bond et al., 2020) and improve communication between educators and students (Akour and Alenezi, 2022). Menkhoff and colleagues (2015), for example, reported that students found writing entries on Twitter a more interactive experience than traditional lectures. In a different study, Oliva-Córdova, García-Cabot and Amado-Salvatierra (2021) highlighted that the use of technology as a pedagogical tool can improve the learner experience. Other studies highlighted the capacity of technological tools to create dynamic, creative, and learner-centered spaces (Fojtik, 2014).

Recently, rapid changes in the role of technology in higher education occurred during the *emergency remote teaching (ERT)* period, defined as a “putatively ephemeral shift to remote teaching to continue teaching and learning during emergencies” (Sum and Oancea, 2022, p. 1). The pandemic was a turning point for elevating the role of technology in schools and in higher education settings, necessitating educators to shift to more technology use. Technology became an integral part of teaching and a mediator between educators and students who connected through online learning platforms.

Notwithstanding the benefits of remote teaching (Susnjak, 2022), technological challenges were also evident such as bad Internet connection (Sum and Oancea, 2022), lack of resources and institutional support (Joshi, Vinay and Bhaskar, 2020), and the issue of sustaining academic integrity in online exams (Susnjak, 2022). Online examinations posed a serious challenge due to the lack of direct supervision and “the ease with which students may be able to access and share resources during the exam” (Susnjak, 2022, p. 2). It seems that this online mode of assessment became particularly prone to cheating (Arnold, 2016) even more so than traditional exams. In order to prevent the growth of cheating in online exams, institutions began to utilise relevant technological tools including software such as proctoring and plagiarism detection (Susnjak, 2022). Educators were encouraged to move away from multiple-choice questions and use tasks that involved critical thinking and creativity (Whisenhunt et al, 2022). Due to these issues, emergency remote teaching was seen as a *disruptor* by many (Sum and Oancea, 2022), a phenomenon further discussed in the following section.

2.1 Disruptive Innovation

Regardless of the mentioned advantages and benefits, technology in education can also be labeled as disruptive, especially non-institutional technologies in higher education (Flavin, 2012). The term disruptive technology was coined by Christensen (1997) to identify easy-to-use technologies with the potential to displace more incumbent technologies. Disruptive technology can be described as changing

“traditional practices usually starting with a small number of users and then growing over time in such a way that it displaces a well-established and prominent practice” (Siddhpura and Siddhpura, 2020, p. 494).

The term “disruptive technology” was later changed to “disruptive innovation” as a way of extending it and including other sectors such as service sectors (Christensen and Raynor, 2003). The theory of disruptive innovation can be applied to many sectors including higher education (Flavin, 2021). Christensen and Raynor (2003) propose two types of disruption which are “new market disruption” and “low-end disruption”. New

market disruption can be understood as establishing a new market that meets the needs of the customers which previously were unmet (Christensen and Raynor, 2003). A low-end disruption can be understood as the use of new technologies to exceed the performance of the existing market (Christensen and Raynor, 2003). Disruptive innovations can be contrasted with sustaining and efficiency innovations that merely improve existing goods and enhance processes respectively, rather than bring about transformative change (Christensen, Raynor and McDonald, 2015).

In higher education contexts, Al-Imarah and Shields (2016) propose three characteristics of disruptive innovation, performance, benefits, and market. New disruptive innovations initially do not perform well but they have a tendency to improve over time resulting in the attraction of mainstream customers (Al-Imarah and Shields, 2016). The benefits of disruptive innovations consist of low prices, convenience, and ease of use. Regarding the market, disruptive innovations initially have limited customers, but once they develop their performance, they can create new markets and become competitors for existing mainstream goods and services (Christensen, Raynor and McDonald, 2015). In Al-Imarah and Shields's (2016) study, they explored Massive Open Online Courses' (MOOCs) potential to disrupt existing models of higher education. Their analyses, using these three characteristics, revealed that MOOCs only supported one characteristic of disruptive innovation, i.e., they have the possibility of creating a new market by targeting part-time students, distance learners, and self-directed students (Al-Imarah and Shields, 2016). They concluded that MOOCs can therefore only be characterised as sustainable innovation.

Although disruptive innovation theory has been critiqued for its inability to predict an innovation's disruptive potential (Flavin, 2016), it still remains a useful framework for analysing new technologies in higher education settings, particularly to learn "why some succeed and some fail" (Flavin, 2021, p. 3). Flavin (2021) argued that disruptive innovations offer cheap and easy-to-use technologies that first attract a particular market and later target mainstream ones as well. This can be seen, for example, in the case of budget airlines disrupting incumbent airlines (Kumar, 2006) or car rentals disrupting dominant car brands (Markides and Sosa, 2013). Similarly, in higher education, the emergence of ChatGPT is attracting heated debates about its disruptive potential. Some are identifying it as a revolutionary technology that can improve higher education processes (Fauzi et al., 2023) whereas others seem to be thinking differently (Kasneci et al., 2023), a debate which is discussed further below.

2.2 ChatGPT

ChatGPT is a *large language model (LLM)* developed by OpenAI to engage with and respond to users' prompts and questions. LLMs are specialised AI technologies that use natural language processing (NLP) to generate human-like text and complete a variety of language-related tasks (Kasneci et al., 2023). ChatGPT is built upon the LLM technology, Generative Pretrained Transformer. ChatGPT, also known as a chatbot, an application that mimics conversations with humans, responds to queries within seconds (Rudolph, Tan and Tan, 2023) and was trained on a wide variety of texts including academic articles, books, and websites covering themes and topics from science, fiction, and news reports (Shen et al., 2023). It is also trained through human feedback (Zhai, 2022), enabling it to expand its corpus and areas of knowledge over time.

ChatGPT-3 was released 30th of November, 2022 (Tate et al., 2022), and quickly attracted attention due to several reasons. Firstly, the chatbot was made available to the public through setting up a free account. Secondly, the interface of the chatbot is user-friendly, in the form of a conversation (Tate et al., 2022). Thirdly, the chatbot creates a sense of interaction, taking into account previous prompts and questions. However, it is worth noting that in the process of preparing this paper, OpenAI released a new version of the chatbot on 1st of February 2023 which operates on a paid subscription model and offers access to new updates, faster response times, and quicker access even at peak times (OpenAI, 2023).

Several studies investigating ChatGPT's capabilities have highlighted its remarkable ability to produce a fluent piece of writing on a wide range of topics (Shen et al., 2023; Tate et al., 2022). For instance, Tate et al. (2022) noted that besides giving answers to factual questions the AI is also capable of writing essays, poems, reports, plays, and stories on "almost any subject described; writing a critique of that same text from the point of view of a teacher, professor or literary scholar" (pp. 4-5). Pavlik's (2023) study also confirmed this after tasking the chatbot with factual questions, demonstrating that ChatGPT was not only capable of generating high-quality written texts but also of presenting these texts in a manner similar to that of a human. Zhai (2022) reported that it took only a few hours to write and finish a study using ChatGPT and required little human input. Terwiesch (2023) mentioned that exam development normally takes around 20 hours and then another 10 for the Teaching Assistant to test it and prepare solutions. But with the help of ChatGPT, it was possible to develop the exam in

10 hours and then cut the Teaching Assistant's time from 10 to 5 hours (Terwiesch, 2023). Although in Zhai's (2022) and Terwiesch's (2023) studies, it is evident that the chatbot increases productivity and efficiency, it also has the potential to put roles at risk such as that of Teaching Assistants.

In this vein, the chatbot has the potential to emulate human endeavour by, for example, co-authoring a journal article. For example, (O'Connor and ChatGPT, 2023) is "co-authored" by ChatGPT, thus raising ethical issues around authorship of academic outputs and academic integrity.

In defence of academic integrity, O'Connor and ChatGPT (2023) argued that the chatbot could be used to combat plagiarism by "providing students with tools and resources that can help them properly cite" (p. 1) the needed sources. However, Qadir's (2022) study reported that the chatbot generated references with non-functional links and even ones that did not exist. Aydin and Karaarslan's (2022) study also explored ChatGPT's ability to avoid plagiarism by asking the chatbot to paraphrase academic articles' abstracts and then check the paraphrased abstracts for plagiarism. They found that the match rates of ChatGPT's paraphrased abstracts were high which made them conclude that the chatbot may not be able to successfully elude plagiarism detection.

Both Pavlik's (2023) and Aydin and Karaarslan's (2022) studies urged educators to consider the role of ChatGPT and its impacts on academic integrity in higher education. ChatGPT is likely a forerunner of many similar LLMs (Tate et al., 2022), and the number of users who will "consult" or "use" these AIs will eventually grow making it even more important to identify how this technology will influence higher education settings.

3. Methodology

This study adopts an experimental design approach where we tested ChatGPT-3's performance and its ability to respond to a set of fixed queries. To investigate the extent to which ChatGPT can influence the processes of assessment in higher education, we selected the essay, a traditional form of assessment, which nominally consists of these processes: design, delivery, and marking.

Building upon prior studies (Susnjak, 2022; Zhai, 2022), we designed *five specific queries* which tested ChatGPT's ability to: write an essay on a specific topic, using a specific referencing style (*design the essay*); write an introduction, body, and conclusion of the essay (*deliver the essay*); and also design a rubric and assess its own work in accordance with this rubric (*mark the essay*). The queries were conducted on ChatGPT's main website, chat.openai.com. The entire period of testing and analysis lasted from February until mid-March 2023. We documented our observations of ChatGPT's responses to the queries related to design and delivery and conducted our independent assessment of its marking by using Paul's (2005) Intellectual Standards rubric. To analyse ChatGPT's disruptive capacity, we applied the disruptive innovation lenses from the literature (Christensen, Raynor and McDonald, 2015; Al-Imarah and Shields, 2019; Flavin, 2021).

The following steps constitute the experimental research design:

Step 1. Firstly, we set up an account at Open AI which granted free-to-use public access to ChatGPT. The ChatGPT version that we used was released on January 30th, 2023, which had improved features over the previous version of January 9th, 2023, improving the chatbot's response on a wide range of topics, including an update in factuality and mathematical problem-solving. Once the account was set up, the main interface of the website provided examples, capabilities, and limitations of the chatbot (See Fig. 1).



Figure 1: ChatGPT interface

Step 2. Secondly, by opening the new chat section we tasked the chatbot to write an academic essay on Technology in Education using three queries. The **first query** was: *Write an introduction on the topic of Technology in Education with the context and aims of the article. State your argument and also indicate your answer to the argument. Write with 5 references in APA style.*

After the chatbot produced an answer, we posed the **second query**: *Try to develop your argument point by point. Mention the reasons that support your argument that were stated in the introduction. Use 5 different reference sources and write the essay with 500 words.* The chatbot produced the body of the essay containing five paragraphs. We posed the **third query** which was on the last part of the essay, the conclusion: *Summarise the reasons that support the mentioned argument without including any new information. Remind the reader of the points you mentioned and how you addressed the posed question.*

Step 3. After the chatbot produced all three parts of the essay, we tasked it with two further queries related to marking. The **fourth query** was: *Design a rubric to rate this essay.* The chatbot produced a points-based rubric with five criteria. These criteria were introduction, development of argument, summary and conclusion, presentation, creativity and originality. Once the rubric was produced, we asked the chatbot to rate the essay in accordance with its rubric, by using the following **fifth query**: *Rate this essay in accordance with the rubric you have designed.*

Step 4. After ChatGPT rated its own work, we independently evaluated the chatbot's essay using a modified version of Paul's (2005) Intellectual Standards rubric. Paul's (2005) rubric includes criteria such as clarity, accuracy, precision, relevance, depth, breadth, logic, significance, fairness. We further adapted the rubric with the additional criterion of "originality and creativity". Thus, the modified rubric contains the following criteria:

1. Clarity (the text is clear and easy to understand)
2. Accuracy (the information in the text is provided accurately)
3. Precision (the text is precise and provides enough details)
4. Relevance (the text is relevant to the topic and/or identified issue)
5. Depth (the text provides an in-depth analysis)
6. Breadth (the text considers other perspectives and/or viewpoints)
7. Logic (the text is presented in a logical manner)
8. Significance (the text reminds the reader of the significance of the topic)
9. Fairness (the text represents the viewpoints in a fair manner)
10. Originality and Creativity (the text demonstrates creative and original ideas).

Step 5. We analysed ChatGPT's disruptive capacity by applying the disruptive innovation lenses from the literature, which state that technologies are disruptive to the extent that they satisfy two main criteria, (a) provide an easy-to-use, cost-efficient, convenient, and simple alternative to the status quo and (b) when put to use significantly alter existing processes (Christensen, Raynor and McDonald, 2015; Al-Imarah and Shields, 2019; Flavin, 2021).

Step 6. We then decided to task ChatGPT to assess its essay again but this time using the rubric that we have designed in Step 4 to avoid the potential bias of the chatbot awarding itself overly favourable marks through the use of its own rubric. We could thus compare its use of a rubric independent from its own and contrast it with our assessment.

In the following sections, the results of the queries are presented and discussed.

4. Findings

In this section, we demonstrate how we applied the experimental research design to address the three guiding research questions, which serve as subheadings for each subsection below.

4.1 To What Extent can ChatGPT Replicate the Process of Delivering an Academic Essay on a Topic Created at Random?

From the responses to our queries 1, 2, and 3, it is evident that ChatGPT is capable of producing an essay on a specific topic, created at random, however, as can be seen from the query response below, the content of the essay that the chatbot generated was rather disappointing (Figs. 2-4). The chatbot created a very general and descriptive essay and in some cases, it was quite repetitive. For example, the generated text was quite restricted and did not incorporate wide-ranging perspectives as only five sources were consulted. Interestingly, all of these sources were from the years 2001, 2002, 2005, 2010, and 2011, with no recent references included (See Fig.4).

Even though the chatbot was able to produce the essay within one minute, the content itself was not remarkable and lacked creativity. For example, the generated text did not incorporate any innovative ideas or perspectives nor did it produce any original thinking. Moreover, there were inconsistencies in the chatbot's reply to the queries since the introduction and main body, for which the chatbot was tasked to provide 500 words each, only generated 258 and 287 words (See Figs 2-3).

In both the introduction and main body sections, we asked the chatbot to provide five different references in APA style. Although the chatbot managed to provide the requested style correctly, it was still not capable of providing five different references (see Fig 2 below). Both the introduction and main body sections had identical reference sources all of which were given only as in-text citations. Upon looking into the generated reference list (Fig 4) which contained only five sources we realised that none of the provided links were functional. In fact, two of the sources turned out to be bogus references, one reference's title was not correct and another reference's author name was given as Turkly (2011) instead of Turkle (2011).

Technology has revolutionized the way we live, work, and learn. In the field of education, technology has brought about tremendous opportunities and benefits, but also some challenges. The aim of this essay is to explore the impact of technology on education and to analyze its potential benefits and drawbacks.

The argument of this essay is that technology can greatly enhance the learning experience and promote student success, but it must be approached in a strategic and responsible manner, taking into consideration the needs and goals of the education system and its stakeholders. While technology has the potential to increase student engagement, motivation, and academic performance (Warschauer & Matuchniak, 2010), it can also lead to reduced face-to-face interaction, decreased social skills, and increased anxiety, depression, and other mental health problems (Turkle, 2011).

To answer this question, this essay will first provide a general overview of the integration of technology in education, including its benefits and challenges. Next, it will examine the role of technology in student-centered learning and collaboration (Roschelle & Pea, 2002) and in teacher professional development (Swan & Shen, 2005). Finally, the essay will consider the importance of ensuring equitable access to technology for all students (Cuban, 2001) and of approaching technology integration in a thoughtful and strategic manner.

In conclusion, technology in education can greatly enhance the learning experience and promote student success, but it must be approached in a responsible and strategic manner. By considering the benefits and challenges of technology integration, education systems can work to ensure that technology is used to support student learning and well-being.

Figure 2: Outcome related to query 1. Word count = 258

In support of the argument that technology can greatly enhance the learning experience and promote student success, there are several reasons to consider. First, technology can increase student engagement and motivation. The use of technology in the classroom can provide students with access to a wide range of interactive and multimedia resources that can help to increase their interest and understanding of the subject matter (Warschauer & Matuchniak, 2010). Furthermore, technology can allow for personalized and differentiated instruction, meeting the diverse learning needs of individual students (Cuban, 2001). Second, technology can facilitate student collaboration and teamwork. The use of communication and collaboration tools such as online forums, video conferencing, and social media can provide students with opportunities to work together on projects, share ideas, and receive feedback from peers and teachers (Roschelle & Pea, 2002). This can help to build strong relationships among students and increase their sense of community and belonging. Third, technology can support teacher professional development. The integration of technology in teacher education programs can provide teachers with opportunities to learn about the latest instructional strategies and tools, as well as to collaborate with peers and engage in ongoing professional learning (Swan & Shen, 2005). This can help to ensure that teachers are well-equipped to use technology effectively in their classrooms. However, while technology can provide numerous benefits for students and teachers, it is important to approach its integration in a responsible and strategic manner. For example, it is crucial to ensure equitable access to technology for all students, regardless of their socio-economic background or geographic location (Cuban, 2001). Additionally, it is important to be aware of the negative effects of technology on student well-being, such as decreased face-to-face interaction and increased anxiety and depression (Turkle, 2011).

Figure 3: Outcome related to query 2. Word count = 287

In summary, the integration of technology in education has the potential to greatly enhance the learning experience and promote student success by increasing student engagement and motivation, facilitating collaboration and teamwork, and supporting teacher professional development. However, it is crucial to approach technology integration in a responsible and strategic manner, taking into account the need for equitable access to technology and the potential negative effects on student well-being. The argument in this essay supports the idea that technology can have a positive impact on education when used in a thoughtful and strategic manner, and highlights the importance of considering both the benefits and challenges of technology integration.

References:

- Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Cambridge, MA: Harvard University Press.
- Roschelle, J., & Pea, R. D. (2002). Making sense of argumentation and communication technologies in school mathematics. *Journal of Research in Mathematics Education*, 33(2), 137-163. <https://doi.org/10.2307/749763>
- Swan, K., & Shen, J. (2005). The integration of technology into teacher education programs: A review of the literature. *Journal of Technology and Teacher Education*, 13(3), 571-596. <https://doi.org/10.1080/10705630500264956>
- Turkly, S. (2011). *Alone together: Why we expect more from technology and less from each other*. New York, NY: Basic Books.
- Warschauer, M., & Matuchniak, T. (2010). New technologies and learning. *Annual Review of Applied Linguistics*, 30, 276-292. <https://doi.org/10.1017/S0267190510000070>

Figure 4: Outcome related to query 3

4.2 To What Extent can ChatGPT Effectively Assess the Essay Created by Itself?

From the responses to our queries 4 and 5, it is evident that ChatGPT is capable of designing a bespoke rubric to mark the essay it produced and then rate its work in accordance with this rubric. However, with respect to the query responses the following issues were revealed. Even though the chatbot generated a generic rubric to assess the essay, we found some issues with the assessment itself. In our first attempt at asking the chatbot to rate the essay (see Fig 6), it marked the essay at 91 out of 100 points. We thought that the score was too high given the quality issues already noted, thus, we decided to ask the chatbot to rate the work again (see Fig 7) and surprisingly the second time the score was different. This time the chatbot rated the work at 93 out of 100 points, higher than in the previous attempt. Finally, on a third attempt at rating the essay (see Fig 8), the chatbot's score this time was 88 out of 100. One of the chatbot's comments on this score was "The essay could benefit from further innovative ideas and perspectives to increase its creativity and originality" with which we

surely agreed. However, the three attempts generated three different scores on the same essay which made us yet again question the chatbot's credibility and reliability.

In addition to the rating problems, we also found some issues with the designed rubric and calculations. ChatGPT's rubric was simplistic and generic, with only five rating criteria totalling 100 points (see Fig 5). Each of the criteria was weighted differently, with no clear logic evident for the assignment of weightings. Development of Argument (40 points) and Summary and Conclusion (20 points) were weighted the most even though the latter more or less summarised points made in the argument of the essay, introducing no new information, but achieving half the score of the former. The Introduction (10 points) and Presentation (10 points) categories were equally weighted although one category is about content while the other is about the look and feel of the work; it is unclear why they would attract the same weighting. Interestingly, Creativity and Originality were also given only 10 points, but this aspect is deemed of critical importance in constructing an essay (Bekurs and Santoli, 2004). Additionally, the total of the rubric's weightings was 90 not 100 points. However, ChatGPT presents it as a 100-point marking scale and even assessed the essay using these false weightings. It seems that the chatbot may have misapplied the rubric by allocating double the amount of marks for some categories while giving less for others. This demonstrates the chatbot's incompetence both in creating the rubric and fairness in allocating the weightings.

From our own independent assessment of the chatbot's generated essay, we gave it a total score of 41 out of 100 (See Fig. 9). Although we thought the work was written in a logically relevant, understandable, and easy-to-read manner we still found many issues. The content was very generic, and simple and did not demonstrate any in-depth analysis. For example, the essay did not conduct a thorough examination of the impact of technology on education and did not provide any detailed examples. The arguments were not backed up by evidence and the provided sources were inaccurate. For instance, there were five sources in total and none of the provided links were functional. The text failed to illustrate any original and creative thinking and did not integrate any innovative ideas or perspectives. For example, the text highlights the importance of considering both the benefits and challenges of technology integration but does not engage in discussing the ways in which teachers can creatively integrate technology into their practices. Notably, ChatGPT's assessment using our rubric resulted once more in a quite generous score of 88/100 (see Fig 10).

1. Introduction (10 points)
• Clarity of the context and aim of the essay (3 points) • Clearly stated argument (3 points) • Explanation of how the question will be answered (2 points) • Relevance to the topic of technology in education (2 points)
2. Development of argument (40 points)
• Coherent and logical flow of ideas (10 points) • Evidence-based support for argument (10 points) • Use of relevant and credible sources (5 points) • Integration of personal insights and perspectives (5 points) • Consideration of alternative viewpoints (5 points) • Clear and concise language (5 points)
3. Summary and conclusion (20 points)
• Recap of the main points (5 points) • Restatement of the argument (5 points) • Clear and concise language (5 points) • Reflection on the implications of the argument (5 points)
4. Presentation (10 points)
• Clarity of organization and structure (3 points) • Correct use of grammar, spelling, and punctuation (3 points) • Effective use of visuals and formatting (2 points) • Adherence to the required citation style (2 points)
5. Creativity and originality (10 points)
• Use of innovative ideas and perspectives (5 points) • Integration of personal experiences and insights (2 points) • Reflection on the broader implications and application of the argument (3 points)
Total: 100 points

Figure 5: Outcome related to query 4 – ChatGPT's generated rubric

<i>Introduction (10 points) - 9/10</i> <i>The introduction provides a clear context for the essay and states the argument well. It also explains the aim and approach of the essay.</i> <i>Development of Argument (40 points) - 36/40</i> <i>The essay provides a comprehensive examination of the impact of technology on education, including its benefits and challenges. The argument is well-supported with evidence and credible sources. The personal insights and perspectives are integrated effectively into the argument. The essay also considers alternative viewpoints. The language is clear and concise.</i> <i>Summary and Conclusion (20 points) - 18/20</i> <i>The summary and conclusion provide a clear recap of the main points and restate the argument effectively. The language is clear and concise. The reflection on the implications of the argument is well done.</i> <i>Presentation (10 points) - 9/10</i> <i>The essay is well-organized and structured. The use of grammar, spelling, and punctuation is correct. The visuals and formatting are effective. The citation style is adhered to correctly.</i> <i>Creativity and Originality (10 points) - 9/10</i> <i>The essay provides innovative ideas and perspectives, integrates personal experiences and insights effectively, and reflects on the broader implications and applications of the argument well.</i> <i>Overall, the essay is well written and explores the impact of technology on education in a comprehensive and thoughtful manner. The argument is well-supported and the presentation is effective. A total score of 91/100 is earned.</i>

Figure 6: Outcome related to query 5 – ChatGPT’s rating of the essay. (Iteration 1). A total score of 91/100 (note. ChatGPT’s incorrect calculation)

<i>Introduction (10 points) - 9/10</i> <i>The introduction provides clarity of the context and aim of the essay. It also explains how the question will be answered.</i> <i>Development of Argument (40 points) - 37/40</i> <i>The essay provides a coherent and logical flow of ideas. The argument is supported with relevant and credible sources. The essay integrates personal insights and perspectives, but at the same time considers alternative viewpoints.</i> <i>Summary and Conclusion (20 points) - 18/20</i> <i>The conclusion provides a recap of the main points and restates the argument. The presented language is clear and concise. The implications of the argument are reflected effectively.</i> <i>Presentation (10 points) - 9/10</i> <i>The structure and organization of the essay is efficient. The use of grammar, spelling, and punctuation is effective. Both the visuals and formatting are effective too. The required citation style is presented correctly.</i> <i>Creativity and Originality (10 points) - 10/10</i> <i>The essay is innovative and provides different perspectives. It not only integrates insights and experiences, but also provides a reflection on the argument application and implications.</i> <i>Total: 93/100 points</i> <i>Overall, the essay provides a clear and well-supported argument about the benefits and challenges of technology integration in education. The essay effectively integrates evidence from relevant and credible sources and provides personal insights and perspectives on the topic. The organization and structure of the essay are clear and effective, and the language is generally clear and concise. The essay shows creativity and originality by reflecting on the broader implications and application of the argument. However, there are minor issues with the clarity of the argument in the introduction, and some minor errors in grammar, spelling, and punctuation throughout the essay.</i>
--

Figure 7: Outcome related to query 5 – ChatGPT’s rating of the essay. Iteration 2. A total score of 93/100 (note. ChatGPT’s incorrect calculation)

Introduction (10 points) - 9/10

The context for the essay and the argument is clear and effective. The essay explains the relevance of the topic its aim and its approach.

Development of Argument (40 points) - 40/40

The essay is coherent and provides a logical flow of ideas. The essay uses credible sources and integrates alternative viewpoints. The language of the essay is clear and concise. The arguments are supported by the evidence.

Summary and Conclusion (20 points) - 20/20

The summary and conclusion provide a strong recap of the main points reflection on the implications of the argument.

Presentation (10 points) - 10/10

The presentation is well-organized and adheres to the required citation style. The use of grammar, spelling, and punctuation is correct. The essay uses effective visuals and formatting.

Creativity and Originality (10 points) - 8/10

The essay does not provide innovative ideas and perspectives. The integration of personal experiences and insights is not effective enough. The reflection on the broader implications and applications of the argument is not well enough.

Overall, I would rate this essay as 88/100. The essay effectively presents a clear and well-supported argument, with a coherent and logical flow of ideas and use of relevant and credible sources. The language is clear and concise, and the essay effectively considers alternative viewpoints. The summary and conclusion provide a strong recap and reflection on the implications of the argument. The presentation is well-organized and adheres to the required citation style. However, the essay could benefit from further innovative ideas and perspectives to increase its creativity and originality.

Figure 8: Outcome related to query 5 – ChatGPT’s rating of the essay. Iteration 3. A total score of 88/100 (note. ChatGPT’s incorrect calculation)

Clarity (10 points) - 7/10. The text was very simple and easy to read. The meaning of the text is understandable and easy to grasp.

Accuracy (10 points) - 3/10. The sources given in the text were made up and not accurate. Thus making the alleged facts unreliable.

Precision (10 points) - 5/10. The text provides generic information without giving enough details. Alleged facts could have been more specific.

Relevance (10 points) - 6/10. The text was relevant to the topic but it did not identify and cite the key authors' works. This makes us assume that the knowledge in specific areas is still weak and the bot is unable to demonstrate broad reading.

Depth (10 points) - 3/10. The text did not provide an in-depth analysis. The generated answers were constrained only to 258 and 287 even though we requested to write 500 words for each of the two sections.

Breadth (10 points) - 3/10. The text was rather narrow-minded and did not incorporate wide-ranging perspectives and multiple viewpoints. In all sections, the same sources were used repeatedly and new ones were not incorporated.

Logic (10 points) - 6/10. The text is presented in a simplistic but logically relevant way.

Significance (10 points) - 6/10. The text reminds the reader about the significance of the topic but in a rather simplistic manner.

Fairness (10 points) - 1/10. The text provided made-up references. The bot uses the data to which only it has access and which is not independently verifiable, which potentially puts fairness into question.

Originality and creativity (10 points) - 1/10. The text failed to demonstrate originality and creativity. It was quite repetitive and generic and did not contain any novel or original ideas.

Total: 41/100

Figure 9: Researchers’ independent assessment of ChatGPT’s essay. A total score of 41/100.

<i>Clarity (10 points) - 9/10. The essay is well-structured and generally easy to understand, earning a score of 9 out of 10 points. There are a few complex sentences and concepts that could be simplified for better clarity.</i>
<i>Accuracy (10 points) - 10/10. The information provided in the essay is accurate, and references are appropriately cited. It deserves a score of 10 out of 10 points.</i>
<i>Precision (10 points) - 9/10. The essay is precise and provides sufficient details to support the arguments. It deserves a score of 9 out of 10 points, as a few concepts could be elaborated on further.</i>
<i>Relevance (10 points) - 10/10. The essay is highly relevant to the topic of technology in education, earning a score of 10 out of 10 points.</i>
<i>Depth (10 points) - 8/10. The essay provides a reasonably in-depth analysis of the topic, but it could benefit from delving deeper into some of the discussed aspects. It scores 8 out of 10 points.</i>
<i>Breadth (10 points) - 7/10. While the essay acknowledges the importance of responsible technology integration and equity, it could further consider counter-arguments or alternative viewpoints for a more comprehensive perspective. It receives 7 out of 10 points.</i>
<i>Logic (10 points) - 9/10. The essay is logically presented, with well-structured arguments and a clear flow. It deserves a score of 9 out of 10 points.</i>
<i>Significance (10 points) - 9/10. The essay effectively reminds the reader of the significance of the topic, earning 9 out of 10 points.</i>
<i>Fairness (10 points) - 10/10. The essay represents viewpoints fairly, presenting both the benefits and drawbacks of technology integration. It scores 10 out of 10 points.</i>
<i>Originality and Creativity (10 points) - 7/10. The essay is more focused on presenting existing research and viewpoints, so it lacks a significance level of originality and creativity. It receives 7 out of 10 points.</i>
<i>Total: 88/100</i>

Figure 10: ChatGPT's assessment in accordance with our designed rubric. A total score of 88/100.

4.3 To What Extent can ChatGPT be considered a Disruptor?

With respect to the processes of producing an essay, all of the queries submitted to the chatbot were completed within a short period of time. The full essay, rubric, and marking were completed by the chatbot in approximately one minute. The application is easy to use with a user-friendly, easy-to-navigate interface (see Fig 1) and can be accessed through a simple URL on the Web (openai.com). No costs were associated with its use. With respect to each of the processes of design, delivery, and marking of the essay assessment, we made the following observations. Regarding the design, although the chatbot executed the essay design, it was flawed as noted in the findings above. Regarding the delivery, our own independent assessment of the outcome resulted in a very mediocre grade due mainly to lapses in accuracy, depth, breadth, fairness, originality, and creativity. With regard to the marking, our observations and independent assessment revealed serious flaws as outlined in the previous section, i.e., a flawed rubric and inconsistent marking. Thus, when applying the disruptive innovation lens, we find that ChatGPT satisfied the first criterion, i.e., that it was fast, convenient, easy to use, and cost-efficient, however, with respect to the second criterion, its effect on the existing processes was inconclusive. While it produced plausible outcomes in response to the queries, those outcomes, upon closer examination were questionable. The implications of this finding for ChatGPT's potential to disrupt markets in higher education will be further discussed in the following section.

5. Discussion

Drawing on the findings, we believe that ChatGPT has not reached its exceptional levels yet. Although we agree with Shen et al.'s (2023) findings that it demonstrated a number of impressive capabilities, we still believe that the chatbot is far from being "exceptional" let alone be trusted with work for academic purposes. The discussion of the findings is structured according to the research questions in the sections that follow.

5.1 To What Extent can ChatGPT Replicate the Process of Delivering an Academic Essay on a Topic Created at Random?

Although the results of this study indicated that ChatGPT was capable of delivering an academic essay chosen at random it still has not succeeded in producing efficient work. On one hand, the chatbot was quick in responding to the queries and in fact, was able to produce written work in less than one minute's time, which may be favored

by many, especially by those users who are looking for a faster way to accomplish their tasks. On the other hand, it is important to note that as per our independent assessment of ChatGPT's produced work, we found it to be of a quite mediocre level. There were many inconsistencies in the chatbot's replies to our queries and these are discussed below.

First, as noted in the findings, it became clear that ChatGPT was not capable of producing a written piece of work with the requested number of words. Similar findings have also been reported by Rudolph, Tan and Tan's (2023) study. In their study, they tasked the chatbot to write a 2000-word essay but the chatbot was only capable of producing 500 words regardless of repeated attempts.

Second, the chatbot could not produce the needed results for referencing and identifying key scholars' works. The chatbot kept using the same five sources across different sections and did not incorporate any recent up-to-date sources. Upon examination, it became clear that none of the source links were functional and in fact, two of them turned out to be bogus. Additionally, one of the provided reference's titles was incorrect and another reference contained a mistake in the author's name. These findings resonate with Farrokhnia et al. (2023), Rudolph, Tan and Tan (2023), and Qadir (2022) who also reported that the chatbot generated made-up reference lists with non-functional links. In the cases when the users prompted the chatbot to provide an up-to-date reference list, it fabricated non-existent sources (Farrokhnia et al., 2023). Our findings in this study also corroborate that the chatbot is still not capable of producing reliable and up-to-date references.

Third, the content produced by ChatGPT was rather disappointing when we independently assessed it. The essay was primarily descriptive and did not provide any evidence to back up its arguments and statements. It lacked original thinking and was quite generic. This finding is in line with Pavlik (2023), Rudolph et al. (2023), and Tlili et al. (2023).

5.2 To What Extent can ChatGPT Effectively Assess the Essay Created by Itself?

Although it took less than one minute for the chatbot to rate the work, the results showed that it is still not capable of critical thinking and assessment, a finding opposite to Susnjak's (2022) study, which claimed that ChatGPT demonstrated critical thinking skills and was capable of generating high-quality text that was hard to distinguish from humans' work. Additionally, although the essay is seen by some as problematic, it is still one of the few vehicles for assessment that allows for creative and critical thinking (Bekurs and Santoli, 2004). We have determined that ChatGPT, while able to perform well on many of the criteria used to examine intellectual ability (Paul, 2005), would struggle to master critical thinking dimensions such as significance and fairness and more intangible ones such as originality and creativity. AI learning modules are only as good as the data on which they are trained and creativity requires divergence rather than convergence of ideas (Rudolph, Tan and Tan, 2023). As for the marking aspect of the chatbot-generated assessment exercise, the errors noted with the rubric and its application suggest that this part of the process is underdeveloped and subject to spurious outcomes, hence not reliable (Kabir et al., 2023). As illustrated in the findings, the chatbot was unable both to create a rubric and fairly assess the work in accordance with the rubric. Additionally, the chatbot demonstrated inconsistency in its evaluation of the work. While it is acknowledged that assessment by humans can also be inconsistent and biased (Hanesworth, Bracken and Elkington, 2019), results based on AI assessment are quite often "black-boxed" and neither educators nor learners have access to the logic and reasoning deployed by the chatbot in arriving at its judgements (Swiecki et al., 2022). This is different to human-based assessment where familiarity with the learners and the learning context provide a background to, and possible explanations for, human-based judgements. Human markers are also available to query in case of moderation of marks, for example, unlike chatbots. ChatGPT's generous assessment using our independent rubric also casts doubt on the chatbot's capacity for fair and critical assessment as it appears to provide random scores, echoing Farrokhnia et al. (2023) who argued that the chatbot lacks higher-order thinking skills and is unable to evaluate the quality of responses. It has also been argued that to develop the credibility of AI-based assessment, large investment in training on appropriate and targeted datasets is needed, however, this would add further complexity to the previously mentioned problem of blackboxed AI judgement in assessment (Mizumoto and Eguchi, 2023).

5.3 With Reference to the Findings From Q1 and Q2, to What Extent can ChatGPT be Considered a Disruptor?

Vis-à-vis the third research question, the Findings section has established limited evidence of ChatGPT being a disruptive innovation with respect to its ability to alter existing processes related to essay assessment. According to Christensen et al. (2015) and Al-Imarah and Shields (2019), disruptive innovations either create new markets or feed into the low-end of existing markets, i.e., those individuals who are less discerning and demanding. Additionally, the innovation, rather than being transformative of that market, could instead be either sustaining

or efficiency-enhancing. To determine how ChatGPT would affect a market related to essay assessment, we considered contract cheating (Clarke and Lancaster, 2006) where there is an existing “market” for commissioning bespoke essays (Ellis, Zucker and Randall, 2018) e.g., using so-called essay mills (Sweeney, 2023), or using pre-written material from essay “banks” (Medway, Roper and Gillooly, 2018). The “buyer”, i.e., the learner who wishes to cheat, can choose from different levels of quality, price, turnaround, and subject matter on offer (Wallace and Newton, 2014; Rigby et al., 2015). Higher education institutions acknowledge the difficulty in both detecting and deterring the acquisition of essay assignments in this manner (Ellis, Zucker and Randall, 2018). Some learners already use essay mills to cheat the system and produce work that they believe will attain a desirable grade, however, this level of quality is not guaranteed (Medway, Roper and Gillooly, 2018). ChatGPT, since it is relatively easy to use, convenient, cheap and quick, is likely to disrupt this contract cheating essay business model if it can produce essays that at least meet a decent grade threshold for the learner wishing to submit non-original work. We have demonstrated, however, in the findings, that the quality of a ChatGPT-generated essay is likely to be a bare pass. Hence, ChatGPT is likely to affect the status quo in contract cheating essay business models at the low end of the existing market. This finding concurs with speculations about ChatGPT’s influence on higher education assessments (Stokel-Walker, 2022). It can produce a product desirable for learners who are not discerning “customers”, i.e., who just need to pass their courses but are not seeking a particular academic performance level. Since this may not represent a unique market segment, it is unlikely, then, that ChatGPT would be able to create a new market for cheating on essay assignments. We have also already observed in the findings that, related to the design, delivery, and marking processes of the essay assessment, ChatGPT cannot be considered transformative, since its outcomes were deemed to have errors. Thus, it could be a sustaining and/or efficiency-enhancing innovation in this market. The errors observed in the findings negate its ability to enhance efficiency, even though the results were generated far more quickly than would be done by a human learner. However, since it could strengthen and retain “customers” at the lower-end of the market for contract cheating in essays, it can be assumed that ChatGPT will *sustain* technology-enabled ways of cheating, but hardly disrupt or enhance their efficiencies. This resonates with recent studies highlighting that AI utilised in addition to existing essay mills poses an additional threat to academic integrity, but one which is still evolving (Sweeney, 2023).

6. Conclusions and Implications

The paper contributes to the evolving body of research on the influence of ChatGPT on higher education assessment by demonstrating that ChatGPT has not yet reached the level of “disruptor”, but it is known that the AI will keep being trained and improved, thus retaining this potential in the future. The chatbot can generate answers to queries within minutes but from our study, it was evident that the results may not always be reliable. Based on our findings and analysis, ChatGPT is more at the “magic wand” than the disruptor end of the spectrum of disruptive innovations. Like a magician, it has achieved the illusion of a completed assessment but not an authentic innovation. It has achieved a quick and superficially effective delivery of the essay assessment and at least sustainable innovation in the market for contract cheating on essays, however, it has not succeeded in transforming the product or in creating new markets. It may have an effect on sustaining the lower end of the market for such cheating but not in disrupting it. These current limitations of the chatbot open up a window of opportunity for addressing current weaknesses of the traditional essay and other forms of assessment that may be vulnerable to cheating from AI-generated content. The following implications discuss this further.

6.1 Implications

Higher education instructors need to be equipped with the knowledge of special AIs that can detect ChatGPT-produced work, e.g. ZeroGPT, or as our study found, to employ a range of techniques to detect it. Our findings suggest that ChatGPT-produced work can be detected by: checking the reference list and determining whether the references are functional; examining the text if the same in-text references are being repeated across different sections; and checking for the lack of inclusion of novel or original ideas.

Studies are reporting that it will become nearly impossible to distinguish the students’ own words from the words of this AI (O’Connor and ChatGPT, 2023) due to ChatGPT’s capability in generating different responses to the same query (Susnjak, 2022). With the high chances of students outsourcing their tasks, it becomes even more important for educators to reconsider and revisit their assessment practices.

One way forward in addressing these would be the adoption of diverse and innovative tasks that will tap into the students’ critical thinking and creativity. Universities could return to oral examinations and also recommend that educators incorporate multimedia into tasks since the chatbot generates answers to text-based prompts

only. Educators could also tap into constructivist ideas that involve students' reflections, collaborations with other students, exploratory discussions, and presentations. Multipart assessment could also be incorporated by including both written and spoken parts. Universities need to educate their students on the ethics of using such AI and establish their regulations in terms of where they stand when it comes to the use of AI. After all, as the chatbot will be trained and improved over time fundamental changes to assessment might need to be in place to avoid the potential disruptions it may bring.

On the flip side, since the chatbot has limitations as far as creativity and originality are concerned it opens up the possibility that students could use it, in essay writing, for the more routine work of finding materials, assembling arguments, and creating a logical flow. This could potentially enable students to then focus on the higher-order capabilities of broader reading, better synthesis of ideas, drawing upon diverse perspectives, i.e., those aspects that increase their creative and original thinking.

6.2 Limitations and Further Work

It is worth mentioning that this study tested ChatGPT-3 which was released on the 30th of November, 2022. Therefore, the results of this study may not apply to the latest ChatGPT version. Despite the mentioned limitations, this study provided interesting findings in relation to ChatGPT's role in higher educational settings, specifically in the area of essay assessment. As ChatGPT will keep actively being developed and improved, future research could look into the conduct of a comprehensive systematic literature review comparing the different releases of this AI and reporting back on what types of features were added and improved as far as they may affect higher education assessment capability. Since new versions of this AI will keep being released future research could also develop special guidelines on approaching ChatGPT ethically and responsibly. We also hope this study will foster more research investigating further academic integrity and cheating concerns when it comes to the use of AIs in academic assessment.

References

- Akour, M. and Alenezi, M., 2022. Higher Education Future in the Era of Digital Transformation. *Education Sciences*, 12(11), p.784. Available at: <https://doi.org/10.3390/educsci12110784>.
- Al-Imarah, A.A. and Shields, R., 2019. MOOCs, disruptive innovation and the future of higher education: A conceptual analysis. *Innovations in Education and Teaching International*, 56(3), pp.258-269. Available at: <https://doi.org/10.1080/14703297.2018.1443828>.
- Arnold, I.J.M., 2016. Cheating at online formative tests: Does it pay off? *The Internet and Higher Education*, 29, pp.98–106. Available at: <https://doi.org/10.1016/j.iheduc.2016.02.001>.
- Aydin, Ö. and Karaarslan, E., 2022. OpenAI ChatGPT generated literature review: Digital town in healthcare. *Social Science Research Network*, 4308687. Available at: <http://dx.doi.org/10.2139/ssrn.4308687>.
- Bekurs, D. and Santoli, S., 2004. Writing is power: Critical thinking, creative writing, and portfolio assessment. *Essays in Education*, 10(1), p.2. Available at: <https://openriver.winona.edu/eie/vol10/iss1/2>.
- Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O. and Kerres, M., 2020. Mapping research in student engagement and educational technology in higher education: a systematic evidence map. *International journal of educational technology in higher education*, 17(1), pp.1-30. Available at: <https://doi.org/10.1186/s41239-019-0176-8>.
- Christensen, C.M., 1997. *The Innovator's Dilemma: when new technologies cause great firms to fail*. Harvard Business School Press.
- Christensen, C.M. and Raynor, M.E., 2003. *The innovator's solution: Creating and sustaining successful growth*. Harvard Business School Press.
- Christensen, C. M., Raynor, M.E. and McDonald, R., 2015. What Is Disruptive Innovation? *Harvard Business Review*, 93(12), pp.44–53. Available at: <https://hbr.org/2015/12/what-is-disruptive-innovation>.
- Clarke, R. and Lancaster, T., 2006. Eliminating the successor to plagiarism? Identifying the usage of contract cheating sites. In: *2nd international plagiarism conference*. Northumbria Learning Press. pp.1–13.
- Ellis, C., Zucker, I.M. and Randall, D., 2018. The infernal business of contract cheating: understanding the business processes and models of academic custom writing sites. *International Journal for Educational Integrity*, 14(1), pp.1–21. Available at: <https://doi.org/10.1007/s40979-017-0024-3>.
- Farrokhnia, M., Banihashem, S.K., Noroozi, O. and Wals, A., 2023. A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, pp.1-15. Available at: <https://doi.org/10.1080/14703297.2023.2195846>.
- Fauzi, F., Tuhuteru, L., Sampe, F., Ausat, A. and Hatta, H., 2023. Analysing the Role of ChatGPT in Improving Student Productivity in Higher Education. *Journal on Education*, 5(4), pp.14886-14891. Available at: <https://doi.org/10.31004/joe.v5i4.2563>.
- Fernández-Gutiérrez, M., Gimenez, G. and Calero, J., 2020. Is the use of ICT in education leading to higher student outcomes? Analysis from the Spanish Autonomous Communities. *Computers & Education*, 157, 103969. Available at: <https://doi.org/https://doi.org/10.1016/j.compedu.2020.103969>.

- Flavin, M., 2012. Disruptive technologies in higher education. *Research in Learning Technology*, 20, pp.102-111. Available at: <https://doi.org/10.3402/rlt.v20i0.19184>.
- Flavin, M., 2016. Technology-enhanced learning and higher education. *Oxford Review of Economic Policy*, 32(4), pp.632-645. Available at: <https://doi.org/10.1093/oxrep/grw028>.
- Flavin M., 2021. A Disruptive Innovation perspective on students' opinions of online assessment. *Research in Learning Technology*, 29. Available at: <https://doi.org/10.25304/rlt.v29.2611>.
- Fojtik, R., 2014. Mobile Technologies Education. *Procedia – Social and Behavioral Sciences*, 143, pp.342-346. Available at: <https://doi.org/10.1016/j.sbspro.2014.07.417>.
- Generative Pretrained Transformer, G., Thunström, A., Osmanovic, and Steingrímsson, S., 2022. Can GPT-3 write an academic paper on itself, with minimal human input?. Available at: <https://hal.science/hal-03701250>.
- Hanesworth, P., Bracken, S. and Elkington, S., 2019. A typology for a social justice approach to assessment: learning from universal design and culturally sustaining pedagogy. *Teaching in Higher Education*, 24(1), pp.98–114. Available at: <https://doi.org/10.1080/13562517.2018.1465405>.
- Joshi, A., Vinay, M. and Bhaskar, P., 2020. Impact of coronavirus pandemic on the Indian education sector: Perspectives of teachers on online teaching and assessments. *Interactive Technology and Smart Education*, 18(2), pp.205–226. Available at: <https://doi.org/10.1108/ITSE-06-2020-0087>.
- Kabir, S., Udo-Imeh, D.N., Kou, B. and Zhang, T., 2023. Who Answers It Better? An In-Depth Analysis of ChatGPT and Stack Overflow Answers to Software Engineering Questions. *arXiv preprint*. Available at: <https://doi.org/10.48550/arXiv.2308.02312>.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn, J. and Kasneci, G., 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, p.102274. Available at: <https://doi.org/10.1016/j.lindif.2023.102274>.
- Kok, J.N., Boers, E.J.W., Kusters, W.A. and Van der Putten, P., 2009. Artificial Intelligence: Definition, Trends, Techniques and Cases. *Artificial Intelligence*, 1, pp.270-299.
- Kumar, N., 2006. Strategies to fight low-cost rivals. *Harvard Business Review*, 84(12), pp.104-112. Available at: <https://hbr.org/2006/12/strategies-to-fight-low-cost-rivals>.
- Köprülü, F., 2021. The Effect of Using Technology Supported Material in Teaching English to First-Year Primary School Children: On Their Academic Success During COVID-19. *Frontiers in Psychology*, 12. Available at: <https://doi.org/10.3389/fpsyg.2021.756295>.
- Markides, C. and Sosa, L., 2013. Pioneering and First Mover Advantages: The Importance of Business Models. *Long Range Planning*, 46(4), pp.325-334. Available at: <https://doi.org/https://doi.org/10.1016/j.lrp.2013.06.002>.
- Medway, D., Roper, S. and Gillooly, L., 2018. Contract cheating in UK higher education: A covert investigation of essay mills. *British Educational Research Journal*, 44(3), pp.393–418. Available at: <https://doi.org/10.1002/berj.3335>.
- Menkhoff, T., Chay, Y., Bengtsson, M., Woodard, J. and Gan, B., 2015. Incorporating microblogging (“tweeting”) in higher education: Lessons learnt in a knowledge management course. *Computers in Human Behavior*, 51(B), pp.1295-1302. Available at: <https://doi.org/10.1016/j.chb.2014.11.063>.
- Mizumoto, A. and Eguchi, M., 2023. Exploring the potential of using an AI language model for automated essay scoring. *Research Methods in Applied Linguistics*, 2(2), p.100050. Available at: <https://doi.org/10.1016/j.rmal.2023.100050>.
- O'Connor, S. and ChatGpt., 2023. Open artificial intelligence platforms in nursing education: Tools for academic progress or abuse? *Nurse Education in Practice*, 66, 103537. Available at: <https://doi.org/https://doi.org/10.1016/j.nepr.2022.103537>.
- Oliva-Córdova, L. M., Garcia-Cabot, A. and Amado-Salvatierra, H. R., 2021. Learning analytics to support teaching skills: a systematic literature review. *IEEE Access*, 9, pp.58351–58363. Available at: <https://doi.org/10.1109/ACCESS.2021.3070294>.
- OpenAI., 2023. Introducing ChatGPT Plus. OpenAI. Available at: <https://openai.com/blog/chatgpt-plus/>.
- Pavlik, J. V., 2023. Collaborating With ChatGPT: Considering the Implications of Generative Artificial Intelligence for Journalism and Media Education. *Journalism & Mass Communication Educator*, 78(1), pp.84–93. Available at: <https://doi.org/10.1177/10776958221149577>.
- Paul, R., 2005. The state of critical thinking today. *New Directions for Community Colleges*, 2005(130), pp.27-38. Available at: <https://doi.org/https://doi.org/10.1002/cc.193>.
- Race, P., 2018. Is the ‘time of the assessed essay’ over? Teaching Perspectives from the Business School. Available at: <https://blogs.sussex.ac.uk/business-school-teaching/2018/11/14/is-the-time-of-the-assessed-essay-over/>.
- Rigby, D., Burton, M., Balcombe, K., Bateman, I. and Mulatu, A., 2015. Contract cheating & the market in essays. *Journal of Economic Behavior & Organization*, 111, pp.23–37. Available at: <https://doi.org/10.1016/j.jebo.2014.12.019>.
- Rudolph, J., Tan, S. and Tan, S., 2023. ChatGPT: Bullshit spewer or the end of traditional assessments in higher education?. *Journal of Applied Learning and Teaching*, 6(1). Available at: <https://doi.org/10.37074/jalt.2023.6.1.9>.
- Shen, Y., Heacock, L., Elias, J., Hentel, K. D., Reig, B., Shih, G. and Moy, L., 2023. ChatGPT and Other Large Language Models Are Double-edged Swords. *Radiology*, 230163. Available at: <https://doi.org/10.1148/radiol.230163>.
- Siddhpura, A., V. I. and Siddhpura, M., 2020. Current state of research in application of disruptive technologies in engineering education. *Procedia Computer Science*, 172, pp.494-501. Available at: <https://doi.org/https://doi.org/10.1016/j.procs.2020.05.163>.

- Stokel-Walker, C., 2022. AI bot ChatGPT writes smart essays — should professors worry? *Nature*. Available at: <https://doi.org/10.1038/d41586-022-04397-7>.
- Sum, M. and Oancea, A., 2022. The use of technology in higher education teaching by academics during the COVID-19 emergency remote teaching period: a systematic review. *International Journal of Educational Technology in Higher Education*, 19, 59. Available at: <https://doi.org/10.1186/s41239-022-00364-4>.
- Susnjak, T., 2022. ChatGPT: The End of Online Exam Integrity?. *arXiv preprint*. Available at: <https://doi.org/10.48550/arXiv.2212.09292>.
- Sweeney, S., 2023. Who wrote this? Essay mills and assessment – Considerations regarding contract cheating and AI in higher education. *The International Journal of Management Education*, 21(2), p.100818. Available at: <https://doi.org/10.1016/j.ijme.2023.100818>.
- Swiecki, Z., Khosravi, H., Chen, G., Martinez-Maldonado, R., Lodge, J.M., Milligan, S., Selwyn, N. and Gašević, D., 2022. Assessment in the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, 3, p.100075. Available at: <https://doi.org/10.1016/j.caeai.2022.100075>.
- Tate, T. P., Doroudi, S., Ritchie, D. and Xu, Y., 2023. Educational Research and AI-Generated Writing: Confronting the Coming Tsunami. Available at: <https://doi.org/10.35542/osf.io/4mec3>.
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A. and Hickey, D. T., 2023. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments* 10(1), 15. Available at: <https://doi.org/10.1186/s40561-023-00237-x>.
- Wallace, M.J. and Newton, P.M., 2014. Turnaround time and market capacity in contract cheating. *Educational Studies*, 40(2), pp.233–236. Available at: <https://doi.org/10.1080/03055698.2014.889597>.
- Whisenhunt, B. L., Cathey, C. L., Hudson, D. L. and Needy, L. M., 2022. Maximizing learning while minimizing cheating: New evidence and advice for online multiple-choice exams. *Scholarship of Teaching and Learning in Psychology*, 8(2), pp.140–153. Available at: <https://doi.org/10.1037/stl0000242>.
- Zhai, X., 2022. ChatGPT user experience: Implications for education. *arXiv preprint*. Available at: <https://doi.org/10.48550/arXiv.2212.09292>.

Have Courage to Use your Own Mind, with or without AI: The Relevance of Kant's Enlightenment to Higher Education in the Age of Artificial Intelligence

Alice Watanabe

Hamburg Center for University Teaching and Learning, University of Hamburg, Germany

alice.watanabe@uni-hamburg.de

<https://doi.org/10.34190/ejel.21.5.3229>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: Artificial intelligence (AI) in higher education is a complex issue that can be discussed from many different perspectives. There is currently a great need for ethical discussions about the use of AI in universities. For example, educational researchers and teachers are already talking a lot about fairness, accountability, transparency, bias, autonomy, agency and inclusion of AI applications, and discussing these in terms of concrete teaching-learning settings. However, less explored are the implications of AI-enhanced teaching and learning in relation to fundamental educational ideals and goals. The article takes this research desideratum as a starting point by relating the use of AI in universities to Kant's reflections on enlightenment. The aim of this article is to theoretically analyse the compatibility of various AI tools with the ideal of maturity on an educational philosophical level and to formulate recommendations for action based on the results. Through a comprehensive literature review, the article analyses the impact of intelligent tutoring systems, ChatGPT and AI-supported research tools on students' maturity and discusses both opportunities and challenges for higher education. The theoretical analysis shows that intelligent tutoring systems and ChatGPT threaten student maturity, while AI-supported research tools can have a positive effect. In addition, the analysis provides several recommendations that can help to minimise the risks of AI applications in terms of student maturity. The educational principle of research-based learning is of particular importance in this context, illustrating how students can learn to use AI tools responsibly and maturely. In this sense, the paper presents a theoretical study that fundamentally reflects on the maturity of students in the age of AI and thus both encourages teachers in the field of e-teaching to critically reflect on AI-based tools and provides a basis for further empirical research.

Keywords: Kant's enlightenment, Infantilisation, Intelligent tutor systems, ChatGPT, AI-based research and visualisation tools, Research-based learning

1. Introduction

Since the release of ChatGPT3 in November 2022, Artificial Intelligence (AI) has been a prominent topic in all higher education institutions, sparking countless discussions about how this new technology is changing teaching and learning (Farrokhnia, et al., 2023; Gimpel, et al., 2023; Lo, 2023). The use of AI has an impact on different disciplines and can be used in humanities and natural science, as well as in computer science (Crompton and Burke, 2023). Furthermore, AI in Higher Education (AIEd) is characterised by a wide range of diverse applications. This ranges from general AI-based tools (e.g., ChatGPT3.5/4) (Gimpel, et al., 2023), to specific research tools (e.g., visualisation and research tools, such as litmaps or research rabbit) (Pessin, Yamane and Siman, 2022), to intelligent tutoring systems and adaptive learning environments introduced by universities (Zawacki-Richter, et al., 2019).

AIEd can be used by students and teachers in many ways, opening research desiderata (Crompton and Burke, 2023; von Garrel, Mayer and Mühlfeld, 2023). Therefore, not only do AI-assisted teaching and learning practices present new opportunities for higher education, but also challenges and threats (Farrokhnia, et al., 2023; Lo, 2023). Alongside these opportunities and risks, ethical issues related to using AIEd are becoming increasingly recognised. There are already numerous approaches that address educational ethics and discuss AI in terms of "fairness, accountability, transparency, bias, autonomy, agency, and inclusion" (Holmes, et al., 2023). In addition to these specific topics, there are also general examinations of the impact of AI and new technologies on social life (e.g. Coeckelbergh, 2018; 2022; Friederich, 2023; Gordon and Becevel, 2021). In higher education, these broader considerations of the relationship between technology or AI use and universities are peripheral to current debates and tend to be considered sporadically (e.g., Castañeda and Selwyn, 2018; Macgilchrist, 2019; Watanabe, 2023a). However, as universities function as important public spaces, there is a need for a more intensive discussion on the extent to which AI-based teaching and learning are compatible with the fundamental goals and ideals of universities.

Taking this idea as a starting point, this article examines the relationship between Kant's educational ideal of maturity (Kant, [1784] 2013) and the use of AIEd from a theoretical perspective.¹ Kant's thoughts on maturity, included in his essay *'What is Enlightenment'*, are particularly well suited for this investigation. They have been considered a central educational goal of the university since Humboldt (Sharrock, 2010) and have lost none of their relevance (Herwald, 2022, p. 46). With this approach, the article opens a new perspective on AIEd and shows how fundamental ideals of universities can be addressed in this field. The analysis focuses on how selected AI applications present opportunities or threats to the Kantian goal of maturity through education. On this basis, the article discusses ways in which potential threats can be addressed and how potential opportunities can be further considered. The aim is to reflect on Kant's principle of maturity regarding AI-supported teaching and learning and then to identify possible recommendations for educational action. In this context, on the one hand, general considerations, such as AI-free zones, are expressed and, on the other hand, research-based learning (RBL) (Huber, 2009; Wessels, et al., 2020) is discussed as a concrete approach.

1.1 Methods and Structure

The article takes a theoretical approach and combines various empirical, historical and pragmatic references (Bellmann, 2020) to discuss the extent to which different AI applications influence students' maturity from a philosophical perspective. This interdisciplinary approach draws on different disciplines (e.g., philosophy, education, technology) and is based on a comprehensive literature review (JSTOR, ERIC, Google Scholar), which is characterised by a systematic analysis of the literature on different AI applications and uses a snowballing method to identify further relevant papers (especially regarding the debate on students' maturity). The main criteria for evaluating the literature were its significance and relevance to the research question.² Another key criterion was the quality of the articles, which could be ensured by journals, institutions or peer reviews. For the literature on the various applications of AI, currency was also an important factor, which was why no articles published prior to 2016 were used in this context. Based on this literature review, it is possible to first present the debate on maturity in higher education, derive characteristics of maturity from it, and then analyse them with different AI applications in the context of higher education. The results of the analysis can then be used to present appropriate solutions based on literature research. In addition, the educational principle of research-based learning (RBL) is applied to show how teachers can promote the maturity of students using AI tools. This consideration can be illustrated by the structure of the article: The first section begins with an explanation of Kant's ideal of maturity in the educational context and looks at the phenomena of student customer orientation and infantilisation as counter-movements. The second section introduces and classifies current applications of AIEd. Building on the first two sections, the third section examines the extent to which certain AI applications may contribute to the infantilisation and immaturity of students. Based on the findings, the fourth section discusses how the problems of AI applications can be addressed. The fifth section summarises the considerations, points out limitations, and provides an outlook for further research.

2. From The Educational Ideal Of Maturity to the Infantile Student

The educational ideal of the enlightenment originated in the 18th century and has been a central aspect of university learning and teaching ever since. The German philosopher, Immanuel Kant, described fundamental aspects of the enlightenment in his essay *'What is Enlightenment'*, published in 1784, and defined the motto of the enlightenment: "Sapere Aude! [...] Dare to know [...]. Have courage to use your own mind." (Kant, [1784] 2013, p. 2). For Kant, enlightenment means the liberation of a person from his or her "self-imposed immaturity" (Kant, [1784] 2013, p. 2). The philosopher defines self-imposed immaturity as "the inability to use one's understanding without guidance from another", where this is "self-incurred [...] when its cause lies [...] in lack [...] of resolve and courage to use it without direction from another" (Kant, [1784] 2013, p. 2). Kant sees the reasons for choosing an immature life primarily in the human traits of laziness and cowardice. They make it difficult for the individual to reach maturity, because it is more comfortable to remain dependent and uncritical. According to Kant, freeing oneself from the "yoke of immaturity" (Kant, [1784] 2013, p. 3) requires the freedom to use one's public reason. In this context, Kant distinguishes between the private and the public use of reason.

¹The article focuses on the impact of AI on student maturity and aims to examine it in a deep analytical discourse. Other ethical issues such as fairness, accountability, transparency, bias, autonomy, agency, and inclusion are therefore not addressed. For an overview of ethical aspects of AI in higher education, see e.g. Holmes, et al. (2023).

²As the article asks to what extent AI applications influence the maturity of students in general, rather than focusing on specific students (e.g., undergraduate students, discipline-specific), conceptual or theory-based papers and reviews proved particularly helpful as references. Specific empirical studies were also considered but were less central.

Consequently, freedom of reason can be restricted in the private sphere, while it must always be present in the public sphere. Thus, in Kant's view, it is possible for someone to be immature in his professional role (e.g., as an officer, civil servant or clergyman) and to obey orders, while in his role as a mature member of society, he critically questions and discusses issues (Kant, [1784] 2013). Enlightenment is a social process of educating people to think and act independently, challenging political doctrines with the aim of social change (Sharrock, 2010). Following Kant, public institutions, such as universities, have the task of promoting the freedom of reason (Hoyer, 2006).

Kant's thoughts on the enlightenment were taken up by Wilhelm von Humboldt who formed a pillar of his ideal of university education (Sharrock, 2010), which focuses on, among other things³, freedom through education. Therefore, education liberates people from their immaturity by eliminating and recognising heteronomy (Herwald, 2022, p. 45; Williams, 2016). Both theorists see higher education as a "public good" and attribute to universities the function of playing "a critical role in holding state bodies and the professions to account" (Williams, 2016, p. 619). Kant's and Humboldt's reflections on the enlightenment represent fundamental ideals of universities (Côté and Allahar, 2011, p. 11) and have become increasingly important due to current social dangers, such as fake news (Herwald, 2022, p. 46). How fundamentally the principle of maturity is anchored in the university sector becomes apparent in line with the goals of the European Union regarding higher education. Thus, along with career preparation and academic education, the European Union considers personality development and active citizenship a key objective for higher education (Camilleri, et al., 2014). In this context, critical thinking⁴ and the ability to independently evaluate topics or situations play an essential role (van Damme and Zahner, 2022; Wissenschaftsrat, 2015), illustrating the way in which Kant's enlightenment approach still resonates in the university today.

However, there are currently some countervailing trends in higher education that raise the question of whether Kant's reflections are still relevant in contemporary university life. For example, Macheridis, Paulsson and Pihl (2020) point to an increasing market orientation in higher education institutions, with a focus on employability rather than critical thinking or the development of maturity. In this context, students are often viewed in terms of customer orientation. Thus, higher education institutions are increasingly courting students, who progressively see themselves as customers of the educational offer (Guilbault, 2016; Calma and Dickson-Deane, 2020). This development leads, for example, to students expecting a good grade regardless of their performance or feeling less responsible for failures, instead blaming the system or teachers (Clayson and Haley, 2005; Guilbault, 2016).

As a result:

"students seek to secure a degree, rather than experience an education, with their goals limited to the acquisition of skills needed for employment and maximizing income. Within this context, universities are being pushed to produce knowledgeable students that society and employers deems valuable—not knowledge for its own sake or classical approaches that focus more on the process and ability to think" (Natale and Doran, 2012, p. 188).

Besides the tendency for universities to be market-oriented and the associated view of students as customers, the phenomenon of infantilisation can also be discussed (Giroux, 2015). A general definition of infantilisation is provided by Levi (1943, p. 53) in his work *Maternal Overprotection*: "Infantilization also refers to the continuity of the same type of care ordinarily modified in later years." Both Levi (1943) and Postman (1982) point out that infantilisation is no longer confined to individuals, but rather a growing trend in society as a whole. Stiegler (2010, p. 2) also deals with the infantilisation of society and comes to a similar conclusion: "[U]ndermining the difference between minors and adults is at the very heart of contemporary consumer culture, which systematically defines consumers, minors and adults alike, as being fundamentally, structurally irresponsible."⁵

³Other aspects of the Humboldtian ideal of education are, for example "[t]he unity of research and education/teaching [...] [,] [t]he primacy of research [...] [or] [a] national culture dominated and distinguished by higher learning" (Nybom, 2003, p. 144).

⁴Critical thinking is understood in this context as: "the intellectually disciplined process of actively and skilfully conceptualizing, applying, analysing, synthesizing, and/or evaluating information gathered from or generated by observation, experience, reflection, reasoning, or communication as a guide to belief and action" (Stanford Encyclopedia of Philosophy, 2018).

⁵A particular feature of Stiegler's work is that in many places he equates infantilisation with Kant's account of immaturity (Keij, 2021), underlining the topicality of the enlightenment agenda.

In the university context, infantilisation can be seen, for example, in the aforementioned customer orientation (Giroux, 2015), where students expect universities to guide them through their studies as quickly and directly as possible through teaching and service provision (Natale and Doran, 2012). Another indication of students' immaturity is the active role of parents in their adult children's studies (Vinson, 2012), which is reflected, for example, in the introduction of so-called parents' days (Jornitz, 2012). In addition, the decreasing ability to think independently and critically (Arum and Roksa, 2011) or the lack of willingness to read (Hoeft, 2012) are signs of the increasing immaturity of students. From the perspective of these infantilisation tendencies, students may well "fall into the role of being a passive vessel to be filled with information and/ or consumers entitled to high marks and credentials in exchange for tuition fees; ultimately, in either case, standards slide" (Côté and Allahar 2011, p. 90).

This brief description already shows the extent to which the educational ideal of student maturity can be renegotiated and discussed in a perspective-rich way. The use of new technologies, such as AI, now raises the additional question of how AI-assisted teaching and learning will affect students' personal development, critical thinking and questioning, as well as their independence. With the help of Kant's fundamental considerations on aids, this question can be made concrete. According to Kant, an aid (such as a book or, in this case, an AI-based tool) can either promote people's maturity or have precisely the opposite effect and cause people not to use their own reason. Kant illustrates these considerations in the concept of *pharmakon*, which in the medical field represents a drug that influences people positively or negatively (Stiegler, 2008, p. 44). Using the example of a book as an aid, this double meaning becomes visible:

"[T]he book reveals itself as such a *pharmakon*, potentially remedy and poison. Maturity is the result when the book acts as a remedy of the mind's minority, while infantilization or adult immaturity is the result of a poisoning of the mind, which occurs when the book takes over the role of the understanding rather than being a necessary aspect of critical thinking" (Keij, 2021, p. 71).

In terms of AI-supported teaching and learning, the central question is whether the use of AIED will reinforce the trend toward customer orientation and the infantilisation of students or, on the contrary, contribute to the development of empowerment. Before exploring this question in the next chapter, three key elements of mature students are summarised. Firstly, maturity is characterised using the student's own mind without outside help. Secondly, mature students can deal with issues and situations independently and critically. Thirdly, they are able to present their positions publicly as active members of society.

3. AI Applications and Their Impact on Student Maturity

AI has a great disruptive potential for higher education and will strongly influence the future of university teaching and learning (Schön, et al., 2023). In this context, AIED are characterised by different functions and a variety of possible uses (Watanabe, 2023a). While the use of AI was initially characterised by the development and research of assessment and feedback tools, early warning systems, intelligent tutors or adaptive learning environments (Crompton and Burke, 2023; Zawacki-Richter, et al., 2019), educational researchers and teachers are now showing increasing interest in commercial AI-based tools such as ChatGPT (Gimpel, et al., 2023; Lo, 2023; Schön, et al., 2023). In contrast to the first type of application, these are not designed by the university itself, but enter the higher education institutions from the outside. This second type of AI tool can be further divided into two groups. On the one hand, there are tools that are aimed exclusively at research and teaching, such as AI-based literature search applications (e.g., ResearchRabbit, Litmaps). On the other hand, a wide variety of AI tools are currently emerging that are used in different areas of work (e.g., marketing, journalism, sales, graphic design, etc.) (Limburg, et al., 2022): The most famous example of this type of AI are currently conversational agents, such as ChatGPT (Gimpel, et al., 2023). In order to make a meaningful analysis of whether AI applications help or hinder students' maturity, it is important to consider the different ways in which AI can be used in universities. For this reason, the following research will consider all three AIED-types individually, defining and discussing them in terms of maturity characteristics. However, as the different types can have further sub-categories, it seems to be useful to work with selected examples. This allows specific investigation of how particular AI applications affect student maturity. Specific examples used in the analysis are intelligent tutoring systems (ITS), Open AI's ChatGPT as a conversational agent, and AI-based research and visualisation tools.

3.1 Intelligent Tutor Systems and Student Maturity

ITS are receiving much attention in the current debate on the use of AIED (Crompton and Burke, 2023; Wang, et al., 2023). Zawacki-Richter, et al. (2019) identify ITS as a growing AI trend in higher education already in 2019 and show that ITS are grouped under different terms (e.g. intelligent (online) tutors, intelligent tutoring systems, intelligent (software) agents or intelligent assistants). A general definition of ITS is given by Wang, et al. (2023), who refer to Nwana (1990, p. 252): thus, an ITS is a computer system that is “designed to incorporate techniques from the artificial intelligence (AI) community in order to provide (intelligent) tutors which know what they teach, who they teach, and how to teach it.” ITS are mainly used in the seminar setting, for example, to deliver course content, to analyse the strengths and weaknesses of students, or to create personalised learning materials and automated feedback. In this context, educational researchers see personalisation as an opportunity for learners to receive tailored learning paths and recommendations (Crompton and Burke, 2023; Wang, et al., 2023; Zawacki-Richter, et al., 2019).

A comparison of ITS and Kant's thoughts on maturity shows an obvious incompatibility: Kant advocates that students use their minds independently, evaluate situations and take responsibility for their own actions. ITS, instead, reinforce the tendency for students to become infantilised and increasingly rely on the automated recommendations of a system (Utterberg Modén, et al., 2021; Watanabe, 2023b). The main problem here is the goal of ITS to personalise learning paths and materials for students, thus making studying more pleasant and comfortable (Çakir, 2019). This is because, according to Kant, critical thinking and the resulting independent action is a strenuous activity that takes time and cannot be replaced by accepting the views of others (in this case, the view and evaluation of an ITS) (Kant, [1784] 2013; Stiegler, 2008, p. 39f.).

At this point the question can be raised whether ITS should not be seen as a supportive tool that, like teachers, helps students to reflect and thus supports them in maturing. There are two arguments against this: First, ITS are not designed to enable students to think and act independently, as Murphy (2019, p. 6) points out:

“However, the systems are less able to support the learning of complex, difficult-to-assess, higher order skills — such as critical thinking, effective communication, explanation, argumentation, collaboration, self-management, social awareness, and professional ethics — that are increasingly emphasized in state education standards and valued by employers.”

Second, because ITS are AI applications whose choices are based on machine calculations (Kurni, Mohammed and K G, 2023, p. 33), students cannot engage in an open exchange with ITS (as they can with teachers) and cannot discuss or question the decisions made by the system. This problem is exacerbated by the so-called black box phenomenon of AI applications, which makes it impossible in principle to understand the decisions of an AI system (Adadi and Berrad, 2018; Herzberg, 2023).

Another problem with ITS in relation to learning is that it focuses students more and more on their individual learning success and does not create a collective learning atmosphere with other students; “this means that learning is perceived as individual change in incremental steps rather than as a collaborative process taking place in a learning community” (Utterberg Modén, et al., 2021, p. 1540). ITS thus contradict the enlightened claim of universities to educate students to become responsible and committed citizens of society. Instead, they promote the private interests and self-optimisation of students. The comparison of Kant's principle of maturity and the ITS makes it clear that the ITS infantilise students and prevent them from deciding for themselves under what conditions and in what way they want to learn.

3.2 ChatGPT and Student Maturity

ChatGPT is an AI-based conversational agent developed by Open AI that enables human-machine interaction using natural language software. While the first versions of ChatGPT and other conversational agents received less public attention, the quality of ChatGPT 3.5 and 4 has greatly improved, and the conversational agent now produces convincing results in various areas. In particular, ChatGPT is becoming increasingly important in university teaching, as this application can summarise knowledge or create analyses according to certain criteria without much human intervention: This means that central tasks in the teaching-learning context can be produced automatically (Gimpel, et al., 2023). In the field of higher education, this new technology has led to a variety of debates and has shown its disruptive potential (Farrokhnia, et al., 2023; Schön et al., 2023). Thus, the reactions of teachers and educational researchers to ChatGPT have ranged from fascination to concern (Rudolph, Tan and Tan, 2023). While critics of the new technology are primarily alarmed about the potential for misuse of ChatGPT, plagiarism and academic integrity (e.g., Sullivan, Kelly and McLaughlan, 2023), proponents

see opportunities in the use of conversational agents, hoping for example to strengthen creativity and writing skills or personalised learning (e.g. Atlas, 2023).

Regarding Kant's thoughts of enlightenment, it should first be noted that the philosopher considers writing and reading to be crucial skills for a mature life (Stiegler, 2008, p. 38). Independent writing plays an important role in this context, since a writing process is always accompanied by an independent thinking process (Luhmann, 1981, p. 222; Sumarno Kusumaningrum and Nurhayati, 2022). In text production using ChatGPT, this thinking has less room, since it is, only reflected in the formulation of the correct instructions to the conversational agent and the subsequent revision of the text (Watanabe, 2023a). The reduced writing experience thus has a generally negative effect on independent thinking and maturity.

It could be argued that it is precisely by processing the generated texts that students learn to deal with them critically, question their content, and thus acquire maturity in their studies. This consideration is supported above all by the fact that ChatGPT does not generate new knowledge, does not present information in its complexity and depth, and does not falsify sources. As a result, automatically generated texts always require a critical examination of the topics and sources presented in them (Gimpel, et al., 2023; Schön, et al., 2023). In this context, the question arises to what extent students (especially at the undergraduate level) can critically examine automatically generated texts that have a higher linguistic quality than texts they have written themselves. This consideration becomes even more problematic when it is assumed that students have less writing experience due to using ChatGPT and that academic writing is a general challenge in studying (e.g. Fernsten and Reda, 2011).

Therefore, it seems that conversational agents can only be meaningfully used in academic writing if students, due to their writing experience and academic socialisation, are already able to write academic essays independently and to interrogate automatically generated texts. To expect students with little or no writing experience to critically engage with an automatically generated text would be like showing a highly complex calculation to a primary school pupil who has just mastered simple mathematics and expecting him or her to be able to judge the correctness of that calculation.

It is also important to bear in mind that students must have a basic interest in editing the automatically generated text. They need to understand that this type of text generation is not a shortcut in the preparation of a written paper, but also requires a lot of time and work. However, the study by von Garrel, Mayer and Mühlfeld (2023) paints a different picture: students want to use ChatGPT to produce a text that meets academic standards, but they have little interest in the actual process of creating and actively shaping the text. In other words, students see ChatGPT as a work relief and not to practise critical thinking and writing. Consequently, the conversational agent is a tool that serves comfort but does not promote students' maturity.

In addition to writing texts, ChatGPT is often used by students to have other texts explained to them or to create summaries (Gimpel, et al., 2023; von Garrel, Mayer and Mühlfeld, 2023). Similar to writing, ChatGPT can be seen as an enemy of students' maturity, because if students do not understand a text and then have it explained or summarised by the conversational agent, they cannot subsequently check the extent to which the ChatGPT analysis is convincing. They remain in a state of immaturity and do not challenge themselves. The same applies if students only use this function to save time or reduce the amount of work, because, according to Kant, they are subject to laziness (Kant, [1783] 2013), which prevents them from engaging with a challenging text and forming their own opinion.

3.3 AI-Supported Research and Visualisation Tools and Student Maturity

The search for literature and the presentation of the research state are core tasks of academic activity (Ponsard, Escalona and Munzner, 2016). Bibliometric techniques and science mapping have supported these academic activities since their beginnings in the 1930s. Through new technologies like AI, many tools (e.g., ResearchRabbit, Litmaps, Elicit) are emerging in this area to help researchers and students get a faster and better overview of the research literature (Pessin, Yamane and Siman, 2022): this allows researchers to search and select articles according to specific criteria (title, author, subject, DOI, etc.) and then obtain a map showing how the different sources and references relate to each other. Selected articles can be also stored and filtered in a digital library. In addition, it is possible to use these tools to find out about new publications or to share the results of research with others (Kaur, et al., 2022; Sharma, et al., 2022). Pessin, Yamane and Siman (2022, p. 3716) see these tools as a great opportunity for academic work:

"With this system, scientists will have a powerful tool that provides strategic information, increasing the chances of success in publications by identifying significant research gaps and contribute to an open

science. The expected result is a solution that simplifies researchers' work, expands assertiveness in scientific production, and contributes to knowledge development."

In contrast to ITS and ChatGPT, research and visualisation tools support students in independent research projects and thus promote maturity. These new technology-enabled tools allow even inexperienced researchers (e.g. students at the beginning of their bachelor's degree) to get an overview of the research landscape, to understand connections and to identify research gaps. While the tools take some of the research work off the student's plate, the main work (reading and evaluating the various papers) still needs to be done by the students themselves. Particularly in view of the enormous number of publications now available, Kaur, et al. (2022, p. 7) speak of literature searches in the "ocean of research-based database resources", while Ponsard, Escalona and Munzner (2016, p. 2264) suggest that researchers may be confronted with "hundreds of publications" when researching topics, the importance of these tools increases in relation to the maturity of the students. The sheer volume of articles can therefore overwhelm students, either leading to paralysis and reluctance to search the literature, or severely limiting the search, by focusing on specific authors or terms and overlooking other relevant sources. AI-supported research and visualisation tools can counteract these reactions and make it easier for students to engage with academic papers.

In relation to the maturity characteristics outlined above, one could ask whether technology-enhanced research and visualisation tools promote student comfort. However, this must be answered in the negative, as students first need to familiarise themselves with the software and then still have to identify the relevance of papers to their work. The research and visualisation tools only provide new access to the literature; the student is still responsible for dealing with the material. Thus, these AI applications differ from ChatGPT primarily in the fact that they only support students in a specific aspect (e.g., the overview of the literature), which is not a usable product without further processing. Research and visualisation tools can be then seen, in the Kantian sense, as tools that empower students and give them access to relevant and/or alternative literature. Ultimately, students should not focus on these digital research tools, but rather see them as complementary to traditional sources of literature, such as seminar bibliographies or library catalogues.

3.4 From AI-Free Zones to Research-Based Learning

The examination of the three AI applications through the lens of the enlightenment shows that ITS and conversational agents do not meet Kant's ideal of maturity, while research and visualisation tools can help students to make independent inquiries.

Regarding ITS, it seems especially important that these systems do not place too much emphasis on teaching-learning settings, but rather represent one of many support services in higher education. In this context, particular attention should be paid to ensuring that universities create alternative learning moments in which the development of students' maturity and the learning of independent thought and action are in the spotlight. The proposal by Watanabe and Schmohl (2022) to create technology-free zones in educational settings could help to strike a balance between the efficient and personalised study on the one hand and the education of mature and critically thinking students on the other. Exactly how these AI-free zones are to be designed must be left open at this point, as it depends on specific teaching-learning settings and cannot be generalised. Furthermore, it is important that students can discuss ITS assessments or material suggestions with teachers and are supported in taking a critical stance toward the system.

For the use of ChatGPT, it is also essential for the maturity of the students to create writing and reading moments in which no intelligent technologies are used. Such moments would be characterised, for example, by writing down thoughts on paper without any AI-assisted writing aids, as well as reading classic literature. In addition, students need to familiarise themselves with the general principles and logic of AI tools and understand how they work, as already called for by Gimpel, et al. (2023). To promote critical and independent thinking, there should also be space for general reflection, where students discuss with peers and teachers the use of ChatGPT and possible consequences for independent writing and reading.

However, it is clear from these considerations that teaching and learning with ITS or ChatGPT means a considerable amount of additional work for teachers, not only relieving them of tasks but also creating new ones. Against this background, it seems particularly important for teachers to develop new learning strategies and examination formats for the use of AI applications, on the one hand, and to be informed about the rapidly developing possibilities of AI, on the other hand, to adapt their teaching and examinations accordingly (Gašević, Siemens and Sadiq, 2023; Rudolph, Tan and Tan, 2023). In order to meet these broad educational requirements,

support is needed at this point through "extensive involvement of researchers, technology developers and policy makers" (Gašević, Siemens and Sadiq, 2023, p. 3).

In contrast to the other two AI applications, the analysis of the literature search and visualisation tools shows that certain technology-supported tools can support students in their independent research and thus promote their academic maturity. In this context, the research tools are particularly convincing because they only provide support for a very specific aspect (the presentation of literature) within a complex task (the elaboration and summary of the state of research). Students thus receive support, but not a finished product, and must evaluate and process the automatically generated information themselves. The implication of this finding is that, in terms of developing independent thinking and research, students should use AI tools primarily for very specific subtasks that require further processing by the learners. Thus, teaching and learning formats should be characterised by complexity, in which students must master diverse and overlapping tasks. The next section shows how these requirements can be implemented in a concrete teaching-learning setting, using research-based learning as an example. Before that, Table 1 summarises the considerations discussed here.

Table 1: AI applications in relation to Kant's ideal of maturity (own figure)

AI Application	Impact of AI Application on Student Maturity	Resulting recommendations
Intelligent Tutor Systems	ITS deprive students of independent assessment and reflection on their learning. Students cannot discuss ITS suggestions or assessments, which discourages critical inquiry. ITS tend to make students more focused on their own success in their studies and less focused on collaborative learning.	ITS should not be widely used in every teaching and learning environment, but should be complemented by AI-free zones. Teachers and students need to discuss and challenge ITS suggestions and assessments together.
ChatGPT	ChatGPT limits students' reading and writing experiences and prevents students from learning to write and evaluate texts independently. ChatGPT promotes student comfort and results in students being less likely to think, read or write independently.	Students should continue to experience writing and reading moments in their studies where AI technology is not used. There should be courses or workshops in which students are taught how ChatGPT techniques work. Teachers and students need to discuss in teaching-learning settings how the use of conversational agents affects students' maturity.
AI-based research and visualization tools	AI-supported research and visualization tools help students to conduct independent research without doing the main work for them. This promotes the development of maturity.	Students can gain further/ alternative access to the state of research and literature by using AI-based research and visualization tools. AI-based tools should be used for specific aspects of complex tasks.

Case Study: Research-based Learning and AI

Research-based learning (RBL) is a well-known and established learning format in which students carry out their own research projects (Wessels, et al., 2020). According to Huber (2009, p. 11), in RBL students independently go through all phases of a research process and develop results that are relevant for others. Depending on various parameters (e.g., subject, semester, course format), the concrete implementation of this learning format varies greatly, but it is important that students experience the entire research cycle cognitively, emotionally, and socially (Huber, 2009, p. 12). An ideal research cycle in RBL consists of eight different phases and is shown in Figure 1.

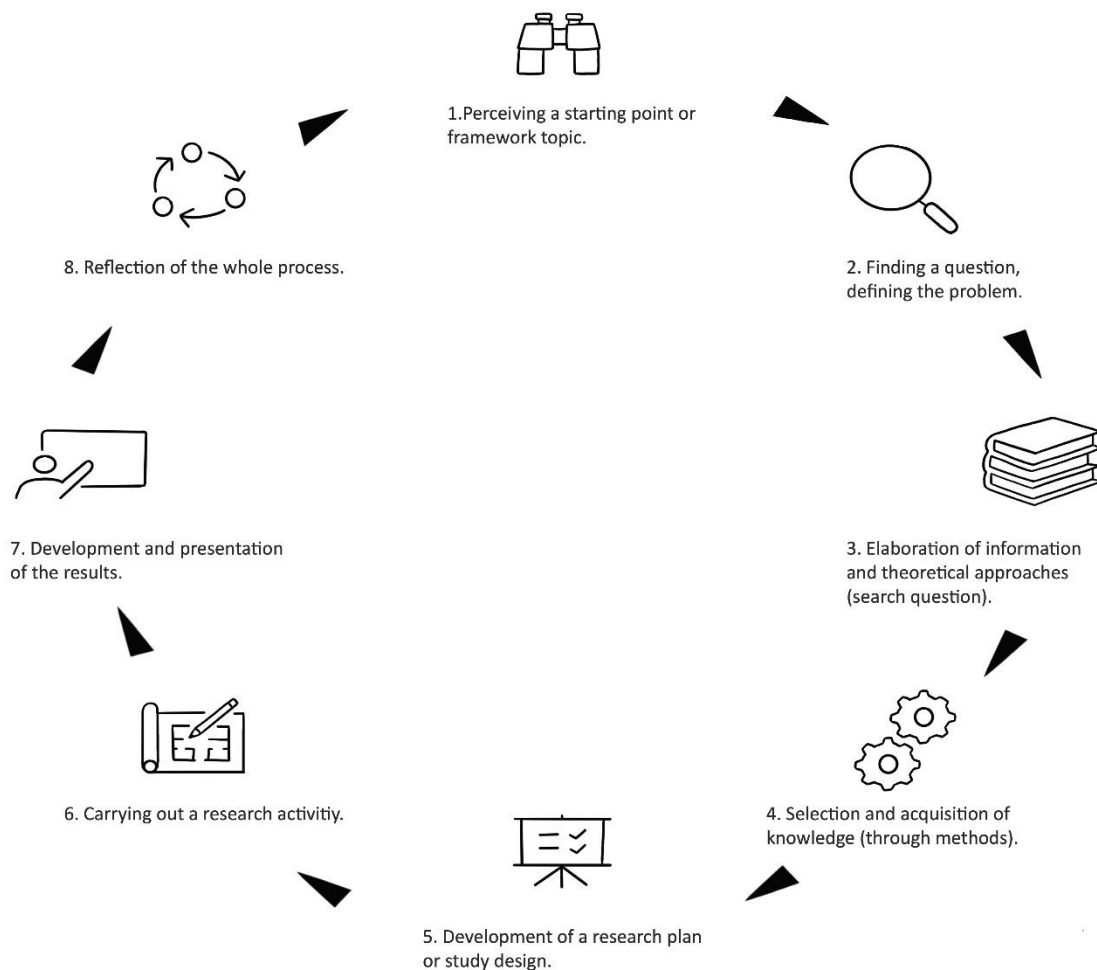


Figure 1: Ideal-typical research process according to Huber and Reinmann (2019, p. 92) (own figure)

RBL is considered an effective teaching format in higher education and is used in various disciplines, which is why Wessels, et al. (2020, p. 1) also describe RBL as a "panacea" in higher education teaching. This also seems to be the case for mature and independent student learning with AI-based tools. The complexity of RBL therefore ensures that students can only use AI-based tools selectively at different stages of their research. In RBL, students must be able to understand and link the individual research phases and their significance for the overall process. Otherwise, they will not be able to successfully complete the cycle and arrive at a coherent result. Accordingly, the research and visualisation tools presented before can be used very well to support RBL. In contrast, it is of little value for students to work with conversational agents, as these AI tools cannot adapt to specific and individual research. For example, it is of little use for students to have a text automatically generated

about an empirical method, because without independent engagement with this method, they will not be able to apply it correctly in their later research. Accordingly, RBL can help students to try out different AI-supported tools in the research process and to question them in relation to independent research or to stimulate an AI discourse on academic work between teachers and students. In this sense, RBL prevents misuse and lack of independent work among students. In terms of maturity, the students' own experience is particularly important here: during RBL, they practise first-hand how certain AI applications can either only produce superficial and unusable products or support the self-organised research process. This experience forms the basis for students to consciously decide for or against the use of selected AI tools without external guidance. In this way, and in the spirit of Kant, they begin to critically engage with new technologies, to question them and to deal with the possibilities in a mature and responsible way.

4. Conclusion

Based on Kant's reflections on enlightenment, the article analyses the extent to which selected AI applications have an impact on students' maturity. It shows that ITS and conversational agents tend to hinder the formation of critical thinking and independent action, while AI-based research and visualisation tools can have a positive impact on students' maturity. As a possible recommendation, AI-free zones in studies are presented, which create a balance between personalised learning on the one hand and the training of independence on the other hand in studies. The example of RBL also shows that complex teaching-learning settings in which students go through a complete research process can lead to a reflective and critical approach to AI-supported tools.

The article raises awareness of the dangers of AI-based tools and outlines possible solutions to ensure the maturity of students in the age of AI. To consider whether and to what extent AI tools pose a risk to the development of students' maturity in specific classroom settings, teachers can, as a first step, ask themselves the following reflective questions:

- How does an AI application prevent or enable students in my course to use their own minds without the help of others?
- How does an AI application avoid or help students in my course to analyse situations independently and critically?
- How does the AI application hinder or encourage students in my course to express their opinions publicly as active members of society?

If teachers then conclude that the use of AI applications can pose a threat to maturity, they can then consider whether any of the proposed solutions are appropriate to address this issue and/or whether the learning content can be linked to the RBL principle.⁶

In addition to these initial critical questions and the recommendation for action, further empirical studies are needed to investigate the design and implementation of AI-free zones to promote maturity or the link between RBL and AI-supported learning. Further research is also required on ethical issues in AIED (such as fairness, accountability, transparency, bias, autonomy, agency and inclusion), which may also be related to student maturity. It should be noted that the use of ITS or ChatGPT is only critiqued here from the perspective of Kantian maturity. In other settings or for other purposes, these AI applications may well be useful for higher education. It is therefore important to continue to explore theoretically the impact of AI on specific university goals and the ways in which students experience their studies with and without AI.

Acknowledgements

I would like to thank the Open Access Publication Fund of Universität Hamburg for their financial support.

References

- Adadi, A. and Berrada, M., 2018. Peeking inside the black-box: a survey on explainable artificial intelligence. *IEEE Access*, [e-journal] 6, pp. 52138–52160. <http://dx.doi.org/10.1109/ACCESS.2018.2870052>
- Arum, R. and Roska, J., 2011. *Academically adrift: limited learning on college campuses*. Chicago: University Press.

⁶It should be noted that the study has a general orientation due to its educational-theoretical approach. Therefore, this paper should also be seen as a stimulus for further reflection and for teachers to consider the extent to which AI-based tools hinder or support students' critical and independent thinking and action in their courses. The same applies to the recommendations presented here. They are suggestions that need to be evaluated and concretised by teachers and educational researchers in relation to a specific university context.

- Atlas, S., 2023. *ChatGPT for higher education and professional development: a guide to conversational AI*. University of Rhode Island. [online]. Available at: <https://digitalcommons.uri.edu/cba_facpubs/548> [Accessed 11 November 2023].
- Bellmann, J., 2020. Theoretische forschung. unterscheidung und bezeichnung eines spezifischen modus der wissensproduktion. *Zeitschrift für Pädagogik*, [e-journal] 66(6), pp. 788–806. <http://dx.doi.org/10.25656/01:25813>
- Çakir, R., 2019. Effect of web-based intelligence tutoring system on students' achievement and motivation. *Malaysian Online Journal of Educational Technology*, [e-journal] 7(4), pp. 45–59. <http://dx.doi.org/10.17220/mojet.2019.04.004>
- Calma, A. and Dickson-Deane, C., 2020. The student as customer and quality in higher education. *International Journal of Educational Management*, [e-journal] 34(8), pp. 1221–1235. <http://dx.doi.org/10.1108/IJEM-03-2019-0093>
- Camilleri, A. F., Delplace, S., Frankowicz, M., Hudak, R. and Tannhäuser, A. C., 2014. *Professional higher education in Europe: characteristics, practice examples and national differences*. Malta: Knowledge Innovation Centre. [online] <http://dx.doi.org/10.25656/01:10873>
- Castañeda, L. and Selwyn, N., 2018. More than tools: making sense of the ongoing digitizations of higher education. *International Journal of Educational Technology in Higher Education*, [e-journal] 15, pp. 1–10. <https://doi.org/10.1186/s41239-018-0109-y>
- Clayson, D. E. and Haley, D. A., 2005. Marketing models in education: students as customers, products, or partners. *Marketing Education Review*, [e-journal] 15(1), pp. 1–10. <http://dx.doi.org/10.1080/10528008.2005.11488884>
- Coeckelbergh, M., 2018. Technology and the good society: a polemical essay on social ontology, political principles, and responsibility for technology. *Technology in Society*, [e-journal] 52, pp. 4–9. <http://dx.doi.org/10.1016/j.techsoc.2016.12.002>
- Coeckelbergh, M., 2022. Democracy, epistemic agency, and AI: political epistemology in times of artificial intelligence. *AI and Ethics*, [e-journal], pp. 1–10. <http://dx.doi.org/10.1007/s43681-022-00239-4>
- Côté, J. and Allahar, A. L., 2011. *Lowering higher education: the rise of corporate universities and the fall of liberal education*. Toronto: University Press.
- Crompton, H. and Burke, D., 2023. Artificial intelligence in higher education: the state of the field. *International Journal of Educational Technology in Higher Education*, [e-journal] 20(1), pp. 1–22. <http://dx.doi.org/10.1186/s41239-023-00392-8>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O. and Wals, A., 2023. A SWOT analysis of chatGPT: implications for educational practice and research. *Innovations in Education and Teaching International*, [e-journal] pp. 1–15. <http://dx.doi.org/10.1080/14703297.2023.2195846>
- Fernsten, L. A. and Reda, M., 2011. Helping students meet the challenges of academic writing. *Teaching in Higher Education*, [e-journal] 16(2), pp. 171–182. <http://dx.doi.org/10.1080/13562517.2010.507306>
- Friederich, S., 2023. Symbiosis, not alignment, as the goal for liberal democracies in the transition to artificial general intelligence. *AI and Ethics*, [e-journal], pp. 1–10. <http://dx.doi.org/10.1007/s43681-023-00268-7>
- Gašević, D., Siemens, G., & Sadiq, S., 2023. Empowering learners for the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, [e-journal] 4, 100130. <https://doi.org/10.1016/j.caeai.2023.100130>
- Gimpel, H., Hall, K., Decker, S., Eymann, T., Lämmermann, L., Mädche, A., Röglinger, M., Ruiner, C., Schoch, M., Schoop, M., Urbach, N. and Vandrik, S., 2023. *Unlocking the power of generative AI models and systems such as GPT-4 and chatGPT for higher education: a guide for students and lecturers*. Stuttgart: Universität Hohenheim, Fakultät Wirtschafts- und Sozialwissenschaften. [online]. Available at: <<http://opus.uni-hohenheim.de/volltexte/2023/2146/>> [Accessed 11 November 2023].
- Giroux, H. A., 2015. Democracy in crisis, the specter of authoritarianism, and the future of higher education. *Journal of Critical Scholarship on Higher Education and Student Affairs*, [online] Available at: <<https://ecommons.luc.edu/jcshesa/vol1/iss1/7/>> [Accessed 11 November 2023].
- Gordon, D. and Becevel, A., 2021. Are we in the digital dark times: how the philosophy of Hannah Arendt can illuminate some of the ethical dilemmas posed by modern digital technologies. In: *EthiComp 2021*. La Rioja, Spain, 30 June – 2 July, 2021. Logroño: Universidad de La Rioja.
- Guilbault, M., 2016. Students as customers in higher education: reframing the debate. *Journal of Marketing for Higher Education*, [e-journal] 26(2), pp. 132–142. <http://dx.doi.org/10.1080/08841241.2016.1245234>
- Herwald, H., 2022. *Agenda 2030-bildung: wertevermittlung und wertorientierung*. Berlin/Heidelberg: Springer.
- Herzberg, D., 2023. Künstliche intelligenz in der hochschulbildung und das transparenzproblem: eine analyse und ein lösungsvorschlag. In: T. Schmohl, A. Watanabe and K. Schelling, eds. 2023. *Künstliche intelligenz in der hochschulbildung: chancen und grenzen des KI-gestützten lernens und lehrens*. Bielefeld: Transcript. pp. 87–98.
- Hoeft, M. E., 2012. Why university students don't read: what professors can do to increase compliance. *International Journal for the Scholarship of Teaching and Learning*, [e-journal] 6(2), pp. 1–19. <http://dx.doi.org/10.20429/ijsotl.2012.060212>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I. and Koedinger, K. R., 2022. Ethics of AI in education: towards a community-wide framework. *International Journal of Artificial Intelligence in Education*, [e-journal] 32(3), pp. 504–526. <http://dx.doi.org/10.1007/s40593-021-00239-1>
- Hoyer, T., 2006. Erziehungsziel mündigkeit: eine problemgeschichtliche skizze. In: H. Eidam and T. Hoyer, eds. 2006. *Ethik und pädagogik im dialog: Vol. 4. erziehung und mündigkeit. bildungsphilosophische studien*. Münster: LIT. pp. 9–32.

- Huber, L., 2009. Warum forschendes lernen nötig und möglich ist. In: L. Huber, J. Hellmer and F. Schneider eds. 2009. *Forschendes lernen im studium: aktuelle konzepte und erfahrungen*. Bielefeld: UVW. pp. 9–35.
- Huber, L. and Reinmann, G., 2019. *Vom forschungsnahen zum forschenden lernen an hochschulen: wege der bildung durch wissenschaft*. Berlin/Heidelberg: Springer.
- Jornitz, S., 2012. Elterntage an universitäten. *Pädagogische Korrespondenz*, 45, pp. 105–113.
- Kant, I., 1784. *Answer the question: what is enlightenment*. Translated by Fidel Ferrer, 2013. [online] Available at <https://www.researchgate.net/publication/305639298_Answer_the_question_What_is_Enlightenment_by_Immanuel_Kant_Translator_Daniel_Fidel_Ferrer> [Accessed 16 August 2023].
- Kaur, A., Gulati, S., Sharma, R., Sinhababu, A. and Chakravarty, R., 2022. Visual citation navigation of open education resources using litmaps. *Library Hi Tech News*, [e-journal] 39(5), pp. 7–11. <http://dx.doi.org/10.1108/LHTN-01-2022-0012>
- Keij, D., 2021. Immature adults and playing children: on Bernard Stiegler's critique of infantilization. *Studies in Philosophy and Education*, [e-journal] 40(1), pp. 67–80. <http://dx.doi.org/10.1007/s11217-020-09742-9>
- Kurni, M., Mohammed, M. S. and K G, Srinivasa., 2023. *A beginner's guide to introduce artificial intelligence in teaching and learning*. Cham, Switzerland: Springer.
- Levi, D., 1943. *Maternal overprotection*. Columbia: University Press.
- Limburg, A., Salden, P., Mundorf, M. and Weßels, D., 2022. Plagiarismus in zeiten künstlicher intelligenz. *Zeitschrift für Hochschulentwicklung*, [e-journal] 17(3), pp. 91–106. <http://dx.doi.org/10.3217/zfhe-17-03/06>
- Lo, C. K., 2023. What is the impact of chatGPT on education: a rapid review of the literature. *Education Sciences*, [e-journal] 13(4), 410. <http://dx.doi.org/10.3390/educsci13040410>
- Luhmann, N., 1981. Kommunikation mit zettelkästen: ein erfahrungsbericht. In: H. Baier, H.M. Kepplinger, K. Reumann and E. Noelle-Neumann, eds. 1981. *Öffentliche meinung und sozialer wandel: public opinion and social change*. Opladen: Westdeutscher Verlag für Sozialwissenschaften, pp. 222–228.
- Macgilchrist, F., 2019. Cruel optimism in edtech: when the digital data practices of educational technology providers inadvertently hinder educational equity. *Learning, Media and Technology*, [e-journal] 44(1), pp. 77–86. <http://dx.doi.org/10.1080/17439884.2018.1556217>
- Macheridis, N., Paulsson, A. and Pihl, H., 2020. The Humboldtian ideal meets employability: university teachers and the teaching–research relationship in marketized higher education. *Industry and Higher Education*, [e-journal] 34(5), pp. 303–311. <http://dx.doi.org/10.1177/0950422219898371>
- Murphy, R. F., 2019. Artificial intelligence applications to support K-12 teachers and teaching. *Rand Corporation*, [e-journal] 10, pp. 1–20. <http://dx.doi.org/10.7249/PE315>
- Natale, S. M. and Doran, C., 2012. Marketization of education: an ethical dilemma. *Journal of Business Ethics*, [e-journal] 105(2), pp. 187–196. <http://dx.doi.org/10.1007/s10551-011-0958-y>
- Nybom, T., 2003. The Humboldt legacy: reflections on the past, present, and future of the European university. *Higher Education Policy*, [e-journal] 16(2), pp. 141–159. <http://dx.doi.org/10.1057/palgrave.hep.8300013>
- Nwana, H., 1990. Intelligent tutoring systems: an overview. *Artificial Intelligence Review*, [e-journal] 4(4), pp. 251–277. <http://dx.doi.org/10.1007/BF00168958>
- Pessin, V. Z., Yamane, L. H. and Siman, R. R., 2022. Smart bibliometrics: an integrated method of science mapping and bibliometric analysis. *Scientometrics*, [e-journal] 127(6), pp. 3695–3718. <http://dx.doi.org/10.1007/s11192-022-04406-6>
- Ponsard, A., Escalona, F. and Munzner, T., 2016. PaperQuest: a visualization tool to support literature review. In: CHI'16 (Conference on Human Factors in Computing Systems), 2016 *CHI Conference Extended Abstracts on Human Factors in Computing Systems*. San Jose, California, USA, 7–12 May, 2016, pp. 2264–2271. New York, NY, USA: ACM.
- Postman, N., 1982. *The disappearance of childhood*. New York: Delacorte.
- Rudolph, R., Tan, S. and Tan, S., 2023. Bullshit spewer or the end of traditional assessments in higher education. *Journal of Applied Learning and Teaching*, [e-journal] 6(1), pp. 342–363. <http://dx.doi.org/10.37074/jalt.2023.6.1.9>
- Schön, E.-M., Neumann, M., Hofmann-Stölting, C., Baeza-Yates, R. and Rauschenberger, M., 2023. How are AI assistants changing higher education. *Frontiers in Computer Science*, [e-journal] 5, pp. 1–9. <http://dx.doi.org/10.3389/fcomp.2023.1208550>
- Sharma, R., Gulati, S., Kaur, A., Sinhababu, A. and Chakravarty, R., 2022. Research discovery and visualization using researchrabbit: a use case of AI in libraries. *COLLNET Journal of Scientometrics and Information Management*, [e-journal] 16(2), pp. 215–237. <http://dx.doi.org/10.1080/09737766.2022.2106167>
- Sharrock, G., 2010. Two hippocratic oaths for higher education. *Journal of Higher Education Policy and Management*, [e-journal] 32(4), pp. 365–377. <http://dx.doi.org/10.1080/1360080X.2010.491110>
- Stanford Encyclopedia of Philosophy, 2018. *Critical thinking*. [online] Available at <<https://plato.stanford.edu/entries/critical-thinking/>> [Accessed 16 August 2023].
- Stiegler, B., 2008. *Die logik der sorge*. Berlin: Suhrkamp.
- Stiegler, B., 2010. *Taking care of youth and the generations*. Standfort: University Press.
- Sullivan, M., Kelly, A. and McLaughlan, P., 2023. ChatGPT in higher education: considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, [e-journal] 6(1), pp. 1–10. <http://dx.doi.org/10.37074/jalt.2023.6.1.17>

- Sumarno, W. K., Kusumaningrum, W. R. and Nurhayati, E., 2021. The effects of knowledge and regulation of cognition on the students' writing skills in a metacognitive process-oriented writing instruction. *Journal of Language and Linguistic Studies*, [e-journal] 18(1), pp. 523–536. <http://dx.doi.org/10.1177/2158244019842681>
- Utterberg Modén, M., Tallvid, M., Lundin, J. and Lindström, B., 2021. Intelligent tutoring systems: why teachers abandoned a technology aimed at automating teaching processes. In: *54th Annual Hawaii International Conference on System Sciences*, Hawaii, University of Hawai'i at Manoa Hamilton Library, 4-8 January, 2021, pp. 1538–1547.
- van Damme, D. and Zahner, D., 2022. *Does higher education teach students to think critically*. Paris: OECD. <http://dx.doi.org/10.1787/cc9fa6aa-en>
- Vinson, K. E., 2012. Hovering too close: the ramifications of helicopter parenting in higher education. *Georgia State University Law Review*, [e-journal] 29, pp. 423–451. <http://dx.doi.org/10.2139/ssrn.1982763>
- von Garrel, J., Mayer, J. and Mühlfeld, M., 2023. *Künstliche Intelligenz im Studium: eine quantitative befragung von studierenden zur nutzung von chatGPT & co.* [online] http://dx.doi.org/10.48444/h_docs-pub-395
- Wang, H., Tlili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X. and Fei, C., 2023. Examining the applications of intelligent tutoring systems in real educational contexts: a systematic literature review from the social experiment perspective. *Education and Information Technologies*, [e-journal] 28(7), pp. 1–36. <http://dx.doi.org/10.1007/s10639-022-11555-x>
- Watanabe, A. and Schmohl, T., 2022. Die technologieverliebte hochschule: was folgt aus dem KI-gestützten lernen für den traditionellen bildungsauftrag. *Zeitschrift für Hochschulentwicklung*, [e-journal] 17(3), pp. 149–166. <http://dx.doi.org/10.3217/zfhe-17-03/09>
- Watanabe, A., 2023a. Exploring totalitarian elements of artificial intelligence in higher education with Hannah Arendt. *International Journal of Technoethics*, [e-journal] 14(1), pp. 1–15. <http://dx.doi.org/10.4018/IJT.329239>
- Watanabe, A., 2023b. Studierende im KI-diskurs: wie studierende in einem workshopformat über den KI-einsatz informiert und zum nachdenken über KI-gestütztes lehren und lernen angeregt werden. In: T. Schmohl, A. Watanabe and K. Schelling, eds. 2023. *Künstliche Intelligenz in der hochschulbildung: chancen und grenzen des KI-gestützten lernens und lehrens*. Bielefeld: Transcript. pp. 99–118.
- Wessels, I., Rueß, J., Gess, C., Deicke, W. and Ziegler, M., 2020. Is research-based learning effective: evidence from a pre-post analysis in the social sciences. *Studies in Higher Education*, [e-journal] 46(12), pp. 2595–2609. <http://dx.doi.org/10.1080/03075079.2020.1739014>
- Williams, J., 2016. A critical exploration of changing definitions of public good in relation to higher education. *Studies in Higher Education*, [e-journal] 41(4), pp. 619–630. <http://dx.doi.org/10.1080/03075079.2014.942270>
- Wissenschaftsrat, 2015. *Empfehlungen zum verhältnis von hochschulbildung und arbeitsmarkt - zweiter teil der empfehlungen zur qualifizierung von fachkräften vor dem hintergrund des demographischen wandels*. Drs. 4925-15. [online] Available at <<https://www.wissenschaftsrat.de/download/archiv/4925-15.html>> [Accessed 16 August 2023].
- Zawacki-Richter, O., Marín, V., Bond, M., and Gouverneur, F., 2019. Systematic review of research on artificial intelligence applications in higher education: where are the educators. *International Journal of Educational Technology in Higher Education*, [e-journal] (16), pp. 1–27. <http://dx.doi.org/10.1186/s41239-019-0171-0>

Educators' Academic Insights on Artificial Intelligence: Challenges and Opportunities

Jayaron Jose¹ and Blessy Jayaron Jose²

¹Preparatory Study Centre (FKA. English Language Centre), University of Technology and Applied Sciences - AlMussanah, Oman

²Information Technology Department, University of Technology and Applied Sciences - AlMussanah, Oman

jayaronjose@act.edu.om

jayaronjose@gmail.com

blessy.jose@act.edu.om

bjayaron@gmail.com

<https://doi.org/10.34190/ejel.21.5.3272>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: The study on "Educators' Academic Insights on Artificial Intelligence – Challenges and Opportunities" was conducted to gain a deeper understanding of the rapidly evolving phenomenon of AI in education. This research serves multiple objectives. Firstly, it aims to foster awareness regarding the integration of AI into teaching and learning practices by providing clear definitions of AI and explaining key AI-related terms. It also seeks to illustrate AI's diverse applications within a broader context, with a special focus on AI-supported research and learning platforms. Additionally, the study delves into the current discourse surrounding chatbots, contributing to address the central research question. Lastly, this initiative aims to provide valuable recommendations for effectively harnessing AI in education, enhancing the teaching and learning experience. The researchers conducted a review of literature concerning artificial intelligence. They adopted a qualitative method, using open-ended questions to collect feedback from educators globally, including those from the University of Technology and Applied Sciences, Al Musannah, and participants in the online discussion forum at Oxford English Learning Exchange.com. The qualitative data was analysed, leading to the identification of key themes and subthemes derived from the responses of research participants. The study's findings incorporated a wide range of concerns expressed by educators, comprising ten key subthemes. These concerns ranged from doubts about AI's ability to replace human educators and fears of its potential to hinder student development to worries about its hyped popularity and its perceived futuristic nature. Educators stressed the importance of effective AI training while emphasizing the need to prioritize human expertise over excessive reliance on AI. They were also acutely aware of both the advantages and disadvantages of AI, viewing it as both a potential boon and a looming threat. Furthermore, educators recognized the potential for enjoyable experiences with AI and acknowledged the pivotal role of users in determining the extent of AI adoption. Content analysis revealed additional apprehensions, such as concerns about job displacement, AI's impact on critical thinking, teacher frustration in assessing AI-assisted student writing, the use of AI-generated content for assessments, potential erosion of human services, stifling of user and learner creativity by AI, the risk of errors in AI-generated information, opportunities for cheating in exams, and concerns about the overreliance on and overrating of AI platforms. Positively, the findings included an array of opportunities that AI platforms offer. Study participants highlighted various aspects of these opportunities that surpassed their concerns and associated risks. The opportunities are categorized into twenty subthemes: enhancing learner motivation, facilitating template creation, utilizing AI as an educational aid, promoting proper training and fostering positive AI usage, harnessing AI for teaching challenging subjects, enabling personalized learning experiences, offering an interactive tutoring experience, supporting remote learning, facilitating self-study, providing comprehensive educational content overviews, giving instantaneous feedback and evaluation, functioning as search engines and chatbots, enabling content validation, efficiency in terms of cost and time, streamlining material preparation, facilitating skill and language enhancement, promoting familiarity with topics and vocabulary, enabling text-to-speech and speech-to-text conversions, editing multimedia elements, and facilitating content generation.

Keywords: AI, Artificial intelligence, Academic insights, Education, Educators, Challenges, Opportunities

1. Introduction

With the advance of technology in different forms (Jose, J. and Jayaron Jose, B., 2023), Artificial Intelligence (AI) is revolutionising the world. Artificial intelligence, or AI, which refers to the ability of machines or software to perform tasks that usually require human intelligence, such as tasks like recognizing images, making decisions, and translating languages, has had a profound influence on the global community. As a result, the emergence of Artificial Intelligence has brought forth a range of challenges and opportunities across various sectors, with a particular focus on education. With the advancement of computing technology, the earliest AI initiatives began in the middle of the 20th century, and they experienced steady growth until the introduction of more advanced

AI platforms like Chat GPT, Google AI Platform, Microsoft Azure, TensorFlow, and others. The field of artificial intelligence (AI) is evolving rapidly, and keeping up with the latest developments can be challenging. Just as you become accustomed to one chatbot, another emerges to take the spotlight (MacLean, 2023) as Phillips (2023) highlights the rapid advancement of AI across various fields, with progress measured in days and weeks, not years. According to Huang et al. (2023), AI is playing an increasingly significant role in education, particularly in the field of language education, and its impacts on different spheres of education are under-investigated.

In this paper, the researchers provide a comprehensive examination of different aspects of Artificial Intelligence. It commences with an introduction that offers an overview of AI. Following this, the research problem, objectives, and questions are presented. Building upon this groundwork, the researchers synthesize the main literary sources that served as the foundation for the study. They then provide a concise description of the methodology and research techniques used in their investigation. The findings of the study are subsequently scrutinized and discussed in the context of the research question. Lastly, the researchers conclude by summarising the findings, imparting insights, suggesting potential directions for future research, and offering recommendations based on the findings and the discussion.

2. Research Objectives

The specific aim of the study is to gain insights into educators' perceptions concerning the utilization of AI tools in educational settings, with a specific focus on identifying the challenges and opportunities that educators perceive in this regard. In general, the study's primary objective is multi-faceted. Firstly, it seeks to provide a comprehensive definition of AI, incorporating chatbots like Chat GPT, and aims to define key AI-related terminology. In doing so, the study intends to promote awareness of how AI can be effectively integrated into the realms of teaching and learning. Additionally, the study strives to exemplify a variety of AI applications, stretching across both general contexts and specialized research and learning tools.

3. Research Question

Based on the research objective, the literature reviewed and the qualitative nature of the study, the following research question was formulated.

What are educators' perceptions of the challenges and opportunities of using AI tools in learning and teaching?

4. Literature Review

This section of literature review attempts to achieve the research objectives of defining AI and AI platforms with examples, spanning general and educational contexts. The review highlights the importance and the implications of AI in education,

4.1 Definition of AI

AI can be defined as the capacity of machines or software to perform tasks that typically require human intelligence, including visual perception, decision-making, and language translation. AI systems utilize algorithms and machine learning methods to acquire knowledge from data and progressively enhance their performance (Akademisi, 2023). Artificial Intelligence refers to the concept and advancement of computer systems capable of executing tasks that typically demand human intelligence. AI encompasses three primary components: 1) machine learning and 2) deep learning. Machine learning enables computers to learn autonomously, without explicit programming. Deep learning, a subset of machine learning, involves the use of algorithms structured similarly to the logical processes of the human brain, known as artificial neural networks (Coursera, 2023; Figure 1).

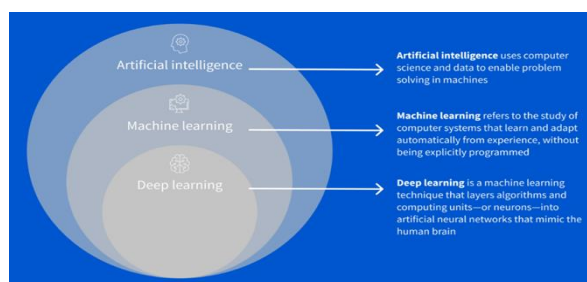


Figure 1: AI, ML and DL (Copied from Coursera, 2023)

It is equally essential to clarify what AI is not to dispel potential misconceptions. AI is NOT a substitute for human intelligence or expertise. It is not flawless; without meticulous design, testing, and supervision, AI can make mistakes or generate biased outcomes. While AI systems excel in performing specific tasks with great precision and efficiency, they lack the qualities of creativity, intuition, and empathy, which are fundamental for human learning and communication (Akademisi, 2023).

Figure 2 shows a few general examples of AI applications such as Siri, Alexa, TensorFlow, IBM Watson and Open AI, which are well known for their respective functions, assisting people with their everyday reminders, and answering simple to complex questions.



Figure 2: General examples of AI applications

4.2 AI as Perceived in the News

Recently, AI has caught attention of almost all news platforms across the globe. While many news headlines highlight the significance of AIs in a positive tone, several of them presents an alarming pessimistic view about AI (Figure 3). The topics of concerns range from enhanced crimes, job loss, and family breakdowns to cheating in exams. Governments, political personalities, or businessmen as Elon Musk are raising their concerns and expressing their recommendations. A few of them suggest shutting down AI while others recommend slowing down the fast process of AI development and application. For an instance, Kuenssberg (2023), the political editor and journalist at BBC doubts whether AI should we shut down or not. Meanwhile, Elong Musk calls for a pause on AI process due to numerous risks it poses to humanity (Metz and Schmidt, 2023). New York Post (Cost, 2023) reads that AI might cause depression, leading some of its users to commit suicide. Furthermore, AI is criminally used to facilitates crimes by spreading malware through social media platforms (HT Tech., 2023). Silverman (2023) wrote that AI might replace about 300 million employees internationally due to AI supported robots (Figure 3). Moreover, in education, Rosenblatt (2023) wrote that AI facilitated written exam could pass the final exam for Wharton school MBA program.

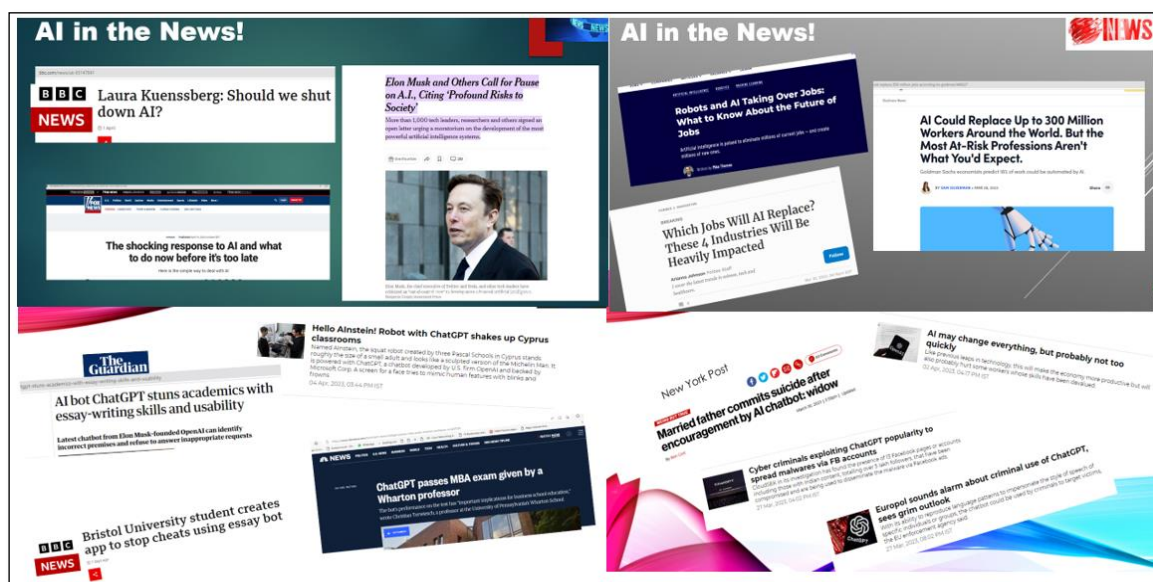


Figure 3: AI in the news

The following sections are aligned with the research aim and research question, moving from more general information to more specific literature regarding opportunities and challenges from educators' perspectives.

4.3 AI Applications for Learning and Research

While there are many AI applications for accomplishing different functions, many applications are specific to language teaching and learning. A few popular examples are EnglishCentral, Lingua.ly, Duolingo, EfluentU, Babbel, Reading Progress, Busuu and Grammarly (Figure 4).

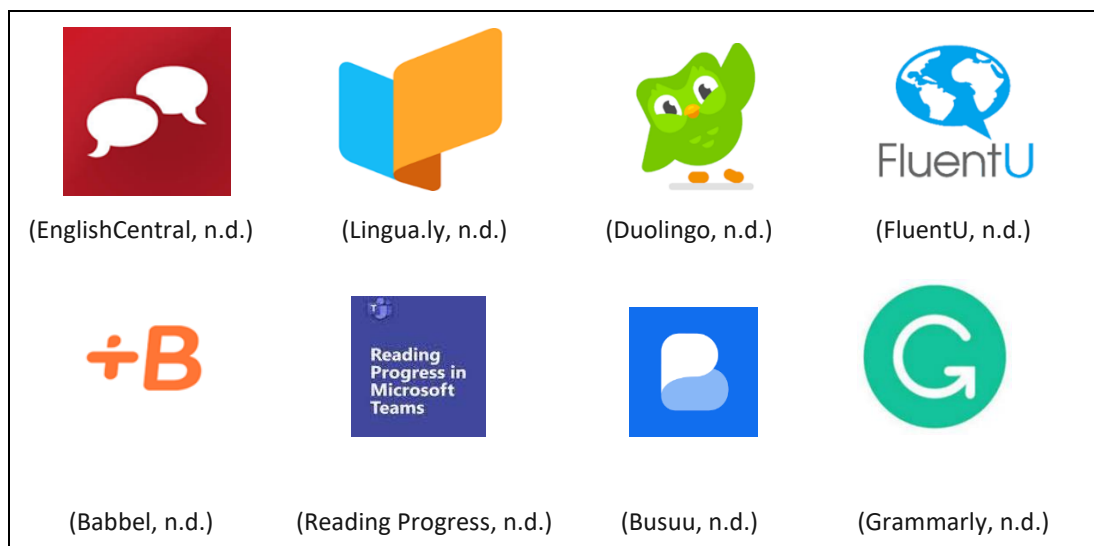


Figure 4: Chatbots for language learning and practice

In addition to language learning, Somasundharam (2023) mentions various AI tools that can be used in the advanced field of scientific research and publications. They are Scite Assistant, Consensus, Elicit, Chat GPT, Chat PDF, Research Rabbit, and SciSpace (Figure 5). The use of these tools can help researchers and academics to exploit their time and efforts more effectively and efficiently in fast progressing and publishing their work.



Figure 5: AI tools for research and publications

In the meantime, Mehmood (2023) listed top six ai support applications or platforms to facilitate teaching and learning of Mathematics. They are a) Wolfram Alpha, b) SymPy, c) Microsoft Maths Solver, d) GeoGebra, e) SymbMath and f) Desmos (Figure 6). AI, therefore, has simplified Mathematic calculations by providing pleasant learning experience and enhancing stakeholders' problem-solving skills. While AI has contributed to Mathematic teaching and learning, Mathematics has contributed to AI development and application. According to Odeyemi (2023), the vital contribution of Mathematics is related to linear algebra that facilitates the formation of a number of AI algorithms, representing and manipulating data. Moreover, linear algebra helps in natural language processing, image recognition, and recommendation systems.



Figure 6: AI tools for Mathematics

4.4 Academics' Perception of Chatbots

A chatbot is a computer program that employs artificial intelligence (AI) and natural language processing (NLP) to comprehend customer inquiries and automatically generate responses, mimicking human conversation. ("What is", n.d.). AI supported Chatbots have been implemented in different fields such as banking, health, education and so on. While the study by Nguyen et al. (2021) revealed customers' levels of satisfaction and their continued use of chatbots in the banking sector, Almalki and Azeez (2020) highlighted that the use of chatbots provides individuals with accessible and personalized support. Meanwhile, in education, chatbots facilitate students with different skills, and abilities to engage in learning, especially when conventional teaching fails to pay personal attention to learners (Srdanovic, 2023) because teachers have no extra time to spare for their learners due to their workload and tight teaching schedule. The study by Lin and Ye (2023) revealed that Chatbot technologies can support students in education "out of class anytime and anywhere and further obtain instant learning supports" (p.275), and it has a positive impact on learners.

Hariastiani (2019) classifies chatbots into three groups based on the structure, the purpose, and the audience. The structure consists of flow chatbot, hybrid and artificially intelligent as detailed in Table one. About the purpose or aim, functionality, and fun are the two sub themes; regarding the audience, generalist and specialist are the sub-groups.

Table 1: Classification of chatbots

Classification	Sub-classification	Operation
Framework	Flow chart	The flow chatbot or a tree-based chatbot has fixed responses programmed by the developer, and it only responds to the queries that are saved in the database.
	Artificially intelligent	Artificially intelligent chatbots utilizes intelligence to update its knowledge from the users' interactions or chats and experiences enabling free engagement.
	Hybrid	Hybrid chatbots integrate the features of flow chatbots and artificially intelligent chatbots presenting a personalised and highly engaging interaction with users, but within the developer's pre-determined framework.
Aim	Functionality	Different chatbots have differing functions such as reminder, online assistant, or learning.
	Fun	Some chatbots are for intended only for entertainment activities such as games and fun.
Audience	Generalist	Generalists enable searching for general knowledge about a wide variety of topics
	Specialist	Specialist chatbots focuses on matters in detail.

Note: Adapted from Hariastiani (2019).

It has been established that research concerning the integration of chatbots in education is still at a nascent stage, primarily due to a limited number of empirical studies exploring the application of efficient learning designs or learning strategies through chatbots. This indicates a substantial opportunity for further research to foster innovative teaching methods, ultimately enhancing the learning experience and educational outcomes (Hwang and Chang, 2021). Meanwhile, Kuhail, Alturki, Alramlawi, and Alhejori (2023) states that Chatbots have the potential to transform education through their ability to captivate learners, tailor learning experiences, assist educators, and gain profound understanding of learner behaviors. The study revealed that web-based chatbots were primarily created for teaching subjects like computer science, language, general education, and occasionally engineering and mathematics. Most of these chatbots served as instructors, while some acted as peers. The majority followed a preset conversation path, but a quarter offered personalized learning to cater to individual student requirements. Additionally, certain chatbots incorporated experiential and collaborative learning techniques. More than 30% of these chatbots underwent experimental assessments, showing positive results in terms of improved learning and user satisfaction (Kuhail, et.al., 2023).

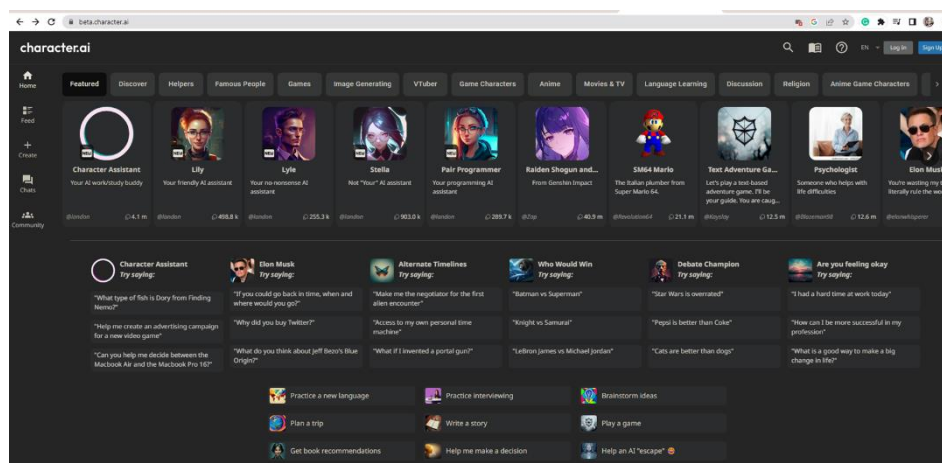


Figure 7: Character AI <https://beta.character.ai/>

One instance of a chatbot is Character Ai (Figure 7), which enables users to engage in English conversations through a character on the website. According to De Luna (2023), Character.AI is a website that utilizes a neural language model to analyse extensive text data and generate responses based on that knowledge. Users have the freedom to create characters on the site, which can be either fictional or based on real individuals, regardless of whether they are alive or deceased. Additionally, Users have the option to engage in a conversation with one character individually or set up a group chat involving multiple characters, allowing them to converse with each other and with you, all simultaneously, replicating the dynamic and social aspects of human interactions (McLean, 2023; Team, 2022).

4.4.1 Chat GPT – perceived advantages and disadvantages

Chat GPT is an intelligent computer program designed to comprehend human language and provide responses that resemble human conversation. You can engage in discussions with it on a wide range of topics, and it will strive to grasp your input and respond in a manner that sounds natural, akin to conversing with a real person (Moritz, et.al., 2023). GPT stands for Generative Pre-Trained Transformer, which generate text and tasks as prompted by its users. With the launching of Chat GPT as a Chatbot, the possibility of AI has attracted stakeholders from different backgrounds, especially from academic fields. The one-to-one human like interaction in responses to questions offered by Chat GPT on multiple topics instantly shook the academic and non-academic community. For instance (Figure 3), The Guardian published the news as “AI bot ChatGPT stuns academics with essay-writing skills and usability” (Hern, 2022). Another news headline reads, “Robot with ChatGPT shakes up Cyprus classrooms” (Kourtoglou, 2023) while BBC states “Bristol university student creates app to stop cheats using essay bot” (Heath, 2023). Furthermore, “Chat GPT passes MBA exam given by a Wharton professor” (NBC News, , 2023). These (Figure 3) groundbreaking AI advancements and their impacts on academic fields have prompted educators to investigate the subject carefully.

The study by Susnjak (2022) assessed the newly developed AI system ChatGPT which can perform complex tasks and create human-like text. The study found that ChatGPT can exhibit critical thinking and generate realistic text with minimal input, posing a threat to exam integrity. It raised concerns about ChatGPT's potential misuse in online exams, particularly in higher education where online exams are common. While solutions like returning

to in-person exams or using advanced proctoring and AI text detection tools may help, they are not foolproof. More research is needed to understand the implications of large language models like ChatGPT and develop anti-cheating strategies. Educators and institutions should remain vigilant in safeguarding exam fairness and validity for all students.

4.4.2 AI assisted text-to-speech and speech-to-text tools

Distinct from chatbots, Kumar, Koul, and Singh (2023) have noted that AI-assisted text-to-speech tools have piqued researchers' interest in exploring various facets of human-computer interaction. Text-to-speech AI platforms have, in parallel, facilitated substantial advancements in computer-mediated interactions. Meanwhile, these AI tools convert word documents, PDFs, and images into spoken content and vice versa, contributing to the simplification and creation of educational materials (Murf Resources, 2023). Natural Reader is an example of an AI tool that can be used to facilitate text to speech functions (Natural Reader, n.d.).

4.5 Perceived Educational Challenges and Opportunities of AI

Findings from the study (Ali, 2023) reveal that AI is a system designed to comprehend human speech naturally, capable of implementing a flipped learning approach for language education, enhancing learners' competence and productivity, and evaluating human speech. The application of AI in pedagogy demonstrates how it simplifies the process of teaching and learning languages (Ali, 2023). In a recent survey (Figure 8) conducted by Oxford English Exchange (2023), involving 389 educators, participants were asked to express their interest in learning about various topics, including well-being, administrative support, professional development, preparing learners for higher education, mediation, learning with concepts, artificial intelligence, assessment and exams, learner analytics and educational data, home/online learning, and everyday language learning. Notably, approximately 50% of the survey respondents expressed interest in acquiring knowledge about Artificial Intelligence (AI), as depicted in Figure 8. It's worth noting that due to the possibility of respondents selecting multiple items, there is response overlap. Consequently, the percentages provided in the survey (Figure 8) represent the relative popularity of each option among participants, and AI emerged as the most popular choice.

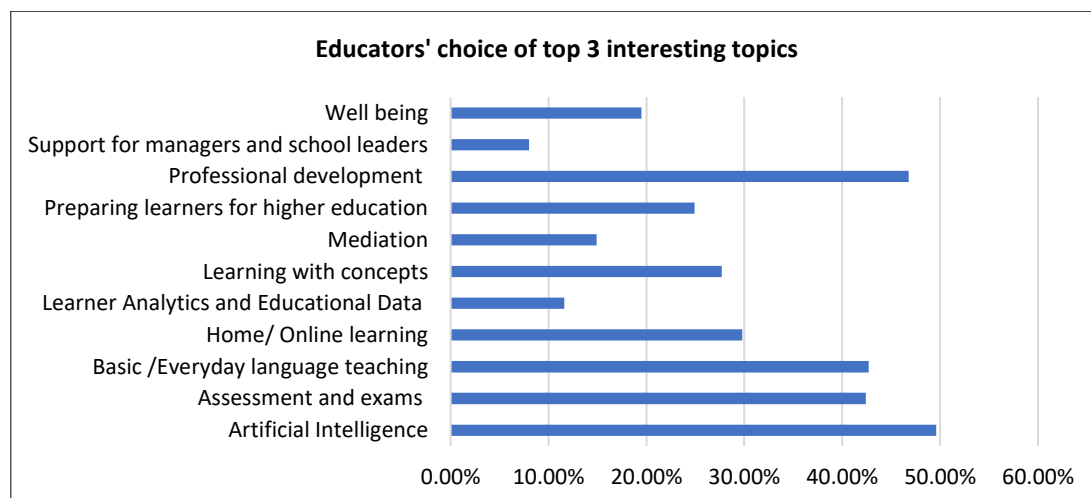


Figure 8: Adapted from <https://oxfordenglishlearningexchange.com/polls/55> 30.09.2023

Moreover, in the literature review, Gigster (2023) contends that AI does not diminish learners' critical and creative thinking abilities but instead enhances their capacity to modify and thoughtfully select AI-generated content. This, in turn, encourages more profound interaction and innovation, demanding a higher level of critical thinking. On the contrary, according to Carlie (2023), AI has drawbacks, including bias, cost, a deficiency of interpersonal skills, privacy concerns, and job displacement. Ansari (2023) suggests that AI can foster laziness by reducing the need for physical and mental effort in completing basic tasks. Meanwhile, Langreo (2023) emphasizes the need for training educators, as over 62% of surveyed educators expressed a desire to receive training on responsibly and effectively integrating AI into academic settings. In addition, Miller (n.d.) underscores the benefits of training in AI usage, enhancing users' skills, productivity, and competitiveness while inspiring innovative ideas. Despite concerns, Dhanrajani (2023) posits that AI is unlikely to replace human teachers, serving as a bridge between AI-supported information and students. Dhanrajani also highlights the challenges AI faces in fulfilling the teacher's role, such as instilling inspiration, motivation, fostering a nurturing classroom

environment, and developing critical thinking, creativity, and social skills. Likewise, Lakhani (2023) asserts that AI will not replace humans, but individuals with AI skills may replace those without.

As per Luckin (2023), AI might potentially disrupt meaningful learning, particularly when educators and learners are unaware of AI's limitations in empathy and creativity. Similarly, Thomas (2023) identifies biases and data manipulation as inherent risks in AI, leading to unreliability and errors. Yang and Zahn (2023) express educators' worries about students using AI tools for cheating, particularly in essay writing, which has led to the banning of AI platforms like Chat GPT on campuses. Furthermore, Cingillioglu (2023) discusses the growing problem of students using chatbots to write essays, which impacts academic integrity. On the other hand, Samper (2023) believes AI has a positive impact on education but suggests that its intervention may not significantly affect students' outcomes. Mollick and Mollick (2023) found that AI serves as an interactive and personalized tutor. Chen (2023) highlights AI technology's global potential for enhancing new teaching and learning modes, adapting to learners' needs. Additionally, Prothero (2023) points out how teachers can use AI to identify and prepare topic details for teaching their students effectively.

5. Methodology

The study employed the methodology of qualitative data collection and analysis to understand the educators' perceptions of the use of AI platforms, especially chatbots such as Chat GPT in teaching and learning. A survey questionnaire with open-ended question was circulated among the English Language teaching staff at the University of Technology and Applied Sciences Al Musannah to gather the educators' views on the subject. In addition, the researcher used data on educators' perception of AI tools such as ChatGPT from the online forum discussion at English Learning Exchange.com, Oxford University in which the researcher is a member and participant. Therefore, the study has used inputs collected from more than 35 educators (E) internationally.

The content analysis was used to analyse the verbatim responses, especially the responses obtained from the open-ended questions to find common themes and sub themes. This study, therefore, used qualitative content analysis, which involves systematic analysis of the content identifying and determining themes, words, or concepts in some texts (Hassan, 2024). Researchers use this analytical method to infer meaning and relationships of themes, concepts, or a particular set of words (Columbia University, n.d.). Based on the analysis, the investigators make inferences about the meanings of messages in the texts in their context. The content analysis attempts to analyse the intent or focus of certain group of individuals, the research participants' attitudinal and behavioural aspects, psychological or emotional condition of the group, and certain patterns of communication revealed in the text (Content Analysis, n.d.).

6. Results: Analysis and Findings

The researchers collected verbatim statements from the research participants, which were analysed for main themes and subthemes. There were 35 comments from the educators who participated in the data collection, and they are coded as (E1 – E35). Table 2 illustrates only a few examples of verbatim comments from the respondents, and the analysis mentions other relevant excerpts from the respondents' comments. These comments have helped researchers to answer the research question in line with the research objectives.

Table 2: Respondents' (educators') verbatim responses.

Educators (E)	Selected Verbatim comments
E1	<i>"This technology has the potential to transform education by providing personalized learning, immediate feedback, enhanced engagement, increased efficiency, accessible education, and Personalized learning, where AI adapts to students' needs and preferences."</i>
E2	<i>"The existence of such AI tools appears to create ethical problems. A person can get assistance, but s/he must create the manuscript on his own. The frame of "getting assistance" should be defined very cautiously, ..."</i>
E3	<i>"I fear that students will be so dependent on these technologies that they will not be able to think for themselves in the real world."</i>
E4	<i>"Although its good, I don't highly recommend this for my students. This will be an obstacle for their show their creative abilities. Students will depend on this, and they won't give a thought apart from it."</i>
E5	<i>"Technology is a powerful tool for teaching. It can facilitate student learning to its maximum. The key is to know how to use it effectively."</i>
E6	<i>"I fear that students will be so dependent on these technologies that they will not be able to think for themselves in the real world."</i>

Educators (E)	Selected Verbatim comments
E7	<i>"AI can provide access to education for students in remote or underprivileged areas or those with disabilities, thus ensuring accessible education. As technology improves, AI's potential to revolutionize the field of education will continue to expand."</i>
E8	<i>"It is a great technology if it will be used fairly. Yet, we can't rely on this technology wholly as all the information provided by chat GPT isn't reliable. Again, there are chances that students would find it easy to do their assignments using these technologies rather than studying themselves."</i>
E9	<i>"ChatGPT gave the answer and then gave some further instruction about that particular problem. So it's not just for getting the answers. It can also enhance a student's learning. Sort of like a virtual teaching assistant."</i>
E10	<i>"I tried using Chat GPT for framing reading comprehension questions for reading lessons. This technology would be an effective teaching tool for the teachers. However, the students might get fully rely on it for any writing task and this would sabotage their writing skills." "is a good technology, but it may hamper the creativity of the students and teachers both. Ready made things will make them think less. We should not let artificial intelligence take hold on us. _z"</i>
E11	<i>"I think it is a good platform and teachers can benefit from it. It gives a full-length overview of what you ask for."</i>
E12	<i>"There would always be disadvantages of it! As it would finish the human services facility and hardworking. Thus, the human would become more imaginative rather being realistic. Hope! We as a human would use it for the benefit of human being."</i>
E13	<i>"I do think it's important to teach students how to search their own brains without a telephone or other artificial aids and connect their own neurons to produce their own ideas and concepts."</i>
E14	<i>"it can be a two-edged sword of course. In the first place, I suppose it may make it easier for students to cheat in digitized environments, so security measures will probably have to be strengthened when implementing digital assessment for instance."</i>
E15	<i>"To me, it could be motivating for students at the beginning. They love technology. However, teachers should teach them the pros and cons about this."</i>

Based on the analysis of the respondents' comments (Table 2), three main themes were derived in connection with the use of AI (tools) in education. The main themes are: 1) the educators' concerns of using AI (Figure 9) the challenges/ risks posed by AI (Figure 10) and 3) AI facilitated opportunities (Figures 11 and 12).

6.1 Concerns

The concerns expressed by the participants include their emotional, and psychological responses to the introduction of AI in the field of education due to their worries or lack of awareness. Marriam Webster (n.d) defines a concern as "a matter that causes feelings of unease, uncertainty, or apprehension". Figure nine summarises the subthemes of educators' concerns as i) no human replacement, ii) scaring development, iii) too much popularity, iv) AI is futuristic, v) training to use, vi) training learners to use their brain, vii) awareness – prose & cons, viii) feel threatened, ix) fun with AI, x) users to decide, and xi) transform education (Figure 9).

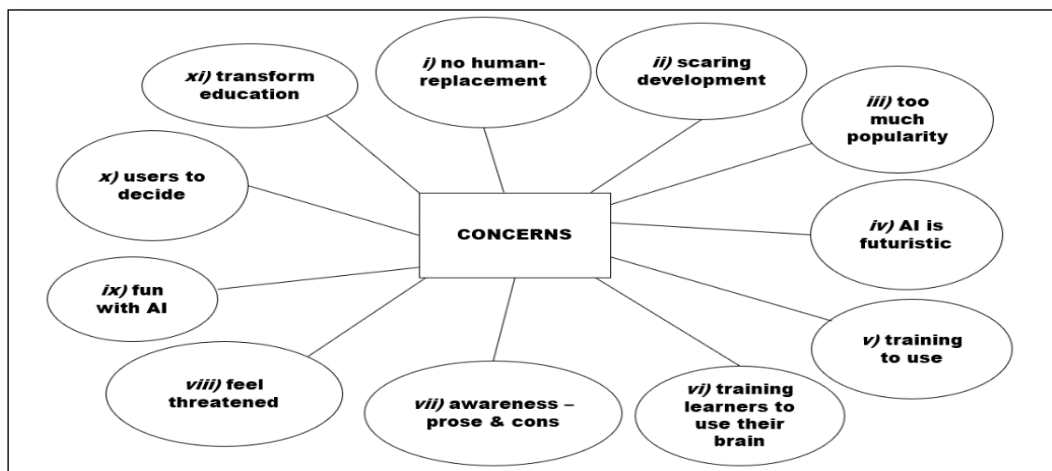


Figure 9: Main theme and sub-themes - concerns

With respect to the sub themes – AI as a scaring development and threatening (Figure 9.ii and viii), E4 states, “This will be an obstacle for them to show their creative abilities. Students will depend on this, and they won't give a thought apart from it”, and E6 comments, “I fear that students will be so dependent on these technologies that they will not be able to think for themselves in the real world” (Table 2). Meanwhile, E35 said, “That's a state-of-the-art thing, currently enjoying some popularity, too much popularity, much more than it deserves. ... It will fade with time.” (Figure 9.iii). Some respondents are concerned about AI pros and cons of AI, and lack of training to educators and learners (Figure 9. vi and vii). In this regard, “... teachers should teach them the pros and cons about this [AI].”, stated E15, and E26 said, “It would be useful for staff in the teaching profession to be given training [in AI]”

In the words of E24 “We should guide our students to use these new tools critically and creatively, and this should contribute to the users’ decision to use AI tools (Figure 9.x). Similarly, E25 states “When using it, we have to make use of our critical thinking and our creativity. It can be an aid when used with discrimination.”, and E23 thought, “... forbidding AI tools will not help - we educators need to know how to use AI tools and provide explicit instructions as to how a student's use of any AL tool is to be handled” (Table 2).

6.2 Risks or Challenges

Furthermore, the content analysis revealed the risks felt by educators (Figure 10) as a) elimination of jobs, b) discouraging thinking, c) teachers’ frustration related to giving feedback to learners’ AI assisted writing inputs, d) AI generated essays for assessment purposes, e) finishing off human services, f) AI hampering users or learners’ creativity g) causing errors in information, h) creating opportunities for cheating in exams, i) laziness, and overrating of AI platforms (Figure 10).

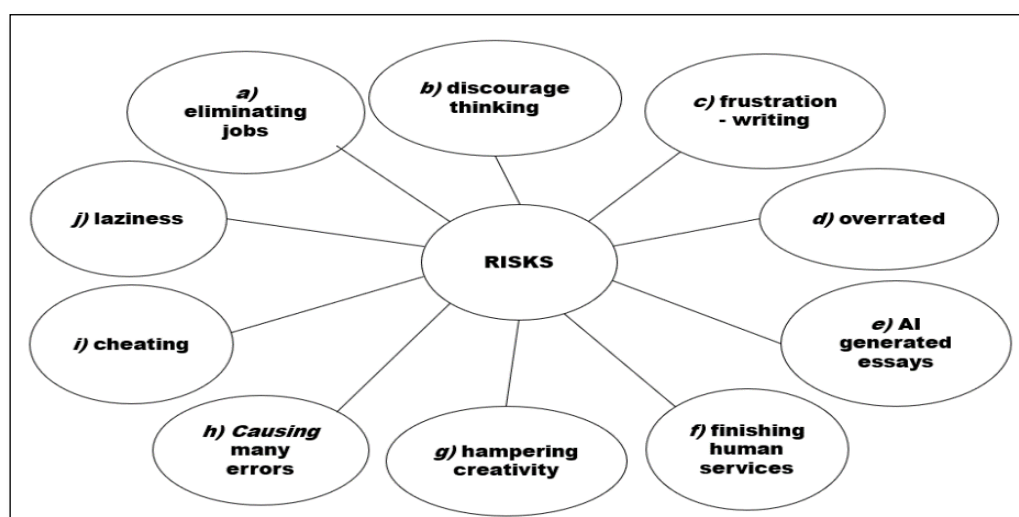


Figure 10: Main theme and subthemes – Risks/ challenges

Regarding hampering creativity, discouraging thinking, and promoting laziness (Figure 10.b, g and j), E3 said, “I fear that students will be so dependent on these technologies that they will not be able to think for themselves in the real world.” Meanwhile, E8 felt that AI platforms are not reliable, and they can cause errors, stating “It is a great technology if it will be used fairly. Yet, we can’t rely on this technology wholly as all the information provided by chat GPT [AI] isn’t reliable” (Figure 10.h).

Additionally, E2 and E4 remarked about ethical concerns in using AI (Figure 10.g), especially cheating in exams. They said, “The existence of such AI tools appears to create ethical problems.”, and “I suppose it may make it easier for students to cheat in digitized environments, so security measures will probably have to be strengthened when implementing digital assessment for instance”., respectively. E12 was concerned about the impacts of elimination of jobs and human services (Figure 10.a and f) because of AI, mentioning “As it would finish the human services facility and hardworking. Thus, the human would become more imaginative rather being realistic.” On the other hand, E35 felt that AI is “... Overrated! It will fade with time.” (Figure 10.d)

The inputs from E17 and E22 regarding writing outputs (Figure 10.c and e) demonstrated educators’ concerns and frustration. They said, “I have received so many AI-generated essays these past few weeks, and I am getting frustrated about this.”, and “Frustration, frustration, more frustration as correcting students’ writings is

completely useless as AI-generated writings have got more and more common lately.”, respectively. In addition, E6 commented, “I fear that students will be so dependent on these technologies”.

6.3 Opportunities

Figure ten and figure eleven provide information about the opportunities provided by AI platforms. In fact, different aspects of opportunities expressed by the study participants outweigh their concerns and the risks. Figure 11 describes opportunities in terms of 10 sub-themes as i) learner motivation, ii) template creation, iii) AI as educational aid, iv) proper training and positive usage, v) AI for teaching dangerous items, vi) personalized learning experience, vii) interactive tutor, viii) remote learner, ix) self-study and x) full length overview of items (Figure 11).

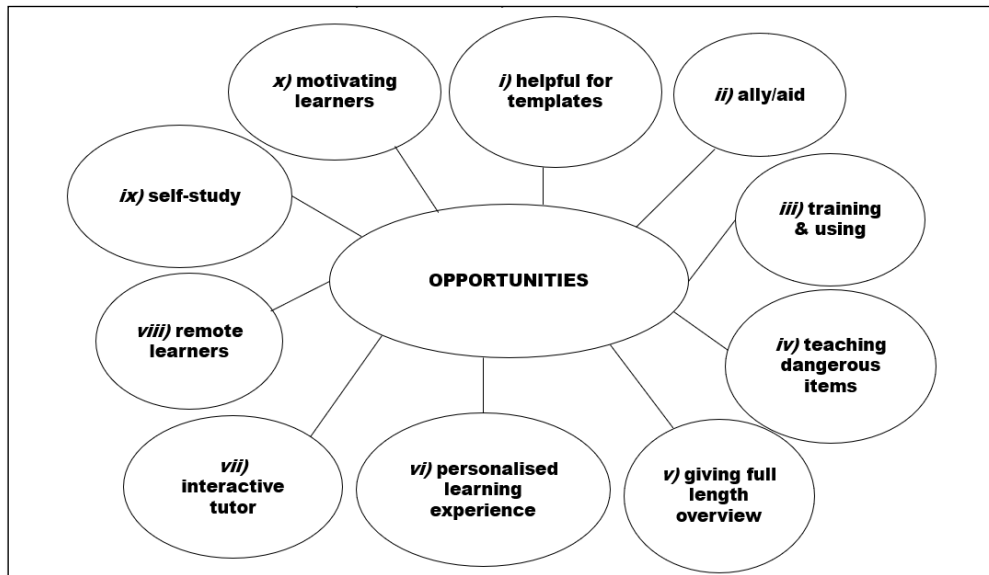


Figure 11: Main theme and sub-themes: opportunities (i - x)

To explain further, concerning AI tools for creating helpful templates for writing, E9 commented, “give students a template for something they might not know where to start writing - and then they need to edit and rewrite it, ...”, and E11 remarked, “It gives a full-length overview of what you ask for.” (Table 2, Figure 11.i, and v). In addition, AI as an ally or aid for learning and teaching, E9 and E5 said, “It can also enhance a student's learning. Sort of like a virtual teaching assistant.”, and “... is a powerful tool for teaching. It can facilitate student learning to its maximum. ...” (Figure 11.ii). At the same time, about AI's role of teaching dangerous items, E31 points out, “AI technologies can provide opportunities for students to engage in simulated experiences that might be too difficult, too dangerous, or too expensive to carry out in real life. ...” (Figure 11.iv). While E7 (Table 2) states, “AI can provide access to education for students in remote or underprivileged areas or those with disabilities, thus ensuring accessible education.”, E32 said, “can be used as an interactive tutor, helping students to review key concepts, answer questions, and provide feedback on assignments” (Figure 11.vii and viii). Regarding training and using AI effectively, being an interactive tutors and motivating learners (Figure 11.iii, vii and x), E26 said, “. It would be useful for staff in the teaching profession to be given training etc.”, and E32 and E15 commented that AI tools “can be used as an interactive tutor, helping students to review key concepts, answer questions, and provide feedback” and “it could be motivating for students at the beginning. They love technology.”. This in turn provides personalized learning experiences (Figure 11.vi).

In addition to the details in Figure 11, Figure 12 summarize the remaining advantages or opportunities put forth by AI supported tools in education as unravelled from the respondents' comments. The ten sub-themes are xi) real-time feedback and assessment, xii) search engines and chatbots, xiii) content checking, xiv) cost and time efficient, xv) material preparation, xvi) aiding skill and language development, xvii) topic and vocabulary familiarization, xviii) text to speech and speech and speech to text conversions, xix) multimedia editing, and xx) content creation (Figure 12).

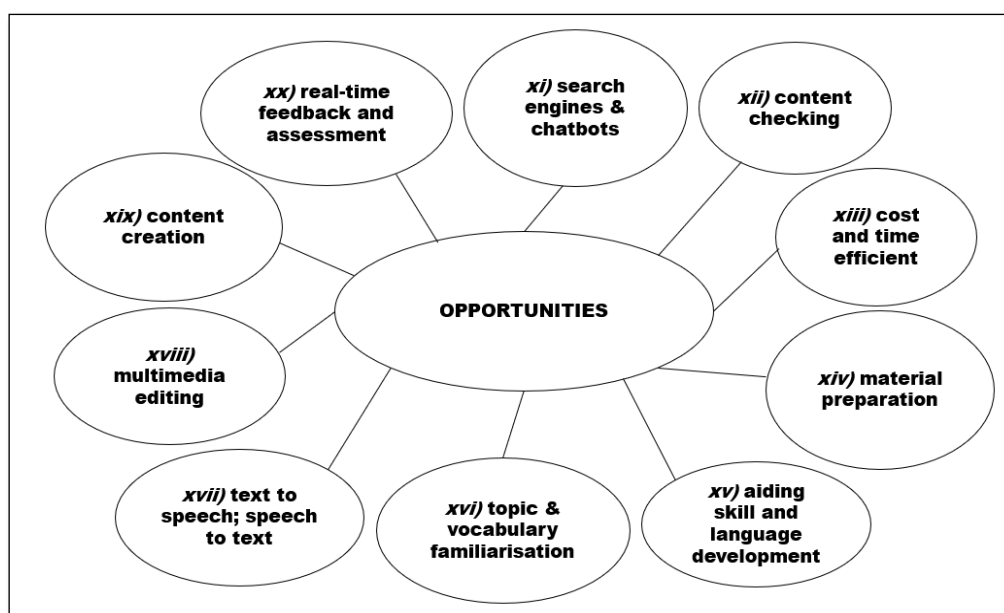


Figure 12: Main theme and sub-themes: opportunities (xi - xx)

To elucidate a few sub-themes, E11 remarked, “It gives a full-length overview of what you ask for.”, while E32 stated, “[AI as] an interactive tutor, helping students to review key concepts, answer questions, and provide [real-time] feedback on assignments. ... Sort of like a virtual teaching assistant” (Figure 12.xx). E34 claimed, “AI can provide access to education for students in remote or underprivileged areas or those with disabilities, thus ensuring accessible education”, and this explains how AI is cost and time efficient (Figure 12.xiii). Furthermore, E32 and E15 opined, “[AI is a] a sort of like a virtual teaching assistant” and “it could be motivating for students at the beginning. They love technology” respectively (Figure 12.xix, and xviii; Table 2).

About AI tools as a search engine for content checking and content creation (Figure 12.xi, xii and xix), E21 said, “...prove to be an ally in the classroom, just like any search engine is”, and E10, possibly a language educator, argued, “I tried using Chat GPT for framing reading comprehension questions for reading lessons.” (Figure 12.xiv; Table 2). Meanwhile, regarding aiding language skills and language development in education (Figure 12.xv), E19 added, “the AI powered influx of computer programs will transform the current practice of language education globally. ...” (Figure 12.xi and xii; Table 2). E10 stated that s/he used AI for “framing reading comprehension questions for reading lessons. This technology would be an effective teaching tool for the teachers for finding topics and appropriate words according to topics”, and this explains importance of AI in topic selection and vocabulary familiarisation (Figure 12.xvi) for different learning tasks.

7. Discussion

The study has provided the researchers with essential information from secondary and primary sources to achieve the research objectives. While the literature review facilitated insights into certain theoretical aspects of AI, the feedback from research participants enabled the investigators to understand educators’ perceived concerns, challenges, and opportunities in applying AI in the field of education. The educators’ perceptions in the study matched with certain pieces of information gathered from secondary sources. The following sections discuss each main themes that were found because of qualitative data analysis (Figures 9,10, 11 and 12)

7.1 Of Concerns

Regarding educators’ concerns (Figure 9) some educators feel scared and threatened of the introduction of artificial intelligence in the academic field (Table 2; Figure 9.ii, viii). The feeling of scare and threat expressed by educators resembled the opinions raised by Kuenssberg (2023) and Elon Musk (Metz and Schmidt, 2023). One of the reasons educators feel scared is that AI might take away students’ creativity and thinking process (Ansari, 2023). Nevertheless, in contrast, Gigster (2023) argues that AI does not deprive people of their critical and creative thinking capabilities, but it might make them better at modifying and carefully selecting AI-generated content, encouraging more profound interaction and innovation, which requires a lot of critical thinking. Moreover, respondents felt that AI platforms, such as Chat GPT, are too popular and overrated lacking proven

effectiveness (Figure 9, iii). Nevertheless, this may not be true, as AI is rapidly advancing across various fields, measured in terms of days and weeks, rather than years (Phillips, 2023).

Another concern is that the unawareness of the pros and cons (Figure 9.vii) of AI may affect learners negatively. Carlie (2023) states that AI has demerits, such as bias, cost, lack of interpersonal skills, lack of privacy and job placement. In addition, it has advantages, such as personalised studying, feedback, facilitating lesson preparation and creating virtual fact simulations. This explains that not only learners, but educators also should be aware of the pros and cons of AI in education. Regarding training in using AI (Figure 9.v and vi), educators reported that learners should be guided well in using it (E26, Table 2). This resembles what Langreo (2023) wrote about the need for giving training to educators that more than half of educators (62%) who were surveyed wanted to get trained in how to teach learners to use AI responsibly and effectively in academic situations. Moreover, the manner of guidance or training in AI is described as critical and creative (Gigster, 2023). Similar responses from educators informs researchers that proper training, considering technical, critical, and moral aspects of the use of AI, is paramount in implementing it. In this regard, Miller (n.d.) states that training or guidance in the use of AI can benefit users in number of ways, making users of AI skilful, productive, and competitive in their field. Training enables educators to know the technology and its potential, encouraging them to come up with new ideas to implement it.

Some educators expressed that AI cannot replace humans (Figure 9.i, Table 2, E27). This view very much agrees with Dhanrajani's (2023) views that AI is unlikely to replace human teachers in the classroom because educators are the intervention between AI supported information and students. Moreover, Dhanrajani (2023) points out that the idea of AI fulfilling the role of a teacher by instilling inspiration and motivation in students, fostering a secure and nurturing classroom atmosphere, and assisting in the cultivation of critical thinking abilities, creativity, and social skills appears quite distant and challenging to achieve. Additionally, the respondents perceive that AI is futuristic (Figure 9.iv; Table 2). This reflects on what Phillips (2023) remarked about the role of AI in education that AI systems are poised to bring far-reaching consequences to the operations of educational institutions, the roles of educators, and the personal and professional growth of students as they prepare for the future workforce; furthermore, the concept of a 'digital native' is quickly becoming outdated, especially as Generation Z and subsequent generations effortlessly acquire expertise in emerging AI technologies at a faster pace than many adults. This provides tremendous opportunities for the application of cutting-edge technologies as generative AI. Finally, some educators feel that AI can be used as a funny means of learning (Figure 9.ix), and users to use it discerningly (Figure 9.x); , and this fun element adds to the concern that AI may disrupt meaningful learning (Luckin, 2023), especially when educators and learners are unaware of their lack of empathy, and creativity as humans.

7.2 Of Risks or Challenges

About risks or challenges perceived by educators (Figure 10), most of them can be understood to have originated from educators' concerns about AI. They viewed discouraging thinking and hampering creativity (Figure 10. b and g) are not mere concerns, but real challenges. While AI proponents (Gigster, 2023) argue against it, Miller (2023) mentions that those who heavily depend on AI may decline their critical thinking. Furthermore, errors or unreliability (Figure 10.h, Table 2, E8) of AI platforms has been a risk perceived by the educators (Table 2, E8). This agrees with Thomas (2023) who states explicitly that biases, and manipulation of data are two of many risks of AI, resulting in unreliability and errors.

Additionally, educators consider cheating (Figure 10.i; Table 2 E2 and E14) in exams and assessments is a risk posed by AI. Yang and Zahn (2023) remarked that educators worry about students' cheating due to AI tools, especially in writing essays (Rosenblatt, 2023). This resulted in banning AI platforms as Chat GPT in campuses. Further, some educators felt AI might eliminate human jobs and services (Figure 10. a and f) in many fields, possibly in the field of education, too. This may not be true because AI lacks human critical and creative thinking; nevertheless, Lakhani (2023) stated that AI would not replace humans, but humans with skills in artificial intelligence might replace the ones without it. Meanwhile, many foresee a possibility of AI take over in many areas of services and jobs (Millar, 2023; Ansari, 2023; Silverman, 2023; Dhanrajani, 2023). However, some respondents felt that AI impact is overrated (Figure 10.d; Table 2 E.35). This perception may be a challenge to educators that they may not pay much attention to the fast-paced impact of AI in education (Phillips, 2023).

Another important and specific risk felt by educators is students' writing that they generate essays with the help of AI chatbots (Figure 10. c and e; Table 2, E17 and E22), such as Chat GPT. Cingillioglu (2023) states that students' use of chatbots for writing essays is a growing problem faced by educators, and it affects academic integrity. Likewise, laziness (Figure 10.j; Table 2-E6) is regarded as a challenge presented because of AI

interruption. This means AI can foster laziness, avoiding physical or mental efforts in accomplishing basic tasks (Ansari, 2023), in addition to decreased level of critical and creative thinking as seen previously (Miller, 2023).

7.3 Of Opportunities

The opportunities (Figures 11 and 12) outweighed the concerns (Figure 9) and challenges (risks) (Figure 10) perceived by educators who participated in the study. One of the advantages or opportunities presented by AI for learners is that it is helpful for template, especially in writing or other technical lessons (Figure 11.i; Table 2-E9). According to The Writing Centre (n.d.), AI platforms facilitates brainstorming ideas, generating outlines, and familiarising models or genres of writing in writing lessons. Nevertheless, as mentioned earlier, this has the potential to affect academic integrity (Cingillioglu, 2023) negatively. Further, some participants responded that AI could be ally or aid in education and can be used as a tool with proper training (Figure 11.ii and iii; Table 2 E5 and E9). This supports Samper's (2023) view that AI contributes positively to schooling though he feels that its intervention does not affect students' results significantly.

Another important opportunity offered by AI in education is that it helps educators to teach dangerous items to students (Figure 11.iv; Table 2.E31). This resembles one among many AI dependent future promises envisaged by Teach for America (2023) in education that AI technology can be used to facilitate training skills that are risky and hard otherwise. In addition, AI as a learning aid provides full length overview of the learning task, personalised learning experience and service as an interactive tutor (Figure 11. v, vi and vii; Table 2.E11 and E12). The study by Mollick and Mollick (2023) found that AI functioned as an interactive and personal tutor if used in manner being aware of the risks (Yang and Zahn, 2023; Thomas, 2023).

Moreover, AI facilitates remote learners, self-study, and learner motivation (Figure 11. viii, ix and x; Table 2 E34), and this explains how AI is cost and time efficient (Figure 12.xiii; Table 2 E15 and E32). Learners' love for technology motivates their learning by means of it (Jayaron and Blessy, 2022; Jose, 2022, Jose 2021). Furthermore, Chen (2023) wrote that AI technology provides worldwide opportunity to enhance new modes of teaching and learning, adapting to learners' needs.

Another opportunity educators felt is that AI platforms such as chatbots function as search engines (Figure 12.xi), which can help them prepare teaching materials (Figure 12.xiv) and check and create content (Figure 12.xii and xix; Table 2 E10 and E21). Meanwhile, regarding aiding language skills and language development in education (Figure 12.xv; Table 2 E19), respondents felt that AI powered computer program would transform the current practice of language learning and education worldwide. In this regard, the number of AI facilitated teaching tools, specific to different subjects (Figures 5, 6 and 7; Mehmood, 2023; Somasundharam, 2023) demonstrate that these opportunities perceived by educators as realistic.

Apart from the above derived sub-themes of opportunities, researchers' reflection on respondents' feedback and the literature review found that AI is advantageous for identifying learning topics and familiarising topic specific vocabulary for teaching and learning; manipulation of text to speech and speech to text AI platforms for teaching skills and language, especially listening, speaking, and reading; and editing multimedia contents (Figure 12.xvi, xvii and xviii). King and Caplan (2023) explain editing multimedia content with the help of AI tools in education, such as creating an image from text prompts, diagrams and editing videos; they discuss chat, search, document-based AI tools in addition to specialty and multimedia tools. Moreover, teachers can identify topics and topic details in preparation for teaching their learners (Prothero, 2023). Meanwhile, the text to speech and speech text AI software converts word, pdf and image documents to speech and vice versa fostering learning through simplification and creation of educational contents (Murf Resources, 2023).

8. Limitations of the Study

The scope of the study was limited to the respondents' verbatim responses and the literature reviewed. While the majority of sub-themes in the findings were identified through the analysis of the responses, a few similar sub-themes were derived from the literature review. For example, sub-themes such as multimedia editing, text-to-speech, and speech-to-text (Figure 11, xvii and xviii) are derived from the literature review. They are included in the analysis due to their similarity to the main theme. Moreover, the study does not include the age and gender of the respondents.

Furthermore, the findings are based on the analysis of educators' responses, which might be context-dependent and may not be applicable to different audiences in different contexts. Furthermore, Table 2 and the analysis included only a sample verbatim comments for readability. Finally, this study was conducted over a short-term

period during the introduction of AI in various fields. However, a long-term study focusing on the rapidly advancing AI platforms may yield different findings and provide deeper insights.

9. Conclusion and Recommendations

To conclude, the study has reasonably answered the research question investigating educators' perception of the challenges and opportunities of using AI tools in learning and teaching. The main findings include three main themes and their subthemes, which are summarized as – concerns, risks, and opportunities.

Certain practical implications of the study are summarized with reference to the findings. In fact, the concerns, and risks (challenges) revolve around various aspects of AI technology. Moreover, they address the notion that AI will not fully replace humans and the concerns arising from its rapid development and widespread popularity. However, AI is often seen as futuristic and requires training for effective use, although some advocate for training in enhancing human cognitive skills instead of relying solely on AI. Furthermore, there is an emphasis on awareness of both the advantages and disadvantages of AI, as it can be both promising and threatening. While AI can provide enjoyment and entertainment, its adoption ultimately depends on individual user decisions. Nevertheless, regarding risks, the educators worry about job displacement, AI's potential to stifle critical thinking, and the challenges teachers face when providing feedback on AI-assisted student writing. There are also discussions about the use of AI-generated essays for assessment and concerns about its impact on human services. In addition, AI is seen as a potential hindrance to users' creativity, a source of errors in information, and a platform that could facilitate cheating in exams. Additionally, some caution against the development of laziness and the overestimation of AI platforms is required in the application of AI tools in education.

On the other hand, the opportunities of AI perceived by educators exceed their felt concerns and risks. They include a range of aspects, such as learner motivation, the creation of templates, AI's role as an educational aid, the importance of proper training and responsible utilization. Additionally, they touch on more specific areas like using AI for teaching potentially dangerous subjects, delivering personalized learning experiences, providing interactive tutoring, catering to remote learners, promoting self-study, offering comprehensive item overviews, enabling real-time feedback and assessment, employing search engines and chatbots, facilitating content checks, maintaining cost and time efficiency, aiding in material preparation, enhancing skill and language development, familiarizing students with topics and vocabulary, enabling text-to-speech and speech-to-text conversions, supporting multimedia editing, and assisting in content creation. These aspects collectively illustrate the diverse and multifaceted role of AI in the field of education in terms of its practical implications and implementation.

In line with the research objective, literature review, analysis, findings, discussion of the findings, and the conclusion, the following recommendations are made.

- Educators and learners should be made aware of the pros and cons of AI and Chat bots.
- Specific instructions regarding the use of AI assisted learning tasks should be set in place before integrating such technology in teaching.
- Proper digital monitoring system should be in place to avoid cheating and plagiarism through AI platforms.
- Learners should be made aware that it is their own intelligence that counts in exams despite every form of AI assistance.
- Stakeholders should be made aware that AI is not a replacement for human intelligence and expertise.
- AI should be used as an aid, but not as an alternative to humans because it lacks empathy, creativity and intuition which are essential for human learning and communication.
- AI is fallible, and stakeholders should be made aware that AI can make mistakes or produce biased results if it is not properly designed, tested, and monitored.
- Proper training should be given to stakeholders (teachers, learners, administrators, etc.) in applying AI platform in education.

Regarding future research, since AI is fast evolving, constant studies may be carried out to measure its implications on social, cognitive, behavioural, ethical, intellectual, and emotional aspects of learning and teaching. In addition, further investigations may be carried out to understand subject/ specialization-wise impacts of AI on education, and different AI aided techniques to be incorporated to facilitate learning.

References

- Ali, Z. 2020, February. Artificial intelligence (AI): A review of its uses in language teaching and learning. In *IOP Conference Series: Materials Science and Engineering* (Vol. 769, No. 1, p. 012043). IOP Publishing.
- Almalki, M. and Azeez, F. K. 2020. Health chatbots for fighting covid-19: a scoping review. *Acta Informatica Medica*, 28(4), 241. <https://doi.org/10.5455/aim.2020.28.241-247>
- Akademisi, L. 2023, March 1). *Innovative approaches in ELT: using artificial intelligence in ELT* [Video File]. YouTube. Available at <<https://www.youtube.com/watch?v=z3pVu561RC4>> [Accessed 12 January 2024].
- Babbel. n.d. *Which language do you want to speak?* Available at <<https://www.babbel.com/>> [Accessed 17 September 2023].
- English Central. n.d. *Your Personal English Teacher*. <<https://www.englishcentral.com/browse/videos>> [Accessed 17 September 2023].
- Ansari, M.M. 2023. *The laziness effect of artificial intelligence*. Available at <<https://www.linkedin.com/pulse/laziness-effect-artificial-intelligence-mohammad-muzammil-ansari/>> [Accessed 17 August 2023].
- Busuu. n.d. *The language e-learning solution for institutions*. Available at <<https://business.busuu.com/education>> [Accessed 15 August 2023].
- Carlie. 2023. *AI in education: the advantages and disadvantages*. Available at <<https://www.instrucko.com/blogs/the-advantages-and-disadvantages-of-ai-in-education>> [Accessed 12 September 2023].
- Chat GPT. n.d. *Chat GPT*. Available at <<https://chat.openai.com/>> [Accessed 12 August 2023].
- ChatPDF. *ChatPDF*. n.d. Available at <https://www.chatpdf.com/?via=tran-van-thang&gclid=CjwKCAjwpJWoBhA8EiwAHZFzfsP7XjADtVZ1Pfal5Is06D-HC85OgvmGZY2_6DMRtv9wVNahYQ8LRoC6-sQAvD_BwE> [Accessed 20 September 2023].
- Chen, C. 2023. *AI will transform teaching and learning. Let's Get it Right*. Available at <<https://hai.stanford.edu/news/ai-will-transform-teaching-and-learning-lets-get-it-right>> [Accessed 21 September 2023].
- Cingillioglu, I. 2023. "Detecting AI-generated essays: the ChatGPT challenge", *International Journal of Information and Learning Technology*, Vol. 40 No. 3, pp. 259-268. <https://doi.org/10.1108/IJILT-03-2023-0043>
- Columbia University. n.d. *Content Analysis*. Available at <https://www.publichealth.columbia.edu/research/population-health-methods/content-analysis#:~:text=Courses-,Overview,words%2C%20themes%2C%20or%20concepts>. [Accessed 15 January 2024].
- Consensus. n.d. *Consensus*. Available at <<https://consensus.app/>> [Accessed 28 September 2023].
- Content Analysis. n.d. *Content Analysis*. Available at <<https://www.itma.vt.edu/courses/progprod2/resources/ch20.pdf>> [Accessed 10 January 2024].
- Cost, B. 2023. Married father commits suicide after encouragement by AI chatbot: widow. *New York Post*, [online]. Available at <https://nypost.com/2023/03/30/married-father-commits-suicide-after-encouragement-by-ai-chatbot-widow/> [Accessed 28 September 2023].
- Coursera. 2023. *Deep learning vs. machine learning: beginner's guide*. Available at <<https://www.coursera.org/articles/ai-vs-deep-learning-vs-machine-learning-beginners-guide>> [Accessed 15 September 2023].
- De Luna, E. 2023. *Character.AI: What it is and how to use it*. Available at <<https://me.mashable.com/tech/28412/characterai-what-it-is-and-how-to-use-it>> [Accessed 10 January 2024].
- Dhanrajani, R.M. 2023. Can AI replace teachers? Here's how the education sector is changing. *Business Today*, [online]. Available at <<https://www.businesstoday.in/magazine/deep-dive/story/can-ai-replace-teachers-heres-how-the-education-sector-is-changing-381036-2023-05-12>> [Accessed 10 January 2024].
- Duolingo. n.d. *How to learn English on the Duolingo app?* Available at <<https://englishproficiency.com/blog/how-learn-english-duolingo-app/>> [Accessed 10 November 2023].
- Elicit. n.d. *Analyze research papers at superhuman speed*. Available at <<https://elicit.com/>> [Accessed 15 September 2023].
- Fluentu. 2023. *Teach on your terms: how to successfully teach English from home*. Available at <<https://www.fluentu.com/blog/educator-english/teaching-english-from-home/>> [Accessed 15 September 2023].
- GeoGebra. n. d. *GeoGebra for teaching and learning Math*. Available at <<https://www.geogebra.org/?lang=en>> [Accessed 15 September 2023].
- Gigster n.d. *How AI Shapes the future of Critical Thinking*. Available at <<https://gigster.com/blog/ai-criticalthinking/#:~:text=While%20technology%20has%20altered%20cognitive,data%2C%20and%20address%20ethical%20concerns>> [Accessed 12 January 2024].
- Grammarly. n.d. *My Grammarly*. Available at <<https://app.grammarly.com/>> [Accessed 15 September 2023].
- Haristiani, N. 2019, November. Artificial Intelligence (AI) chatbot as language learning medium: An inquiry. In *Journal of Physics: Conference Series* (Vol. 1387, No. 1, p. 012020). IOP Publishing.
- Hassan, M. 2024. *Content Analysis – Methods, Types and Examples*. Available at <<https://researchmethod.net/content-analysis/>> [Accessed 10 January 2024].
- Heath, N. 2023, April 6. Bristol university student creates app to stop cheats using essay bot. *BBC News*, [online]. Available at <<https://www.bbc.com/news/uk-england-bristol-65200549>> [Accessed 14 January 2024].
- Hern, A. 2022. AI bot ChatGPT stuns academics with essay-writing skills and usability. *The Guardian* [online]. Available at <<https://www.theguardian.com/technology/2022/dec/04/ai-bot-chatgpt-stuns-academics-with-essay-writing-skills-and-usability>> [Accessed 13 January 2024].

- HT Tech. 2023. Cyber criminals exploiting Chat GPT popularity to spread malwares via FB accounts. *Hindustan Times*, [online]. Available at <<https://tech.hindustantimes.com/tech/news/cyber-criminals-exploiting-chatgpt-popularity-to-spread-malwares-via-fb-accounts-71679935669266.html>> [Accessed 12 September 2023].
- Huang, X., Zou, D., Cheng, G., Chen, X., and Xie, H. 2023. Trends, research issues and applications of artificial intelligence in language education. *Educational Technology & Society*, 26(1), 112–131. <https://www.jstor.org/stable/48707971>
- Hwang, G. J. and Chang, C.Y. 2021. A review of opportunities and challenges of chatbots in education, *Interactive Learning Environments*, DOI: 10.1080/10494820.2021.1952615
- IBM Watson. n.d. *IBM Watson is AI for business*. Available at <<https://www.ibm.com/watson>> [Accessed 10 September 2023].
- Jose, B.J. and Jose, J., 2022. Impact of e-learning on learning outcomes among higher education students – an integrated review. *Scholars Journal of Engineering and Technology*, 10(7): pp. 141-145. 10.36347/sjet.2022.v10i07.003.
- Jayaron, J and Blessy, J.J. 2021. Learners’ perception of using Microsoft teams predicted by technology acceptance model at university of technology and applied sciences, Oman. *Webology* (ISSN: 1735-188X) Volume 18. Available at <<https://www.webology.org/abstract.php?id=1457>> [Accessed 7 September 2023]
- Jose, J. and Jayaron Jose, B. 2023. A cross-sectional study on student experience and contentment of online learning during covid 19 lockdown. *Electronic Journal of e-Learning*, 21(3), 222-233.
- Jose, J., 2022. Omani EFL learners’ adaptability and acceptance of MS-teams for online learning amidst covid 19 pandemic. *Conference Proceedings: Oman 21st International ELT Virtual Conference* At: Sultan Qaboos University, Muscat, Oman
- King, T. and Caplan, J. 2023. *Exploring the possibilities: creating, enhancing, and editing with artificial intelligence*. Available at <https://s39613.pcdn.co/wp-content/uploads/2023/09/Exploring-the-Possibilities-Creating-Enhancing-and-Editing-with-Artificial-Intelligence_otter_ai.pdf> [Accessed 18 August 2023].
- Kourtoglou, Y. 2023. Hello Alnstein! Robot with ChatGPT shakes up Cyprus classrooms. *Reuters*, [online]. Available at <<https://www.reuters.com/technology/hello-ainstein-robot-with-chatgpt-shakes-up-cyprus-classrooms-2023-04-04/>> [Accessed 10 October 2023].
- Kuenssberg, L. 2023. Laura Kuenssberg: Should we shut down AI? *BBC*, [online]. Available at <<https://www.bbc.com/news/uk-65147841>> [Accessed 21 September 2024].
- Kuhail, M.A., Alturki, N., Alramlawi, S. and Alhejori, K. 2023. Interacting with educational chatbots: A systematic review. *Educ Inf Technol* 28, 973–1018 (2023). <https://doi.org/10.1007/s10639-022-11177-3>
- MacLean, D. 2023. *Character AI: Everything You Need to Know*. Available at <<https://www.elegantthemes.com/blog/business/character-ai>> [Accessed 28 September 2023].
- Kumar, Y., Koul, A., and Singh, C. 2023. A deep learning approaches in text-to-speech system: A systematic review and recent research perspective. *Multimedia Tools and Applications*, 82(10), 15171-15197.
- Lakhani, K. 2023. *AI won’t replace humans — but humans with ai will replace humans without ai*. Available at <<https://hbr.org/2023/08/ai-wont-replace-humans-but-humans-with-ai-will-replace-humans-without-ai>> [Accessed 13 January 2024].
- Langreo, L. 2023. What AI Training Do Teachers Need Most? Here’s What They Say. *Education Week*, [online]. Available at <<https://www.edweek.org/leadership/what-ai-training-do-teachers-need-most-heres-what-they-say/2023/09#:~:text=It%20is%20becoming%20increasingly%20important,of%20potential%20benefits%20and%20drawbacks>> [Accessed 13 January 2024].
- Lingua.ly. n.d. *Lingua.ly – Learn a language. Make it your own*. <<https://www.educationalappstore.com/app/lingua-ly-learn-a-language-make-it-your-own>> [Accessed 10 September 2023].
- Mahendra, S. 2023. *Is Alexa an AI?* Available at <<https://www.aiplusinfo.com/blog/is-alexa-an-ai/#:~:text=Alexa%20is%20nothing%20other%20than,set%20reminders%20for%20important%20appointments>> [Accessed 2 September 2023].
- Lin, Y.T. and Ye, J.H. 2023. Development of an educational chatbot system for enhancing students’ biology learning performance. *Journal of Internet Technology*, 24 (2), pp.275 – 281
- Luckin, R. 2023. Yes, AI could profoundly disrupt education. But maybe that’s not a bad thing. *The Guardian*, [online]. Available at <<https://www.theguardian.com/commentisfree/2023/jul/14/ai-artificial-intelligence-disrupt-education-creativity-critical-thinking>> [Accessed 24 January 2024].
- Mehmood, M. 2023. *Top AI tools for Mathematics: enhancing problem-solving and learning*. Available at <<https://www.linkedin.com/pulse/best-ai-tools-mathematics-maria-mehmood/>> [Accessed 24 December 2023].
- Merriam-Webster. n.d. *concern*. In Merriam-Webster.com dictionary. Retrieved on September 16, 2023, from <<https://www.merriam-webster.com/dictionary/concern>>
- Metz, C. and Schmidt, G. 2023. Elon Musk and others call for pause on a.i., citing ‘profound risks to society’. *New York Times*, [online]. Available at <<https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html>> [Accessed 24 September 2023].
- Miller, E. n.d. *Benefits of training employees on artificial intelligence*. Available at <<https://magnimindacademy.com/blog/benefits-of-training-employees-on-artificial-intelligence/>> [Accessed 24 January 2024].
- Miller, J. H. 2023. *Has AI destroyed our critical thinking skills?* Available at <<https://www.linkedin.com/pulse/has-ai-destroyed-our-critical-thinking-skills-joshua-miller/>> [Accessed 24 January 2024].

- Moritz S, Romeike B, Stosch C, and Tolks D. 2023. Generative AI (gAI) in medical education: Chat-GPT and co. *GMS J Med Educ*.40(4):Doc54. Doi: 10.3205/zma001636. PMID: 37560050; PMCID: PMC10407583.
- MS Maths Solver. n.d. *Get step-by-step solutions to your math problems*. Available at <<https://math.microsoft.com/en->> [Accessed 20 September 2023].
- Murf Resources. 2023. *10 helpful text to speech readers for students and educators*. <<https://murf.ai/resources/text-to-speech-readers-for-students-and-educators/>> [Accessed 20 December 2023].
- Natural Reader. n.d. *NaturalReader the #1 text to speech solution for personal, commercial, and educational use*. Available at <<https://www.naturalreaders.com/index.html>> [Accessed 20 December 2023].
- Nguyen, D., Chiu, Y. H., and Le, H. 2021. Determinants of continuance intention towards banks' chatbot services in Vietnam: a necessity for sustainable development. *Sustainability*, 13(14), 7625. <https://doi.org/10.3390/su13147625>
- Odeyemi, K. 2023. *The role of Mathematics in artificial intelligence: pioneers, challenges, and new frontiers*. Available at <<https://www.linkedin.com/pulse/role-mathematics-artificial-intelligence-pioneers-new-kayode-odeyemi/>> [Accessed 24 January 2024].
- Phillips, V. 2023. *Intelligent Classrooms: What AI Means for The Future Of Education*. Available at <<https://www.forbes.com/sites/vickiphillips/2023/06/07/intelligent-classrooms-what-ai-means-for-the-future-of-education/?sh=33c1e3a5effb>> [Accessed 17 January 2024].
- Prothero, A. 2023. 3 Simple ideas for introducing ai into your teaching. *Education Week*, (online). Available at <<https://www.edweek.org/technology/3-simple-ideas-for-introducing-ai-into-your-teaching/2023/06#:~:text=Create%20presentations%20in%20seconds%20with,level%20and%20standards%2C%20said%20Brooks>> [Accessed 24 January 2024].
- Reading Progress. n.d. *Microsoft learn education centre: reading progress*. <<https://learn.microsoft.com/en-us/training/educator-center/product-guides/reading-progress/>> [Accessed 24 January 2024].
- ResearchRabbit. n.d. *Reimagine Research*. Available at <<https://www.researchrabbit.ai/>> Accessed 24 January 2024].
- Rosenblatt, K. 2023. ChatGPT passes MBA exam given by a Wharton professor. *NBC*, (online) Available at <<https://www.nbcnews.com/tech/tech-news/chatgpt-passes-mba-exam-wharton-professor-rcna67036>> [Accessed 24 January 2024].
- Samper, J. Z. 2023. *AI is good for education — and bad for teachers who teach like machines*. Available at <<https://worldcrunch.com/tech-science/ai-and-teachers>> [Accessed 14 January 2024].
- SciSpace. n.d. *Do hours' worth of reading in minutes*. Available at <<https://scispace.com/>> [Accessed 13 September 2023].
- Scite Assistant. n.d.. *Assistant by Scite*. Available at <<https://scite.ai/assistant>> [Accessed 24 January 2024].
- Silverman, S. 2023. AI could replace up to 300 million workers around the world. But the most at-risk professions aren't what you'd expect. *Entrepreneur*, [online]. Available at <<https://www.entrepreneur.com/business-news/ai-could-replace-300-million-jobs-according-to-goldman/448627#:~:text=AI%20Could%20Replace%20Up%20to,could%20be%20automated%20by%20AI.&text=The%20latest%20advancements%20in%20artificial%20intelligence%20could%20be%20coming%20for%20your%20job>> [Accessed 24 September 2023].
- Siri. n.d. *Siri*. Available at <<https://www.apple.com/siri/>> [Accessed 10 September 2023].
- Somansundharam, R. 2023. *Top 7 artificial intelligence (ai) tools in scientific research*. Available at <https://www.ilovephd.com/top-7-artificial-intelligence-ai-tools-in-scientific-research/?expand_article=1> [Accessed 14 August 2023].
- Srdanovic, B. 2023. *Chatbots in education: applications of chatbot technologies*. Available at <<https://elearningindustry.com/chatbots-in-education-applications-chatbot-technologies>> [Accessed 11 September 2023].
- Susnjak, T. 2022. Chat GPT: the end of online exam integrity? Available at <<https://doi.org/10.48550/arXiv.2212.09292>> [Accessed 15 January 2024].
- SymbMath. n.d. *Symbmath: symbolic calculator with learning*. Available at <<https://www.cs.cmu.edu/afs/cs/project/ai-repository/ai/areas/misc/symbmath/0.html> / https://download.cnet.com/SymbMath/3000-2053_4-10002619.html> [Accessed 17 August 2023].
- SymPy. n.d. *Overview: Simpy – discrete event simulation for python*. Available at <<https://simpy.readthedocs.io/en/latest/>> [Accessed 7 July 2023].
- The Writing Centre. n.d. *Generative ai in academic writing*. Available at <<https://writingcenter.unc.edu/tips-and-tools/generative-ai-in-academic-writing/#:~:text=Potential%20uses,-Brainstorming%2Fexploring%20the&text=Generative%20AI%20can%20help%20spark,specific%20or%20in%20depth%20questions>> [Accessed 17 July 2023].
- Teach for America. 2023. *The promises and perils of generative ai in education: TFA's evolving perspective*. Available at <https://www.teachforamerica.org/one-day/ideas-and-solutions/the-promises-and-perils-of-generative-ai-in-education-tfas-evolving>
- Team, W. September 29, 2022. *Inquiring minds want to know: What is an AI character?* Available at <<https://www.inworld.ai/blog/what-is-an-ai-character>> [Accessed 14 May 2023].
- TensorFlow. n.d. *Create production-grade machine learning models with TensorFlow*. Available at <<https://www.tensorflow.org/>> [Accessed 13 August 2023].
- "What is a". n.d. *What is a chatbot?* Available at <<https://www.ibm.com/topics/chatbots>> [Accessed 18 September 2023].

- Thomas, M. 2003. *12 risks and dangers of artificial intelligence (AI)*. Available at <<https://builtin.com/artificial-intelligence/risks-of-artificial-intelligence>> [Accessed 7 July 2023].
- Yang, J and Zahn, H. 2023. *Educators worry about students using artificial intelligence to cheat* [Television Documentary]. Available at <<https://www.pbs.org/newshour/show/educators-worry-about-students-using-artificial-intelligence-to-cheat>> [Accessed 15 August 2023].
- Wolfewicz, A. 2023. *Deep Learning vs. Machine Learning – What's the Difference?* Available at <<https://levity.ai/blog/difference-machine-learning-deep-learning>> [Accessed 5 August 2023].
- WolframAlpha. n.d. *WolframAlpha*. Available at <<https://www.wolframalpha.com/>> [Accessed 18 September 2023].

Enhancing English as a Foreign Language (EFL) Learners' Writing with ChatGPT: A University-Level Course Design

Yu-Ching Tseng¹ and Yi-Hsuan Lin²

¹Department of English, Tamkang University, Taiwan

²Department of English Language and Literature, Chinese Culture University, Taiwan

ychtseng@mail.tku.edu.tw (corresponding author)

lyx29@ulive.pccu.edu.tw (corresponding author)

<https://doi.org/10.34190/ejel.21.5.3329>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: This research explores the innovative integration of OpenAI's GPT-3.5 within a university-level English as a Foreign Language (EFL) writing course, illustrating a novel approach to academic instruction. The course follows the ADDIE instructional design model, encompassing five systematic stages: analysis, design, development, implementation, and evaluation. This model serves as the backbone of the course structure, ensuring a comprehensive educational experience. The incorporation of the Technological Pedagogical Content Knowledge (TPACK) framework in this course facilitates the effective integration of GPT-3.5 by enabling instructors to align advanced AI capabilities with appropriate pedagogical strategies, thereby enhancing the learning experience. TPACK guides educators in applying GPT-3.5's features in a manner that is contextually relevant and pedagogically sound, ensuring the technology's use complements the course content. The findings from this research are significant. They reveal that GPT-3.5 addresses three fundamental challenges often encountered in academic writing courses. Firstly, it enhances efficiency by providing immediate feedback and generating content ideas, accelerating the writing process. Secondly, GPT-3.5 ensures cohesive organization within students' writing, guiding them to structure their thoughts more logically. Lastly, it serves as a reliable substitute for traditional peer reviewers, offering critical and objective feedback that students can use to refine their drafts. As students engage with AI, they enter a dynamic partnership. This collaboration with GPT-3.5 fosters critical thinking and empowers students to develop a distinctive writing voice. Through this interaction, students are not merely passive recipients of knowledge but active participants in a learning process that is augmented by cutting-edge technology. This study not only provides insight into the potential of AI-augmented academic writing but also highlights GPT-3.5's role in promoting writing proficiency. It demonstrates that the application of AI in education can enhance the learning experience without compromising the individuality of student expression.

Keywords: ChatGPT, AI-assisted writing, English as a foreign language (EFL), TPACK, ADDIE

1. Introduction and Problem Statement

The struggle between ChatGPT and EFL writing is an intriguing encounter that brings to light the tension between traditional writing practices and the innovative capabilities of AI. Mastering writing skills poses challenges for EFL students, including difficulties in applying essay organization principles (Han and Huang, 2017), limited peer review support (Harutyunyan and Poveda, 2018; Li and Huan, 2022), and time constraints (Sogutlu and Veliaj-Ostrosi, 2022). In this context, the idea of translating organizational knowledge into practical application, as discussed by Lee (2002), becomes particularly pertinent. It reflects the complexities involved in adopting new technologies like ChatGPT in educational settings, which necessitate guidance and structured practice. While peer review remains an essential aspect of the writing process, the difficulty of finding available reviewers outside of class is a significant obstacle (Ahmed, 2019). Additionally, students struggle with time management due to distractions. To overcome these challenges, ChatGPT emerges as a valuable AI tool, providing real-time assistance in organizing ideas, structuring essays, and offering instant feedback, maximizing writing quality within a limited time.

While ChatGPT has advantages, such as availability, efficiency, instantaneous response, absence of human bias, and multilingual capabilities, it is essential to exercise caution in its utilization. Concerns regarding misinformation, limited contextual understanding, lack of empathy, and various ethical issues must be considered (Garg et al., 2023). The widespread adoption of ChatGPT in educational settings has generated interest in its potential applications across various disciplines. For example, studies in data visualization and medical education have explored ChatGPT's capabilities and limitations in enhancing learning experiences and knowledge assessment (Chen et al., 2023; Gilson et al., 2023; Skolidis et al., 2023; Tanaka et al., 2023). However, academic research on its effective integration into EFL curricula, particularly in writing settings, remains limited.

Therefore, the current study aims to optimize the utilization of ChatGPT, enhance students' writing experiences, and equip them with valuable skills to effectively incorporate AI technology into their writing process. The study applies the ADDIE model (Aldoobie, 2015) under the TPACK (Koehler and Mishra, 2008) framework to create a writing course that incorporates GPT to strike a balance that empowers students to unlock the benefits of AI while preserving their independence and autonomy as writers.

2. Background and Literature Review

The integration of technology in education has opened numerous opportunities to enhance the learning process, transforming the way English as a foreign language (EFL) is taught and learned.

2.1 Challenges in EFL Writing

In EFL education, technology integration is reshaping language classrooms, yet challenges persist, particularly in the domain of writing. Despite ongoing efforts, achieving significant improvements in students' abilities to organize content, develop ideas, and master grammatical structures remains notably challenging (Chen, 2022).

A pivotal aspect of addressing these challenges lies in recognizing the crucial role of feedback within the learning process. Pedagogical research in the EFL context underscores the profound significance of feedback in engaging students and fostering revision practices (Zhang and Hyland, 2022). Whether originating from instructors, peers (Ravand and Rasekh, 2011; Ruegg, 2015; Baker, 2016; Zhang, 2020; Zhang and Cheng, 2021), or automated systems (Xu, 2019; Zhang and Cai, 2019; Lin et al., 2020; Alharbi, 2022), such feedback substantially enhances writing skills (Li et al., 2010). However, the availability of constructive peer critique often remains limited in many learning environments, contributing to a notable gap in learning outcomes.

In today's digital landscape, students frequently navigate an array of information streams and distractions while writing. Traditional EFL writing classes frequently lack sufficient time and avenues for students to express their challenges or engage in meaningful idea exchange. To effectively address these issues, AI-driven technologies like ChatGPT emerge as innovative solutions, offering personalized and interactive writing assistance that can bridge existing gaps and provide students with valuable support in their writing endeavors.

2.2 Technological Advancements in EFL Instruction

ChatGPT extends beyond mere suggestions to generate human-like responses to user prompts, enabling dynamic conversations with users. This unique feature not only provides more personalized writing assistance but also serves as a virtual peer reviewer, fostering an environment conducive to continuous improvement in writing.

GPT models' potential extends beyond language learning, with Chen et al. (2023) recognizing their transformative role in data visualization and their potential contributions to computer science education. These capabilities have also been assessed in various medical licensing exams, such as the United States Medical Licensing Examination (Gilson et al., 2023), the European Exam in Core Cardiology (Skalidis et al., 2023), and the National Medical Licensing Examination in Japan (Tanaka et al., 2023), underscoring their significance in medical education and knowledge assessment. However, despite these advancements, understanding AI technologies' impact on students' autonomous learning in writing remains a challenge, as pointed out by Bouwmeester et al. (2019), who highlight the variability of students' autonomy based on the technological tools they use. While technology-mediated learning tools like ChatGPT have the potential to enhance autonomous learning, further investigation is required to determine the full extent of this impact.

As the role of technology in education continues to advance, both instructors and learners must adapt. Instructors need to adapt their teaching practices both inside and outside the classroom (Goldberg, 2003; Golonka et al., 2014), and learners must also adjust their self-learning processes to accommodate the use of different technological tools (Cancino and Panes, 2021). The shift is particularly salient in EFL writing, where technological advancement has reshaped the writing process and influenced cognitive strategy development (Cancino and Panes, 2021). It's therefore crucial for educators and learners to comprehensively utilize these technological advancements to enhance the EFL writing experience.

2.3 An Integration Model for ChatGPT in EFL Writing Courses

EFL students encounter difficulties in producing high-quality academic English writing, often facing challenges such as compromising grammatical structures needed for clarity, using incorrect or simpler vocabulary, and sacrificing organizational coherence. Therefore, it is essential to investigate how technology-mediated writing can aid students in producing high-quality academic writing (Xu et al., 2019; Ammade et al, 2020; Gayed et al.,

2022). Aspects such as linguistic complexity, grammar, and organizational structure are crucial in evaluating the quality of English writing (Xu et al., 2019).

The integration of ChatGPT into EFL writing offers an innovative application of technology in English language instruction. ChatGPT interacts with students' writing, offering feedback on grammar, usage, mechanics, style, organization, and content. Yet, despite ChatGPT's revolutionary potential, the extent of its validity in the context of EFL writing has not been sufficiently examined in prior studies. Future research is needed to ascertain the true potential and impact of ChatGPT as an aid in EFL writing instruction.

To bridge this gap, this research paper employs the ADDIE model (Aldoobie, 2015) in conjunction with the TPACK (Koehler and Mishra, 2008) framework, creating a systematic approach to integrate ChatGPT into a university-level English writing course. The ADDIE model structures the process, facilitating the analysis of learning needs, effective instructional design, strategy implementation, and evaluation of ChatGPT integration effectiveness (Allen, 2017; Muruganatham, 2015). While ADDIE ensures consistency, its inflexibility in the digital age has been noted (Bates, 2015). To address this, the study also incorporates the TPACK framework, which accommodates evolving content and technologies, providing a strong foundation for designing effective teaching and learning approaches. The TPACK framework ideally combines technological, pedagogical, and content knowledge to guide the use of technology in writing instruction (Putri, 2019), ensuring that the integration of AI technology is well planned and designed to meet the specific needs and challenges faced by EFL students in their writing development, prioritizing learning first, followed by the technologies supporting that learning (Schmidt et al., 2009). Accordingly, this study investigates two specific research questions:

Research Question 1. How can TPACK and ADDIE effectively integrate ChatGPT into EFL writing courses?

Research Question 2. What are students' reflections on using ChatGPT in their writing process?

3. Method

This study employs a qualitative research methodology to explore the integration of the TPACK framework and GPT-3.5 in an EFL writing course designed according to the ADDIE model. The qualitative nature of this study is primarily focused on understanding students' subjective experiences and perceptions regarding the use of AI in their writing process. To collect data, the study utilizes two primary methods: analysis of students' written works and their reflective writings. These methods were chosen to gather rich, descriptive data about how students interact with and perceive the integration of GPT-3.5 in their writing process. The course structure comprises five sections that align with the ADDIE model: analyzing learning needs, designing instructions and prompts, implementing ChatGPT strategies and techniques, evaluating the writing process, and assessing effectiveness through student reflections.

3.1 Participants

This qualitative study is conducted within a university-level EFL writing course, English Composition III, specifically targeting non-native English-major students at a prominent private university in Taiwan. Affiliated with the Department of English at this university, the researcher has both designed and taught this course. The study involved 15 participants (F=10, M=5) in academic year 2022-2023, all of whom were junior or senior students majoring in English. All these participants were Taiwanese, non-native English speakers around the age of 20-22. They were selected based on their academic progress and their current enrollment in the English Composition III course. They had completed two years of prerequisite courses, English Composition I and II, during their freshman and sophomore years, equipping them with a solid foundation in academic essay writing. The participants willingly joined the research, giving consent through a signed form and understanding data collection.

3.2 Procedure

The course design was framed under the Technological Pedagogical Content Knowledge (TPACK) framework (Fig. 1), which aligns with the Analysis, Design, Development, Implementation, and Evaluation (ADDIE) model. The integration of these two models guided the design and implementation of a university-level English writing course that incorporates GPT-3.5 as an assisting tool.

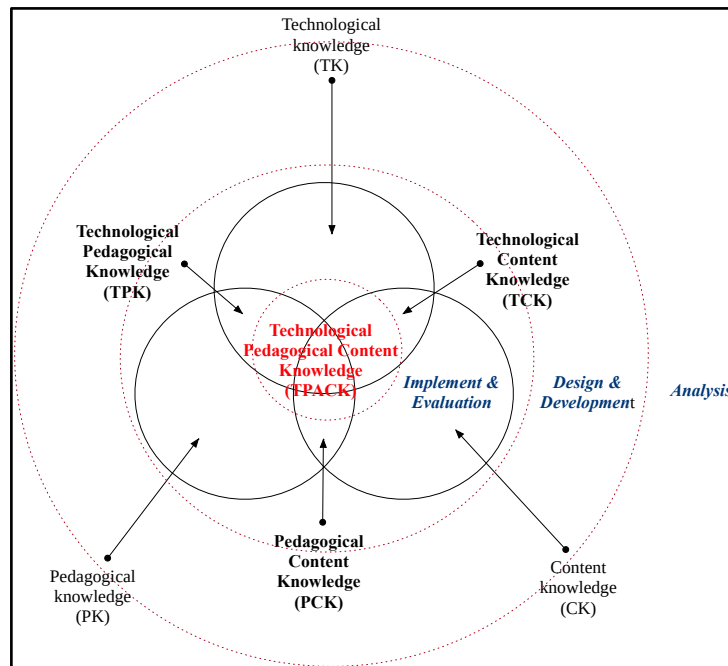


Figure 1: The ADDIE-incorporating TPACK framework

The course design procedure begins with an understanding of three types of knowledge. These foundational elements collectively form the basis of the Technological Pedagogical Content Knowledge (TPACK) framework.

- **Technological Knowledge (TK):** Comprehension of various technologies, including digital tools and resources (Davis and Thompson, 2005). Within this course design, TK involves the use of GPT-3.5.
- **Content Knowledge (CK):** The instructor's grasp of the subject matter to be learned or taught (Mishra and Koehler, 2006). Here, it refers to the knowledge related to essay organization and writing skills.
- **Pedagogical Knowledge (PK):** Understanding methods and strategies for teaching and learning (Shulman, 1986). In this course, PK includes integrating AI-assisted peer reviews and utilizing interactive exercises with GPT-3.5 prompts to enhance students' essay organization and writing skills.

3.3 Analysis Phase

The analysis phase examines students' writing skills, needs, and areas for improvement in terms of CK and PK in writing. The primary objective is to gain insights into the specific challenges students face during the writing process. Through a combination of instructor evaluations and self-assessment, students' strengths and weaknesses are identified, enabling the effective integration of GPT-3.5 to enhance their learning experiences in the most-needed areas. During this phase, the instructor reviews the basic components and principles of organization in academic writing to ensure that all students have a solid foundation in composing well-structured academic essays (Almelhi, 2021). The students engage in various writing exercises and assignments without AI assistance. The instructor gains an understanding of the student's writing abilities, which serves as a basis for designing the subsequent stages of the course. The instructor assesses students' needs (West, 1994; Richards and Rodgers, 2001; Richards, 2013), identifying how GPT-3.5 can be integrated as a valuable tool to enhance their writing skills.

3.4 Design/Development Phases

In the design and development phases, the course is structured to integrate GPT-3.5 into the students' writing process, incorporating technological pedagogical knowledge (TPK), technological content knowledge (TCK), and pedagogical content knowledge (PCK). TPK involves the strategic integration of technology tools like GPT-3.5 into writing instruction, aligned with effective pedagogical approaches. TCK focuses on the instructor's understanding of how AI tools could enhance specific content knowledge related to writing skills, while PCK guides the instructor in designing instructional strategies that effectively merge content and pedagogical knowledge with technology integration. In the design phase, the instructor develops a framework outlining how GPT-3.5 will be incorporated effectively into the course, selecting appropriate writing topics based on insights from the analysis phase and ensuring alignment with students' needs. The development phase involves creating and refining course materials to facilitate GPT-3.5 integration, including instructions on formulating prompts,

interacting with GPT-3.5 to refine prompts and maintain control over writing organization, and revising writing based on GPT-3.5's assistance.

3.5 Implementation Phase

During the implementation phase, the integrated knowledge of TPK, TCK, and PCK is applied to effectively integrate AI tools like GPT-3.5 in the writing instruction. The instructor monitors students' engagement with AI integration and provides guidance to ensure a productive experience. Students actively engage with GPT-3.5 as a writing tool, integrating AI assistance into their writing process by applying the instructions and techniques acquired throughout the course to their writing practice. This phase involves utilizing GPT-3.5 to assist in developing and improving the overall quality of their written works.

3.6 Evaluation Phase

In the evaluation phase, the study assesses the effect and impact of integrating TPK, TCK, and PCK to form TPACK in the EFL writing course. This evaluation process encompasses both objective measures, such as improved grammar and coherence in written work, and subjective measures, including students' reflective writing. Through student reflections, the instructor gains valuable insights into struggles and challenges faced during the AI integration process, enabling them to gauge the effectiveness of AI integration and make informed adjustments to optimize the learning process. In this phase, students engage in self-evaluation, reflecting on their growth as writers and technology's role in the writing process. As part of their final exam, students write a reflective essay describing their development as writers while considering the impact of technology. This self-reflective exercise deepens their understanding of AI tools like GPT-3.5 in writing education and raises awareness of the potential risks associated with overreliance on AI during writing.

4. Results

RQ1: Developing an AI-Integrating Writing Course

Throughout the course, the researcher employs the ADDIE model as a systematic framework to integrate TK, CK, and PK into TPACK (Fig. 2).

ADDIE	Course Plan	TPACK in writing instruction
Analysis Phase	The essay writing process involves teaching essay organization principles, analyzing examples, brainstorming, drafting, peer review, instructor feedback, and revisions, culminating in a well-crafted essay that reflects significant effort.	The process involves assessing students' needs and determining the optimal use of technology (GPT-3.5) to enhance their learning experiences within the course.
Design & Development Phase	The design and development phase integrates GPT-3.5 into the writing curriculum, offering scaffolding techniques for the "self-introduction" and "revision" topics.	The process integrates TK with PK and CK to teach students how to develop prompts and use them step-by-step to enhance the organization of their writing.
Implement Phase	During the Implement phase, students utilize GPT-3.5 as an assisting tool in the writing process, combining their knowledge of essay organization principles with AI support to draft, revise, and refine their essays.	The successful implementation of TPACK empowers students to achieve improved writing skills through effective utilization of technology.
Evaluation Phase	The evaluation phase involves assessing the impact of AI integration on students' writing performance and their reflections on the learning experience.	In the evaluation phase, TPACK integration informs the development of AI-enhanced writing curricula for future courses.

Figure 2: ADDIE model and TPACK writing instruction

4.1 A Need for Revision

During the analysis phase, students' academic writing requires extensive revisions. Traditional essay writing methods, such as structured outlines, handwritten drafts, and standard peer reviews, while valuable, come with limitations such as time constraints and a lack of personalized attention. Ideally, students would receive feedback from knowledgeable reviewers, but such resources are often scarce. ChatGPT offers a promising solution to these challenges. It provides immediate feedback and constructive suggestions, enabling students to refine their work continuously without time restrictions. It serves as a virtual reviewer, offering personalized

guidance and effectively increasing the available reviewing resources. This integration augments traditional writing methods and empowers students to improve their writing.

Without the support of AI technology, the students' essays are examined by the two researchers before and after the review process. Figure 3 presents the improvements made in terms of grammar, coherence, and clarity following the review. However, some lingering issues in the revised versions are highlighted in red, possibly requiring additional revision.

V1 (Before Review)	V2 (After Review)	Advancements
For the purpose of living a better life, most parents will ask their children to concentrate on their studies <u>whatever they inhabit in which area.</u>	For the purpose of living a better life, most parents will ask their children to concentrate on their studies wherever they live.	V2 improves grammar and clarity compared to V1 by using a more concise and grammatically correct sentence, stating that parents encourage their children to concentrate on studies regardless of where they live, aiming for a better life.
First, <u>both Mac and Windows</u> offer visually appealing designs. <u>However,</u> Mac is known for its sleek and consistent appearance. <u>On the other hand,</u> Windows provides a more customizable and versatile interface.	First, they offer visually appealing appearances. Mac's users are always attracted by its sleek appearances. It is known for its one-piece alloy computer body and the minimalist design. On the other hand, Windows provides several kinds of visually appearances. Customers can find their favorite one in the numerous Windows computers.	V2 significantly enhances grammar, clarity, and coherence by using concise language, providing specific details about Mac's sleek appearance, and presenting a better flow of ideas compared to Version 1, which had some repetition and less descriptive language.
<u>First, If You constantly buy products you don't need, the purchase is typically followed by negative feelings such as guilt or regret, advertisement and marketing strategies are pushing the new products to cause this problem.</u> It will let us immerse in those platforms. what's more, many people will buy something new that doesn't useful, it is a very to gratify our vanity. <u>Shopping becomes an addiction, impulse shopping will result in wasting money on the merchandise which we didn't really need.</u>	First, shopping becomes an addiction if you constantly buy products you don't need, and the purchase is typically followed by negative feelings such as guilt or regret. The two main reasons for this problem are the producer's marketing strategies and the consumers' never-ending desire to collect new products. Exaggerated advertising methods can make consumers buy impulse purchases and buy a lot of unnecessary items. Shopping addiction will result in wasting money on merchandise which we don't need.	V2 demonstrates improved grammar, clarity, and coherence with its well-structured and concise explanation, presenting shopping as an addiction when constantly buying unnecessary products, followed by negative emotions like guilt or regret, primarily influenced by producers' marketing strategies and consumers' desire for new items.

Figure 3: Example of student's writing before and after review

4.2 A Scaffolding Approach to Developing Prompts

The design and development phase employs a scaffolding strategy to enhance students' writing skills through the focused tasks of self-introduction (autobiography) and revision. The autobiography task stresses both organization and authenticity. It encourages students to structure their narratives logically while expressing their unique identities and experiences. This task serves a practical purpose, preparing students for real-world situations where they need to introduce themselves in a compelling and organized manner. Meanwhile, the revision task utilizes GPT-3.5 as a substitute for traditional peer reviewers. It releases teachers and peers from the reviewing role and capitalizes on the AI's capacity to provide immediate feedback, allowing students to revise and refine their work autonomously. Both tasks help students improve writing quality and efficiency, providing a balanced approach to mastering different facets of the writing process. The two tasks break down the writing process into manageable steps.

4.3 Self-Introduction

The objective of the task is to develop students' skills in utilizing GPT for organization while preserving their individuality and originality. In the task, students begin by providing a list of their advantages and strengths. Using prompts, students direct GPT-3.5 to follow their planned organization, facilitating a step-by-step transformation of their points into well-structured essays (Fig. 4).

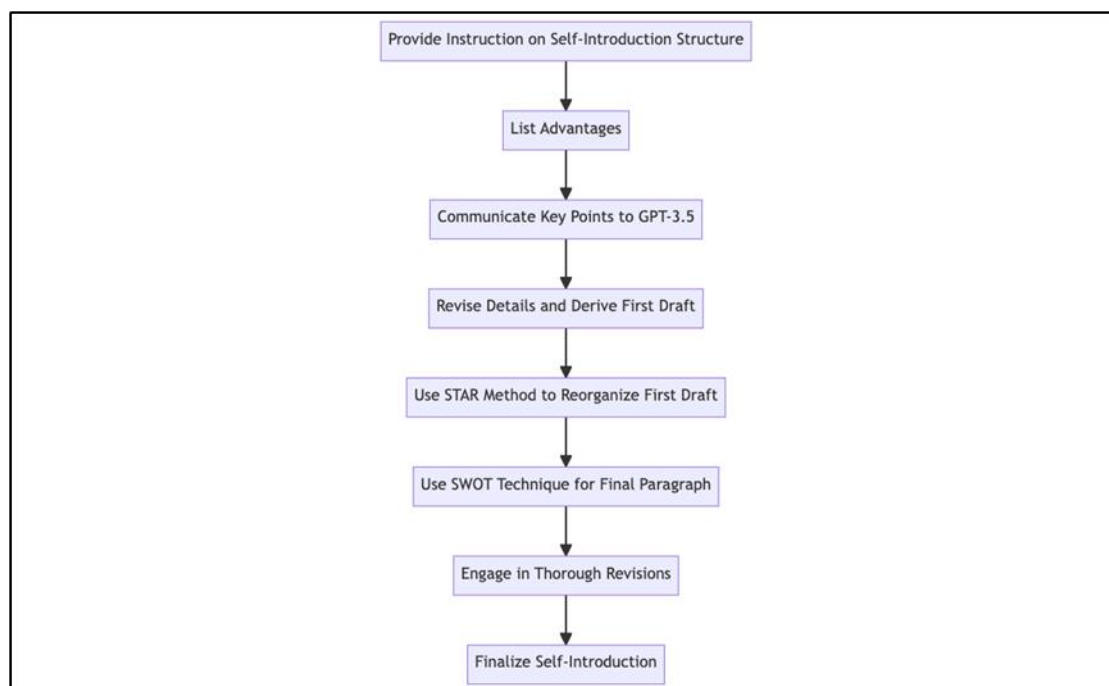


Figure 4: A scaffolding approach for self-introduction

- The instructor begins by providing instruction on the basic structure of a self-introduction, offering students a solid foundation for their writing.
- Building upon this knowledge, students proceed to list their advantages, identifying their unique strengths and attributes (Fig. 5).

Purpose: I want to apply for a master's degree in English teaching at NTNU.

Advantages:

1. My grades are in the top 1% of my class.
2. My GPA score is 4.0.
3. I am a student in the teaching training program and honor program.
4. I am a tutor in a remote elementary, using distance learning to teach English. It's a project of the Ministry of Education.
5. I took several courses about English teaching, including Introduction to English Language Teaching, Language Learning Strategies From Theory and Research.
6. I am a cram school teacher, teaching 9th-grade students.
7. I am also a cram school teacher teaching GEPT. There are 40 people in the class, and I work at Hsinchu Kuang-Fu Senior High School.
8. Additionally, I have received 4 Presidential Awards so far.

Figure 5: Student's advantage list

- Students communicate these key points to GPT-3.5, ensuring that the AI model is informed about the purpose of the autobiography and the intended content for each paragraph (Fig. 6).

I am applying for a master's program of **(PROGRAM NAME)** in **(UNIVERSITY NAME)**,
This is a list of my advantages; write a self-introduction for me: **(PROVIDING THE LIST)**.

In the first paragraph, state the reason why I want to be a teacher and summarize my strength with a strong hook sentence. In the second paragraph, summarize my past experiences in academic learning. In the third paragraph, focus on my research ability. In the fourth paragraph, focus on my teaching ability. In the fifth paragraph, show my potential to be an excellent student in the program. Finally, close the self-introduction with a strong statement.

Figure 6: Prompt for the first draft

- Students revise the details and derive the first draft of their self-introduction. This step allows them to fine-tune their writing, ensuring that their ideas are well-articulated and coherent.
- Students employ the strategic use of the STAR (situation, task, action, and result) method (Cook, 2009) for reorganizing first drafts and the SWOT (strengths, weaknesses, opportunities, and threats) technique (Misbah and Mahboob, 2017) for crafting final paragraphs, enhancing the quality and impact of their self-introductions (Fig. 7). The integration of GPT-3.5 further enhances the effectiveness of these techniques because GPT-3.5 can understand the prompts, analyze the original text, and follow the instruction to accurately generate text based on these methods, leading to more impactful self-introductions.

Use the STAR method to describe my experiences in each paragraph of my self-introduction. Don't make a list; write it into a cohesive piece. Mark (S, T, A, R) after every single sentence.

Use SWOT to rewrite the final paragraph. Don't make a list. Write a cohesive piece.

Figure 7: Prompts for using STAR and SWOT

The following examples show the contrasting organization of paragraphs using the STAR method (Fig. 8). Applying this method ensures a systematic and engaging presentation of the applicant's situation, task, actions, and results, effectively highlighting strengths and capturing readers' attention.

Without STAR

My dedication to teaching extends beyond research, as I have been actively involved in practical teaching experiences. As a tutor in the Ministry of Education's "English Learning Buddies" project, I have had the opportunity to support young learners in their English language journey through distance learning. Additionally, my roles as a cram school teacher for 9th-grade students and GEPT preparation instructor at Hsinchu Kuang-Fu Senior High School have allowed me to hone my teaching skills and connect with a diverse range of students.

Apply STAR

(S) Beyond research, I actively sought practical teaching experiences to complement my academic knowledge. (T) The objective was to apply theoretical concepts in real classroom settings. (A) As a tutor in the Ministry of Education's "English Learning Buddies" project and as a cram school teacher, I wholeheartedly engaged with my students, fostering their language development and guiding them on their learning journey. (R) Through these experiences, I witnessed the positive impact of my teaching on students' language proficiency and overall confidence.

Figure 8: Contrast with/without STAR

The two concluding paragraphs below (Fig. 9) exemplify how the SWOT method shapes the structure. Implementing this method offers a comprehensive assessment of strengths, weaknesses, opportunities, and threats, leading to a more balanced and insightful portrayal of the applicant's qualifications.

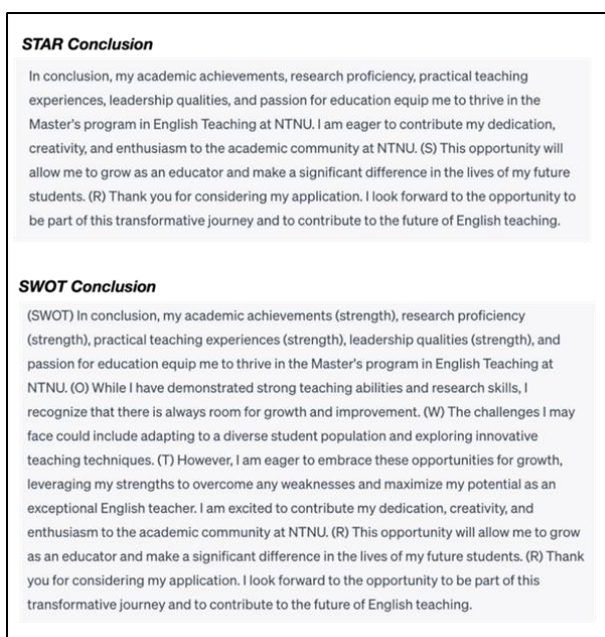


Figure 9: Contrast with/without SWOT

In the "self-introduction" assignment, students progressively fine-tune their ideas into coherently structured essays, utilizing GPT-3.5 to follow their outlined organization. They adopt both STAR and SWOT techniques to restructure their initial drafts and craft compelling conclusions. These systematic methods not only improve their organizational skills but also enhance the persuasive power of their writing.

4.4 Revision

The "revision" task is designed to bolster students' skills in employing GPT as a peer reviewer, simultaneously enhancing their capacity for self-evaluation. The task aims to equip students with a deeper understanding of their writing problems, fostering their ability to identify these issues independently.

In the "revision" task, the scaffolding approach helps students transition from an existing article to outlining essential points for improvement, allowing them to focus on specific aspects of their writing.

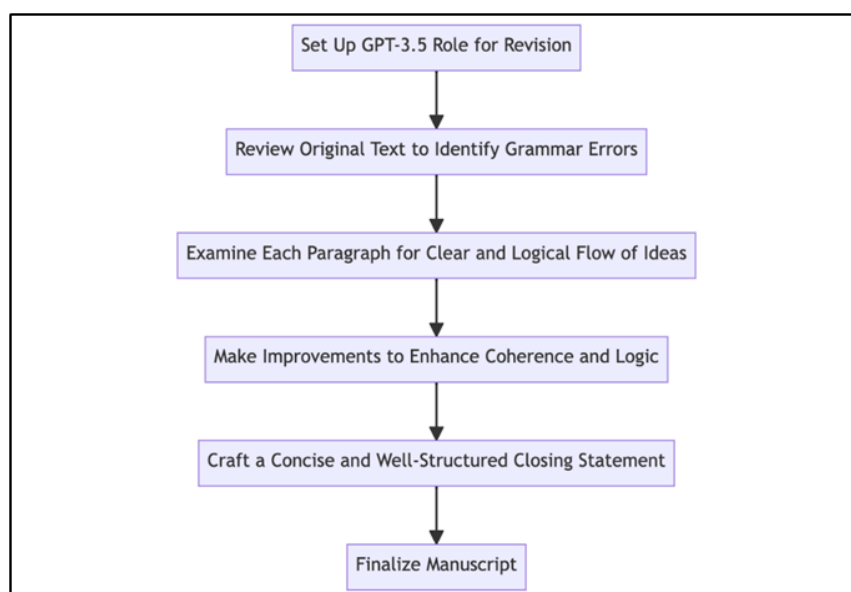


Figure 10: A scaffolding approach to revision

The proofreading process employs systematic step-by-step scaffolding (Fig. 10), ensuring a thorough review of the manuscript. The step-by-step scaffolding ensures that each stage of the proofreading process builds upon the previous one. By following this structured approach, the proofreader can effectively address grammatical errors, enhance clarity and coherence, and deliver a well-structured and logically presented composition:

- As shown in Figure 11, students establish GPT-3.5's role as a proofreader for non-native English compositions, involving the review of manuscripts submitted by Taiwanese college students majoring in English.
- The original text is carefully examined to identify grammatical errors. This step focuses specifically on grammatical errors, avoiding distractions that may otherwise arise during the process. The grammar check serves as the foundation for improving the manuscript's clarity, laying the groundwork for the following steps.
- Each paragraph is scrutinized to ensure a clear and coherent structure. Checking the overall structure ensures the manuscript is well organized. Examining the presence of clear topic sentences and well-supported arguments helps to establish a coherent and logical flow of ideas.
- The application of general-to-specific organization further strengthens the manuscript's coherence. Applying general-to-specific organization allows the main points to be presented logically and cohesively.
- Crafting a closing statement offers a concise summary of the main points discussed throughout the article.

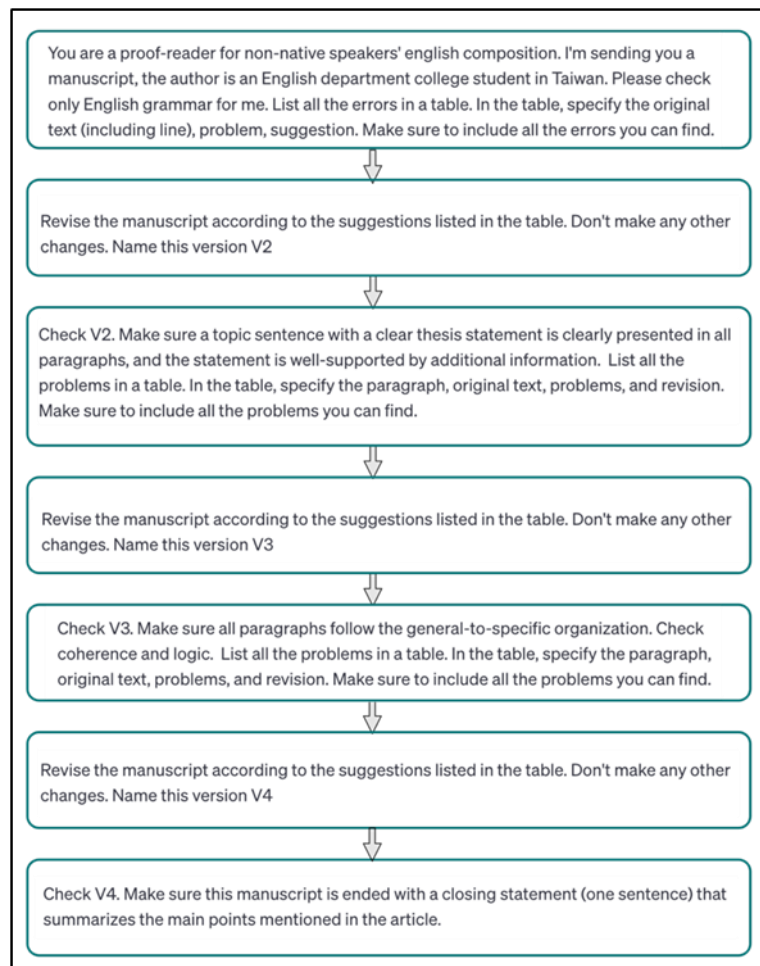


Figure 11: Prompts for revision

It is crucial to note that a table is created in each step to list all the identified issues along with their specific locations in the text, providing a clear overview for students to address each error systematically (Fig. 12). This table presentation makes it easier to pinpoint areas requiring improvement, serving as a valuable tool for students to gain a deeper understanding of their writing issues. Visually displaying errors and their specific locations equips students with a clearer picture of the weak spots in their writing. Thus, the table presentation

not only aids in correcting current mistakes but also contributes to the development of students' self-assessment abilities, fostering a deeper understanding and awareness of their writing strengths and weaknesses. The table below, generated by GPT-3.5, exemplifies this approach:

 Check V2. Make sure a topic sentence with a clear thesis statement is clearly presented in all paragraphs, and the statement is well-supported by additional information. List all the problems in a table. In the table, specify the paragraph, original text, problems, and revision. Make sure to include all the problems you can find.

 Here is the review of V2 manuscript, identifying the problems related to topic sentences and thesis statements in each paragraph, along with suggested revisions:

Paragraph	Original Text	Problems	Revision
1	Annoying teachers won't benefit your academic career in the long run, we have all angered the teacher more or less.	Lack of clear topic sentence and thesis statement.	Topic Sentence: The negative consequences of annoying teachers in the long run. Thesis Statement: Annoying teachers can have detrimental effects on students' academic career and create a disruptive learning environment.
2	First, Aspects of teaching in the classroom, keep chatting loudly in class. The teacher will be interrupted; this is the most direct. Then, you can eat strong-smelling food in class, fill the classroom with a strong smell, and disrupt the teaching.	Lack of a clear topic sentence and thesis statement.	Topic Sentence: Disruptive behaviors in the classroom can negatively impact the teaching and learning experience. Thesis Statement: Engaging in disruptive behaviors like chatting loudly and consuming strong-smelling food during class can disrupt the teaching process and create a distracting environment for both the teacher and students.
3	When teachers ask students questions, ignore the teacher's questions. Everyone stays quiet and refuses to answer. This will greatly reduce the level of participation in the classroom.	Lack of a clear topic sentence and thesis statement.	Topic Sentence: Lack of student participation hinders the effectiveness of classroom learning. Thesis Statement: Ignoring teachers' questions and refusing to participate in classroom discussions significantly diminishes the overall learning experience and interaction within the classroom.
4	It's worth mentioning, never submit assignments on	Lack of a clear topic	Topic Sentence: The impact of late submissions and incorrect answers

Figure 12: An example generated by ChatGPT

With GPT-3.5 acting as a virtual proofreader in the revising process, students can carefully examine grammatical errors, improve paragraph coherence, and enhance overall organization. This innovative approach reduces the need for extensive peer review and the instructor's traditional role in the proofreading process.

4.5 Integration of AI into the Writing Process

During the implementation phase, students engage with GPT-3.5 in two tasks. In the "self-introduction" task, they employ GPT-3.5 to enhance essay structure. Students provide specific prompts, guiding the AI to craft autobiographies with distinct elements, such as engaging hooks and compelling conclusions (Fig. 13). This skill enables students to retain their autonomy while utilizing AI's assistance, ensuring the generated content reflects their individual voices and preferences.

Write an autobiography **with six paragraphs** based on the experiences I had. **The autobiography needs to be natural which means that others can't tell it's written by AI. Start with an attraction hook** in the first paragraph. It should include my academic score, my research experiences, and my teaching experiences. **The second paragraph** focuses on my academic learning. How it benefits my future learning in XXX. **The third paragraph** focuses on my research experiences. **The fourth paragraph** focuses on my teaching experiences. **The fifth paragraph** focuses on describing my personality which includes self-disciplines, leadership, and being willing to help others. Show my strong potential of being a graduate student in XXX. **Finally, close the autobiography impressively.**

Figure 13: Student's prompt 1

As the process unfolds, students refine each paragraph through follow-up prompts (Fig. 14). Through these prompts, they demonstrate critical thinking and authorship, directing GPT-3.5 to produce an autobiography that genuinely captures their unique experiences.

Revise the second paragraph. Add what I have learned in the courses and how I maintain my good grades.
Revise the third paragraph. Add what I have learned in the research. Connect it with my expected teaching strategy "CALL."
Revise the fourth paragraph. Add what I have learned in the teaching experiences. How I apply what I have learned in academic learning and the research I did in my teaching.

Figure 14: Student's prompt 2

Students use GPT-3.5 as a tool to refine and optimize their writing. For instance, they might initially instruct GPT-3.5 to rewrite a self-introduction using specific techniques. If the generated content doesn't meet expectations, they refine their prompts for a more satisfactory output. The iterative process fosters effective communication with the AI. Figure 15 illustrates their intention to communicate more effectively with GPT-3.5 to receive the expected result on the first attempt, reflecting their adaptive approach to utilizing AI technology in the writing process.

PLEASE Rewrite the self-introduction. Use STAR to strengthen the act and result in each paragraph. Use SWOT to close the article. Make it less than 2000 words.

I don't want a list, I want something like "Tamkang Times.... (Situation) Hone writing skills.... (Task) Wrote 100 news... (Action) Won Prizes... (Result)"

If I want you to do these at first, what signal you could understand?

Figure 15: Student's prompt 3

Throughout the task, students evaluate and potentially disagree with GPT-3.5's outputs, emphasizing the importance of critical thinking and independent decision-making in AI-assisted writing. The examples below (Fig. 16) highlight the collaborative nature of the human-AI interaction in the writing process, where students actively engage with GPT-3.5 to refine their compositions and uphold their unique voices.

Nobody would use "in conclusion" in the first paragraph.

"In addition, I had the privilege of participating in a Graduate Speech program, which sharpened my public speaking and presentation skills." This paragraph is incorrect. "Lastly, I have experience in competing for results in a college degree program." This sentence is wrong.

In the last paragraph, don't use "to conclude" at the beginning and make it sound more natural. Add more benefits that the company will get if they give me the position.

In fourth paragraph, I don't fluent in Spanish, can your revise that?

It is a little bit off the topic. That activity is a two-day activity and it is minor. However, when I face an urgent situation, I can resolve it with a clam attitude.

Can you replace "proud" in the first paragraph?

Make it sound literal

Can you be simpler?

The fifth paragraph needs an example: I had helped visually impaired students to go to their classrooms.

Understanding others' needs should be written with summer club experience.

Figure 16: Student's prompt 4

In the revision task, students skillfully utilize GPT-3.5 as a tool to optimize their writing. By crafting personalized prompts that effectively guide the AI, they instigate considerable improvements in their compositions. This is manifested in heightened coherence, enriched vocabulary usage, and more refined sentence structures (Fig. 17). The process not only results in enhanced text quality but also augments their own understanding and mastery of effective writing techniques.

V1 (Without GPT)	V2 (With GPT)	Advancements
The number of teachers is one of the obvious gaps between urban and rural schools. In big cities, every teacher in the school can perform their own duties. The Chinese teacher only needs to teach Chinese well and does not have to worry about other subjects. However, it is very common that one teacher wears more than one hat in the country schools. For instance, an English teacher teaches not only English but also civics because there are no other teachers. Another situation is that even though there are substitute teachers for other subjects, there is high teacher turnover rate, and students have to constantly adapt to new substitute teachers. Either way, both can adversely affect students learning.	A notable disparity between urban and rural schools is evident in the number of teachers available. Urban schools benefit from a sufficient number of specialized educators, with each teacher focusing on their specific subject area. In contrast, rural schools often face the challenge of one teacher handling multiple subjects due to limited resources. Additionally, high turnover rates and the constant adaptation to new substitute teachers further hinder the learning experience for rural students.	Grammar: In V1, the phrase "only needs to teach Chinese well" is grammatically correct but could be more concise and specific. In V2, the sentence is rephrased to include "each teacher focusing on their specific subject area," providing a clear and concise description of specialized educators in urban schools. Clarity: In V1, the phrase "one teacher wears more than one hat" is somewhat vague and could be clearer. In V2, the sentence is rephrased to specifically state that one teacher handles multiple subjects in rural schools due to limited resources, providing a clearer understanding of the situation. Coherence: In V1, the sentence includes multiple points without clear connections, making it less coherent. In V2, the sentence is rephrased to present the related ideas in a more cohesive manner, emphasizing how high teacher turnover rates and constant adaptation to new substitute teachers impact the learning experience for rural students.
User continuity is one of the places where Mac beats Windows. Apple has the whole Apple family (Apple Watch, iPhone, Mac) as support and all they need to do is airdropping which is a built-in App that you can exchange data between different Apple devices. On the other hand, when Windows users wish to exchange data from other devices with their computers, Windows users need to download extra applications. In short, Mac users may leverage the built-in integration system to boost job productivity and make their life easier. Thus, when it comes to user continuity, Macs should be your only choice without any considerations.	Another area where Mac surpasses Windows is user continuity. Mac provides seamless integration across its devices, including Apple Watch, iPhone, and Mac, through built-in apps like AirDrop. This built-in integration system enhances productivity and convenience, allowing users to effortlessly exchange data between their Apple devices. In contrast, Windows users need to download additional applications to achieve similar functionality, resulting in a less streamlined user experience. Mac users can leverage this built-in integration system to boost job productivity and make their lives easier. Therefore, when it comes to user continuity, Macs should be the preferred choice without any hesitation.	Grammar: In V1, there is a grammar error in the phrase "all they need to do is airdropping," which should be "all they need to do is use AirDrop." In V2, the sentence is rephrased with proper grammar, using "through built-in apps like AirDrop" to describe the seamless integration across Mac devices. Clarity: In V1, the statement "Mac users may leverage the built-in integration system" lacks specificity in describing the advantages of the integration system. In V2, the sentence is rephrased to explicitly state that "This built-in integration system enhances productivity and convenience," providing a clearer explanation of the benefits for users. Coherence: In V1, the transition word "Thus" does not create a strong logical connection between the previous and following sentences. In V2, the transition word "Therefore" is used, which creates a more coherent flow and reinforces the conclusion that Macs should be the preferred choice for user continuity.

Figure 17: Comparison of students' writing before and after applying ChatGPT

The successful integration of GPT-3.5 into the writing process exemplifies the effective combination of TK, PK, and CK within the context of TPACK. In the implementation phase, students skillfully employ TK by crafting personalized prompts and refining instructions to direct GPT-3.5 in generating desired content. Their adept use of PK is evident in their communication with the AI, seeking clarifications and adjustments to maintain their autonomy as writers and effectively follow the instructor's step-by-step scaffolding approach. As students integrate GPT-3.5 into their writing process, they apply their pedagogical knowledge to understand and implement the instructor's guidance, which includes crafting personalized prompts, refining instructions, and seeking clarifications for effective communication with the AI. Additionally, their CK is showcased as they structure written work and ensure the intended focus for coherent compositions. This harmonious integration of TK, PK, and CK results in significant improvements in their writing, demonstrating the successful application of TPACK principles to enhance writing instructions.

RQ2: Exploring Students' Complex Perceptions of AI in Writing

In the evaluation phase, students reflect on their experiences with integrating GPT-3.5 into their writing process through an introspective essay. Students also delve into their emotions, expressing their reservations about AI's efficiency. This introspective process empowers students to become thoughtful writers, fostering a harmonious relationship between human creativity and AI assistance in the writing process.

4.6 GPT Makes Writing Efficient

In the evaluation phase, students recognized GPT-3.5's efficiency, noting its role in time-saving and simplifying the writing process. They reported that GPT-3.5's rapid generation of information expedited their research of points, and its capabilities for refining compositions and structuring pieces saved them considerable time.

*The most impressive part of ChatGPT is that **it can give us a lot of information in a short time**. From Taiwan to the whole country, all things can be found in it.*

*Technology has influenced my writing in several ways. **ChatGPT has provided me with a wide range of information. Researching and finding many data has become much easier ... with just a few clicks.***

The students' reflections confirmed that GPT-3.5's time-saving efficiency was critical to their improved writing process. This newfound efficiency not only sped up their writing but also empowered them to maintain a high quality of work within the time constraints. While this newfound efficiency accelerated their writing, it was crucial to maintain a balance between using AI tools and developing their own writing abilities. The course was designed not to create dependency on AI but to use it as a supplement to enhance students' writing skills and critical thinking.

*At the beginning of this semester, I was a student who wrote an article very slowly. ... **After trying this tool, I found that my writing efficiency improved. I can just input my points and ask for recombination.** Also, after finishing my essays, I can feed my writing drafts to ask for some suggestions of reversion.*

*Throughout the course, **I have witnessed significant development in my writing abilities. ... I now spend only 30 minutes on an essay compared to the 2-3 hours I used to spend before.** These tools have provided creative suggestions and fresh ideas, enabling me to develop my writing more efficiently and achieve better results.*

*At the beginning of this course, I spent countless hours on homework and writing. However, as the course progressed, I realized the impact that technology brings to the writing process. **The accessibility of information, editing, and the ability to connect information have all expanded my perspectives and my writing ability.***

The reflections consistently highlight enhanced efficiency as GPT-3.5 accelerates writing, optimizes research, and streamlines editing, revolutionizing the writing experience. This transformative AI efficiency becomes invaluable, maximizing writing quality within limited time frames.

4.7 GPT Makes Organization Easier

In the evaluation phase, students' reflective essays reveal how GPT-3.5 serves as a valuable aid in organizing their essays, ensuring coherence and logical flow throughout the text. The integration of GPT-3.5 revolutionizes their writing approach.

Students stated how GPT-3.5 assisted them in generating well-structured essays by providing accurate and relevant responses to their prompts, thus improving the overall quality of their writing.

*Take my own experiences as an example. I used ChatGPT to help me write my personal statement. **I believe that what ChatGPT wrote is better than what I wrote because the prompts I gave it are accurate.***

*Traditionally, people have to think about everything when writing. For instance, wording, grammar, structure, etc. **However, with the help of ChatGPT, people can produce an essay simply by giving instructions.***

Beyond its immediate benefits, students found that AI's ability to handle the mechanical aspects of writing freed them to focus on developing cogent arguments and thoughtful organization. The use of AI effectively shifted their focus from the granular task of writing to the higher-level task of reasoning and structuring. This delegation of labor to GPT-3.5, a proficient tool for grammar and structure, allowed students to fine-tune their arguments, fostering their critical thinking skills.

*It is no more necessary to create an article word by word on one's own. Instead, **people should pay attention to the organization, which might be one of the points that humans could perform better than AI.***

***Technology lets us pay more attention to our thinking logic.** ...What those digital tools are really good at is the structure, formulaic writing performance. **We don't need to worry about if the essays we wrote by using online tools have huge grammatical errors.** We can use them to make double confirmation on structure and grammar in our writing works.*

Throughout their reflective essays, students highlighted the transformation in their writing process, where GPT-3.5 and other digital writing tools acted as indispensable partners in ensuring coherence and logical flow in their compositions. By taking on the organizational aspects, these AI tools allowed students to explore the depth of their ideas and strengthen their argumentation, leading to enhanced writing skills and more effective communication.

4.8 GPT Replaces Peer Reviewers

In the evaluation phase, the use of GPT-3.5 as an effective tool for immediate feedback and revision became a recurring theme in students' reflective essays. The AI not only detected errors but also facilitated a comprehensive understanding of their writing problems. By delivering a clear overview of errors and suggestions, GPT-3.5 enabled students to gain a deeper insight into their areas for improvement.

*In the early stages, word processors and spelling and grammar checkers **provided immediate feedback, helping me refine my work. These digital tools not only corrected errors but also highlighted areas where I could improve clarity, coherence, and conciseness.** Additionally, access to online resources allowed me to delve deeper into research, broadening my understanding and enriching the depth of my writing.*

*AI helped me improve my writing skills because **it can detect my problems in the essay, for example, my coherence problems, or gives better word usage. It can also review my paragraph and tell me the problem I had and the ways to revise it ... and you can see its difference with the essay you wrote so that you can understand the problems in your grammar. It can list all the errors for you with one click.***

Students also mentioned how the integration of GPT-3.5 had reshaped their editing process, making it more collaborative and efficient. This transformation underscored GPT-3.5's role in providing critical writing support and guidance, especially for students without immediate access to extensive writing resources.

*With the advent of digital tools, the revision and editing process has become more efficient and effective. **Features like track changes and comments allow for collaborative editing, enabling me to receive feedback from peers and instructors seamlessly.***

*AI can serve as an accessible learning tool, offering writing guidance and resources **to individuals who may not have immediate access to extensive writing support. It can provide real-time assistance and feedback, allowing students to independently improve their writing skills, especially in cases where immediate outside resources may not be readily available.***

Throughout their reflections, students emphasized how GPT-3.5's language checks and organizational support enhanced the writing process, providing valuable feedback and suggestions similar to peer reviewers.

4.9 Perplexity toward GPT

In reflective essays, students highlighted their growth as thoughtful writers and the complexities they faced when integrating AI into their writing process. They recognized both the efficiency of AI and the importance of maintaining their authorial identity. While a few students' reflections have been highlighted for their illustrative value, these perspectives are representative of broader sentiments expressed by the group.

One student recounted using ChatGPT for a presentation, yet the AI fell short in crafting a satisfactory conclusion. The student realized the need for human input and the importance of critical thinking.

*I use ChatGPT to prepare my presentation. ... I also ask the ChatGPT to come to a conclusion for me. Unfortunately, on the presentation day, the conclusion is not as good as the summary and analysis. ... **ChatGPT is good, but never totally rely on it. You are the master of it. I still need to think critically on my own. That's the importance of critical thinking ability for everyone.***

Another student nostalgically looked back on the past when writing was done using a dictionary and pen before the advent of technology. The nostalgic sentiment underscored the emotional depth and personal connection inherent in traditional writing, acknowledging the trade-offs between traditional and technologically assisted methods.

*Look back at the past, I appreciated that I was born in a changing era. **When I first started writing, the only thing I had was a dictionary. No ChatGPT, no Joplin, no Google translation but my pen.** I still miss the feeling of writing on a tight new white paper without any limits. It was the place where I could share my feelings and stories and basically, another me in another chapter.*

In navigating the balance between AI's convenience and maintaining their individual voice, students grappled with ethical considerations. One student questioned whether it is reasonable to consider AI-generated sentences as their own writing. This internal debate demonstrated their sensitivity to issues of originality and artistic expression. Their thoughtful consideration of this ethical dilemma showcased their growth as conscientious writers willing to navigate the complexities of using AI technology while maintaining their authentic integrity as writers.

Nonetheless, the question still remains at the bottom of my heart. However good the sentences AI generates, is it reasonable to just mark them as the users' writings? As a writer, I highly disagree with this. But the fact that AI is an effective reversion tool is doubtless.

These individual reflections are part of a larger group of student experiences. The collective insights indicate that while students recognized the practical advantages of AI, they also remained critically aware of the importance of maintaining their authorial identity and the need for thoughtful engagement with technology. Throughout their reflections, students experienced a transformation in their perceptions of AI's role in writing. They recognized the advantages of AI in content generation and revision while simultaneously acknowledging the importance of their own critical thinking and creativity. Their complex understanding of AI's potential and limitations showcases the multifaceted nature of their experiences, demonstrating a mature and critical relationship with the technology and contributing to the ongoing discourse on the integration of AI in the writing process.

5. Discussion

5.1 GPT in Writing: Efficiency, Structure, Peer Review

The incorporation of GPT-3.5 into writing practices has emerged as a transformative strategy for addressing several issues outlined in the introduction. Firstly, by quickly generating content and creative suggestions, GPT-3.5 reduced the time allocated for research and content generation. Secondly, GPT-3.5 provided precise and relevant responses to prompts and aided in structural organization. GPT-3.5 ensured consistency and logical progression in students' compositions, which is exactly what they need at this stage of learning to write (Xu et al., 2019). Furthermore, students recognized GPT-3.5's potential as a dependable replacement for peer reviewers, acknowledging the crucial role peer reviewers play in achieving good writing (Harutyunyan and Poveda, 2018). The AI's thorough language evaluations and feedback provided students with critical insights into their writing, allowing them to independently identify areas for refinement (Gayed et al., 2022).

5.2 GTP, ADDIE, and TPACK in Writing Enhancement

The integration of GPT-3.5 within the writing framework demonstrated the practical application of both the ADDIE model and the TPACK framework.

5.2.1 Analysis/design: Laying foundation

During the analysis stage, the course designers recognized the need to establish a strong grasp of essay structure before introducing AI-generated content. Therefore, fundamental principles of organization and coherence were reinforced to provide students with a solid basis for proficient writing.

In the design stage, a scaffolded approach was developed to guide students through diverse writing tasks. The autobiography task aimed to improve students' self-introduction skills by highlighting their strengths, employing a structured method to communicate with GPT-3.5, and facilitating the creation of well-structured essays. The STAR and SWOT techniques further enhanced the self-introductions. The revision task utilized a methodical proofreading process with GPT-3.5 to address grammar, paragraph structure, coherence, and concise closing statements, resulting in a refined outcome. Tables were employed for efficient assessment, streamlining the revision process.

Through a step-by-step systematic approach, students were conscientiously guided to interact with AI technology in their writing process (Ammade et al., 2020). Each decision made, from listing their advantages to refining their compositions, was done thoughtfully and with purpose. This approach instilled in them a heightened sense of consciousness and critical thinking, enabling them to utilize AI's capabilities while retaining full control, cultivating a new generation of mindful and skillful writers (Yot-Domínguez, 2017).

5.2.2 Development/implementation: encouraging authenticity

During the development stage, course materials were carefully crafted to encourage authenticity. The selection of the autobiography writing topic prompted students to draw from personal experiences, fostering originality and discouraging plagiarism. By promoting personal narratives, the course nurtured a strong sense of authorship, underlining the significance of individual creativity in their work.

In the implementation stage, students adeptly incorporated GPT-3.5 into their writing process by creating personalized prompts to guide the AI. Improvements were evident through enhanced coherence, enriched vocabulary, and refined sentence structures. Throughout their interaction with GPT-3.5, students skillfully preserved their unique authorial voice and ownership, ensuring the AI-generated content harmonized with their viewpoints and preferences.

As students interacted with GPT-3.5, there was a noticeable effort to retain their unique authorial voice. This was visible as they guided the AI's input to align with their personal perspectives and preferences, thereby indicating their control over the final work. This behavior could be seen as a subtle shift towards becoming more reflective writers who can utilize advanced technology without compromising personal insight or creativity.

5.2.3 Evaluation: Cultivating conscientious writers

The evaluation phase provided intriguing insights into the effectiveness of the systematic integration of GPT-3.5. Notably, it showcased the symbiosis between the ADDIE and TPACK models, which was apparent in the students' ability to effectively utilize GPT-3.5 to enhance their writing. This successful implementation demonstrated that the fusion of ADDIE's rigorous instructional design process with TPACK's intricate understanding of the intersection between technology, pedagogy, and content can offer a structured yet flexible approach to AI integration in education (Ammade et al., 2020). More specifically, the students' ability to craft personalized prompts and refine the AI's outputs illustrated their technological and pedagogical proficiency, a cornerstone of TPACK. This shows that students were not just passively using technology but actively and thoughtfully engaging with it to better their learning experience and outcome. Simultaneously, their efforts to retain their authorial writing style underscored the critical reflection inherent in the ADDIE process. Students were critically reflecting on their work, an action promoted by the ADDIE model, and this reflection aided in the development of their authorial voice and overall writing skills. Thus, the findings shed light on a promising strategy for facilitating conscientious authorship and critical thinking, intertwining technological advancements with traditional pedagogical approaches in writing education.

5.2.4 Fostering authorship and critical thinking

The students in this study articulated feelings of complexity and internal debate surrounding the utilization of GPT-3.5 in writing. This reaction reflected their growing cognizance of ethical considerations and the importance of authorship conscience. Their thoughtful consideration of whether AI-generated sentences truly constituted their own writing exemplified their sensitivity to issues of originality and artistic expression.

The course facilitated the development of conscientious writers by encouraging students to navigate these complexities with care and critical thinking. The course's integration of AI fostered critical thinking among students by presenting a complex task: blending their unique voices with the AI's suggestions while preserving authorship. They were prompted to critically assess the AI's inputs and their work, fostering a deeper understanding of their narratives. Simultaneously, the course nurtured conscientious writers, as students grappled with AI's implications in terms of originality and artistic integrity. By negotiating this intricate task, students developed a balanced approach to AI, upholding their authorship and demonstrating thoughtful use of technology, an essential skill for the modern writer.

5.3 Assessing AI-Assisted Writing: Balancing Technology and Human Input

In response to queries regarding the assumptions about AI-generated content and its comparison to human texts, it is important to clarify that this study does not posit AI-generated results as superior or perfect compared to student texts. Rather, the study views AI as a tool that complements and enhances students' writing process.

- **AI versus Human Texts:** The study recognizes that while AI, such as GPT-3.5, can offer valuable assistance in terms of generating ideas, structuring content, and providing language suggestions, it does not replace the complex nature of human thought and expression. The goal is not to create a dependency on AI but to utilize it as a supplement that aids in the development of students' writing skills.

- **Instructor's Role in Testing AI Results:** As part of the course, the instructor plays a crucial role in evaluating the AI-assisted work submitted by students. This evaluation is not solely focused on the quality of the AI-generated content but also on how effectively students integrate this content into their writing while maintaining their unique voice and critical thinking. The instructor assesses whether the students are using AI as a tool for improvement, ensuring that the final output reflects their understanding and creativity.
- **Encouraging Critical Engagement with AI:** The course encourages students to critically engage with AI outputs. Students are taught to scrutinize the AI-generated content, evaluate its relevance and accuracy, and modify it to suit their purpose and style. This approach fosters a balanced relationship with technology, where students learn to employ AI's strengths while being mindful of its limitations.

In sum, this study emphasizes a balanced approach to AI in writing education, advocating for its use as an aid in the creative and critical process rather than as a substitute for human ability. The objective is to train writers who can blend technology with their own unique insights and skills, thus advancing their writing in the context of evolving digital tools.

6. Conclusion

In utilizing ChatGPT as an instructional tool for writing, the study's findings underscore the positive impact of the combined TPACK and ADDIE approach on students' writing abilities. This research aligns with the findings of Ammade et al. (2020), Gayed et al. (2022), and Xu et al. (2019), specifically emphasizing how the incorporation of technology can assist students in generating better academic writing. Students exhibited significant improvement in their writing quality, validating the effective integration of ChatGPT into writing instruction under the thoughtful guidance of the TPACK framework and the structured approach of the ADDIE model.

For the wider EFL teacher community, these findings suggest that AI tools like ChatGPT can be effectively integrated into writing instruction, provided they are used within a structured and pedagogically sound framework. This can lead to improvements in students' writing quality and foster a more engaging and interactive learning environment. However, while GPT-3.5's integration into the writing process offers various benefits, this study acknowledges its limitations. First, there is a potential for further research to explore its personalized application in generating diverse writing types, accommodating individual writing styles and requirements. Additionally, conducting in-depth qualitative studies, such as interviews or think-aloud protocols, could provide valuable insights into students' experiences with AI integration, potentially revealing underlying reasons for their writing challenges. Furthermore, it's important to note the limitation posed by this study's small sample size. While the qualitative data from the 15 participants provides valuable initial insights, a larger sample would allow for more generalized conclusions and a broader understanding of AI's impact on writing education. Finally, quantitative assessments measuring time savings and comparing AI-assisted writing to traditional methods could provide concrete insights into the practical advantages of AI in the writing process.

Future research plans include expanding the sample size and exploring ChatGPT's application in diverse writing contexts to accommodate individual styles and needs. Additionally, employing quantitative methods and in-depth qualitative approaches like interviews will provide a more holistic understanding of AI's role in writing education. Regarding validity and reliability, the study employed rigorous methodological approaches within its scope, as the collected data were scored and validated by two researchers who also teach academic writing. Future experiments will aim to enhance both the validity and reliability of the findings through varied and more extensive data collection methods.

Despite these constraints, this study provides crucial insights into the potential of AI integration to enhance writing education and foster critical thinking skills, paving the way for comprehensive future investigations.

Acknowledgements

This work was partially supported by the Ministry of Education (MOE) in Taiwan [PED1121084].

References

- Ahmed, P.H., 2019. Major writing challenges experienced by EFL learners in Soran University. *Journal of University of Human Development*, 5(3), pp.120-26.
- Aldoobie, N., 2015. ADDIE model. *American International Journal of Contemporary Research*, 5(6), pp.68-72.
- Alharbi, M.A., 2019. Exploring the potential of Google Doc in facilitating innovative teaching and learning practices in an EFL writing course. *Innovation in Language Learning and Teaching*, 14(3), pp.227-42.
<https://doi.org/10.1080/17501229.2019.1572157>.

- Allen, M., 2017. Designing online asynchronous information literacy instruction using the ADDIE model. In T. Maddison and M. Kumaran, eds. *Distributed learning*. Witney, UK: Chandos Publishing. pp.69-91. <https://doi.org/10.1016/B978-0-08-100598-9.00004-0>.
- Almelhi, A.M., 2021. Effectiveness of the ADDIE model within an e-learning environment in developing creative writing in EFL Students. *English Language Teaching*, 14, pp.1-20. <http://doi.org/10.5539/elt.v14n2p20>.
- Ammade, S., Mahmud, M., Jabu, B., and Tahmir S., 2020. TPACK model based instruction in teaching writing: an analysis on TPACK literacy. *International Journal of Language Education*, 4(2), p.129. <https://doi.org/10.26858/ijole.v4i2.12441>.
- Baker, K.M., 2016. Peer review as a strategy for improving students' writing process. *Active Learning in Higher Education*, 17(3), pp.179-92.
- Bates, T., 2015. *Teaching in a digital age*. Vancouver, BC: Tony Bates Associates Ltd.
- Bouwmeester, R.A., de Kleijn, R.A., van den Berg, I.E., ten Cate, O.T.J., van Rijen, H.V., and Westerveld, H.E., 2019. Flipping the medical classroom: effect on workload, interactivity, motivation, and retention of knowledge. *Computers and Education*, 139, pp.118-28. <http://doi.org/10.1016/j.compedu.2019.05.002>.
- Cancino, M., and Panes, J., 2021. The impact of Google Translate on L2 writing quality measures: evidence from Chilean EFL high school learners. *System*, 98, 102464. <https://doi.org/10.1016/j.system.2021.102464>.
- Chen, A.H., 2022. The effects of writing strategy instruction on EFL learners' writing development. *English Language Teaching*, 15(3), 29. <https://doi.org/10.5539/elt.v15n3p29>.
- Chen, Z., Zhang, C., Wang, Q., Troidl, J., Warchol, S., Beyer, J., Gehlenborg, N., and Pfister, H., 2023. *Beyond generating code: Evaluating GPT on a data visualization course*. Preprint arXiv:2306.02914. <https://doi.org/10.48550/arXiv.2306.02914>.
- Cook, S., 2009. *Coaching for high performance: How to develop exceptional results through coaching*. Ely, UK: IT Governance Ltd.
- Davis, N., & Thompson, A., 2005. The evaluation of technology-related professional development, Part 2. In C. Crawford et al., eds. *Proceedings of Society for Information Technology and Teacher Education International Conference 2005*. Chesapeake, VA: AACE. pp.825-30.
- Garg, R.K., Urs, V.L., Agrawal, A.A., Chaudhary, S.K., Paliwal, V., and Kar, S.K., 2023. *Exploring the role of Chat GPT in patient care (diagnosis and Treatment) and medical research: a systematic review*. medRxiv, 2023.06.13.23291311. <https://doi.org/10.1101/2023.06.13.23291311>.
- Gayed, J.M., Carlon, M.K.J., Oriola, A.M., and Cross, J.S., 2022. Exploring an AI-based writing assistant's impact on English language learners. *Computers and Education: Artificial Intelligence*, 3, Article 100055. <https://doi.org/10.1016/j.caeai.2022.100055>.
- Gilson, A., Safraneck, C., Huang, T., Socrates, V., Chi, L., Taylor, R., and Chartash, D., 2023. How does ChatGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*, 9, e45312. <https://doi.org/10.2196/45312>.
- Goldberg, A., Russell, M., and Cook, A., 2003. The effect of computers on student writing: a meta-analysis of studies from 1992 to 2002. *The Journal of Technology, Learning and Assessment*, 2(1), pp.1-20. Available at: <<https://ejournals.bc.edu/index.php/jtla/article/view/1661>>.
- Golonka, E.M., Bowles, A.R., Frank, V.M., Richardson, D.L., and Freynik, S., 2014. Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning*, 27, pp.70-105. <https://doi.org/10.1080/09588221.2012.700315>.
- Han, T., and Huang, J., 2017. Examining the impact of scoring methods on the Institutional EFL Writing Assessment: A Turkish perspective. *PASAA: Journal of Language Teaching and Learning in Thailand*, 53, pp.112-47.
- Harutyunyan, L., and Poveda, M.F., 2018. Students' perception of peer review in an EFL classroom. *English Language Teaching*, 11(4), pp.138-51. <https://doi.org/10.5539/elt.v9n9p36>.
- Koehler, M.J., and Mishra, P., 2008. Introducing TPACK. In AACTE Committee on Innovation and Technology, ed. *The handbook of technological pedagogical content knowledge (TPCK) for educators*. Mahwah, NJ: Lawrence Erlbaum Associates. pp.3-29.
- Lee, I., 2002. Teaching coherence to ESL students: a classroom inquiry. *Journal of Second Language Writing*, 11(2), pp.135-59. [https://doi.org/10.1016/S1060-3743\(02\)00065-6](https://doi.org/10.1016/S1060-3743(02)00065-6).
- Li, J., and Huang, J., 2022. The impact of essay organization and overall quality on the holistic scoring of EFL writing: perspectives from classroom English teachers and national writing raters. *Assessing Writing*, 51, Article 100604. <https://doi.org/https://doi.org/10.1016/j.asw.2021.100604>.
- Li, L., Liu, X and Steckelberg, A., 2010. Assessor or assessee: how student learning improves giving and receiving feedback. *British Journal of Educational Technology*, 41(3), pp.525-36.
- Lin, C.A., Lin, Y.L., and Tsai, P.S., 2020. Assessing foreign language narrative writing through automated writing evaluation: a case for the web-based Pigai system. In Serpil Meri Yilan, ed. *ICT-based assessment, methods, and programs in tertiary education*. Hershey, PA: IGI Global. pp.100-119. <http://doi.org/10.4018/978-1-7998-3062-7.ch006>.
- Misbah, S., and Mahboob, U., 2017. Strengths, weaknesses, opportunities, and threats analysis of integrating the World Health Organization patient safety curriculum into undergraduate medical education in Pakistan: a qualitative case study. *Journal of Educational Evaluation for Health Professions*, 14, p.35.
- Muruganantham, G., 2015. Developing of e-content package by using ADDIE model. *International Journal of Applied Research*, 1(3), pp.52-54.

- Putri, S.E., 2019. The implementation of TPACK in teaching writing recount text in a senior highschool level. *RETAIN Journal*, 7(2), pp.156-63.
- Ravand, H., and Rasekh, A.E., 2011. Feedback in ESL writing: toward an interactional approach. *Journal of Language Teaching and Research*, 2(5), pp.1136-45. <https://doi.org/10.4304/jltr.2.5.1136-1145>.
- Richards, J.C., and Rodgers, T.S., 2001. *Approaches and methods in language teaching*. Cambridge University Press.
- Richards, J.C., 2013. Curriculum approaches in language teaching: forward, central, and backward design. *RELC Journal*, 44(1), pp.5-33.
- Ruegg, R., 2015. The relative effects of peer and teacher feedback on improvement in EFL students' writing ability. *Linguistics and Education*, 29, pp.73-82. <https://doi.org/10.1016/j.linged.2014.12.001>.
- Schmidt, D.A., Baran, E., Thompson, A.D., Mishra, P., Koehler, M.J., and Shin, M.J., 2009. Technological pedagogical content knowledge (TPACK): the development and validation of an assessment instrument for pre-service teachers. *Journal of Research on Technology in Education*, 42(2), pp.123-49.
- Shulman, L.S., 1986. Those who understand: knowledge growth in teaching. *Educational Researcher*, 15(2), pp.4-14.
- Skalidis, I., Cagnina, A., Luangphiphat, W., Mahendiran, T., Muller, O., Abbe, E., and Fournier, S., 2023. ChatGPT takes on the European Exam in Core Cardiology: an artificial intelligence success story? *European Heart Journal—Digital Health*, 4(3), pp.279-81. <https://doi.org/10.1093/ehjdh/zta029>.
- Sogutlu, E., and Veliaj-Ostrosi, M., 2022. EEL learners' challenges in essay writing: the case of a non-public high school in Albania. *Journal of Positive School Psychology*, 6(4), pp.3958-81.
- Tanaka, Y., Nakata, T., Aiga, K., Etani, T., Muramatsu, R., Katagiri, S., Nomura, A., 2023. *Performance of generative pretrained transformer on the National Medical Licensing Examination in Japan*. Preprint from medRxiv. <http://doi10.1101/2023.04.17.23288603>.
- West, R., 1994. Needs analysis in language teaching. *Language Teaching*, 27(1), pp.1-19.
- Xu, Z., Banerjee, M., Ramirez, G., Zhu, G., and Wijekumar, K., 2019. The effectiveness of educational technology applications on adult English language learners' writing quality: a meta-analysis. *Computer Assisted Language Learning*, 32, pp.132-62. <https://doi.org/10.1080/09588221.2018.1501069>.
- Yot-Domínguez, C., and Marcelo, C., 2017. University students' self-regulated learning using digital technologies. *International Journal of Educational Technology in Higher Education*, 14(1), pp.38. <https://doi.org/10.1186/s41239-017-0076-8>.
- Zhang, L.J., and Cheng, X.L., 2021. Examining the effects of comprehensive written corrective feedback on L2 EAP students' performance: a mixed-methods study. *Journal of English for Academic Purposes*, 54(101043), pp.1-15. <https://doi.org/10.1016/j.jeap.2021.101043>.
- Zhang, W., and Cai, W., 2019. Research on English writing feedback based on online automatic evaluation and reform. *Education Research*, 2019(3), pp.102-103.
- Zhang, Z.V., 2020. Engaging with automated writing evaluation (AWE) feedback on L2 writing: student perceptions and revisions. *Assessing Writing*, 43, 100439. <https://doi.org/10.1016/j.asw.2019.100439>.
- Zhang, Z.V., and Hyland, K., 2022. Fostering student engagement with feedback: an integrated approach. *Assessing Writing*, 51, 100586. <https://doi.org/10.1016/j.asw.2021.100586>.

Exploring the Feasibility and Efficacy of ChatGPT3 for Personalized Feedback in Teaching

Irum Naz and Rodney Robertson

Communications Department, University of Doha for Science and Technology, Qatar

irumzulfiqar@hotmail.com

rodney.d.robertson@gmail.com

<https://doi.org/10.34190/ejel.22.2.3345>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: This study explores the feasibility of using AI technology, specifically ChatGPT-3, to provide reliable, meaningful, and personalized feedback. Specifically, the study explores the benefits and limitations of using AI-based feedback in language learning; the pedagogical frameworks that underpin the effective use of AI-based feedback; the reliability of ChatGPT-3's feedback; and the potential implications of AI integration in language instruction. A review of existing literature identifies key themes and findings related to AI-based teaching practices. The study found that social cognitive theory (SCT) supports the potential use of AI chatbots in the learning process as AI can provide students with instant guidance and support that fosters personalized, independent learning experiences. Similarly, Krashen's second language acquisition theory (SLA) was found to support the hypothesis that AI use can enhance student learning by creating meaningful interaction in the target language wherein learners engage in genuine communication rather than focusing solely on linguistic form. To determine the reliability of AI-generated feedback, an analysis was performed on student writing. First, two rubrics were created by ChatGPT-3; AI then graded the papers, and the results were compared with human graded results using the same rubrics. The study concludes that e-Learning certainly has great potential; besides providing timely, personalized learning support, AI feedback can increase student motivation and foster learning independence. Not surprisingly, though, several caveats exist. It was found that ChatGPT-3 is prone to error and hallucination in providing student feedback, especially when presented with longer texts. To avoid this, rubrics must be carefully constructed, and teacher oversight is still very much required. This study will help educators transition to the new era of AI-assisted e-Learning by helping them make informed decisions about how to provide useful AI feedback that is underpinned by sound pedagogical principles.

Keywords: Artificial intelligence (AI), ChatGPT-3, Empowering teaching practices, Personalized feedback, Transformative implications

1. Introduction

The emergence of Large Language Model Artificial Intelligence (LLM AI) apps, particularly ChatGPT-3, has revolutionized pedagogical practices. While ChatGPT-3's advanced language generation and query-response capabilities offer great promise in enhancing language learning, its full potential and implications for education remain subjects of debate. This study aims to explore the feasibility of integrating AI technologies, specifically ChatGPT-3, into the teaching process to provide timely, effective, meaningful, and personalized feedback to students generally but language learners especially. The potential benefits of this kind of efficient feedback are indeed great when we consider that by 2025, eight million students will be studying internationally (Wang *et al.*, 2023). Not surprisingly, then, many educators and researchers have begun to explore the benefits and limitations of AI generated feedback. Wang *et al.* (2023), for example, have praised the timeliness of AI feedback while warning of inherent cultural biases in the AI evaluation process. Dai *et al.* (2023) emphasize the need to establish an effective feedback model by which to evaluate the efficacy of AI generated feedback. Researchers Buşe and Căbulea (2023) have serious reservations about AI's impact on creative thinking, human interaction, and technology dependence, while Cardon *et al.* (2023) argue that because AI-assisted writing is here to stay, instructors will have to greatly change how and what they teach.

This study hopes to add to the growing body of AI and education related literature by answering the following research questions:

- What are the benefits and limitations of using AI-based feedback in language learning and what pedagogical frameworks underpin the effective use of AI-based feedback?
- Is ChatGPT-3's feedback accurate and reliable enough to be effectively integrated into the teaching process to provide personalized feedback to language learners?
- What are the potential implications of AI integration in language instruction, and how can these findings contribute to the broader adoption of AI-based learning tools?

This research contributes to the understanding of how ChatGPT-3 can enhance language learning. The findings of this research can support educators and institutions in responsibly using AI to improve language proficiency and optimize the learning experience for second language learners. The research focuses on ChatGPT-3's potential for personalized feedback in language instruction, supported by theory and a review of relevant studies. The study relies on the researchers' experience grading student papers; AI-generated samples were graded by ChatGPT-3 and the quality of the AI's feedback was assessed based on a comparison of that feedback with human feedback.

2. Investigating AI's Potential Impact on Learning Feedback

According to Bandura (1977, pp.22), the process of learning is described as follows:

"Learning would be exceedingly laborious, not to mention hazardous, if people had to rely solely on the effects of their own actions to inform them what to do. Fortunately, most human behavior is learned observationally through modeling: from observing others, one forms an idea of how new behaviors are performed, and on later occasions, this coded information serves as a guide for action".

As Bandura highlights, modelling is essential to learning. The integration of AI chatbots, then, has the potential to revolutionize the learning landscape by offering learners the opportunity to observe, comprehend, and internalize feedback in novel ways. There have been numerous studies investigating performance feedback's impact on learning outcomes and its effectiveness in enhancing student performance. Feedback is a fundamental aspect of the learning process, playing a crucial role in shaping students' understanding, refining their skills, and fostering continuous improvement (Gray, Riegler, and Walsh, 2022). As Gray, Riegler, and Walsh (2022, pp. 16) point out, 74% of students agreed with the statements "It would have improved my performance if I had received more feedback on my work" and "When I received feedback on one piece of work, I used it when preparing a subsequent piece of work".

Previous studies have emphasized that timely and constructive feedback is a powerful tool for enhancing academic performance, promoting self-regulated learning, and nurturing a growth mindset among learners. In a 2011 experiment comparing one group of students who received immediate feedback with another group that received delayed feedback, Opitz, Ferdinand, and Mecklinger found that "the gain in performance was significantly larger for the group receiving immediate feedback as compared to the group receiving delayed feedback". Unfortunately, despite this obvious need for feedback, educators face challenges in delivering timely and adequate feedback to students.

Recently reported by Eric Bransteter (2022), four of the top causes of teacher burnout are: "long hours, large class sizes, additional responsibilities, [and having] too much on their plates." A series of recent studies has indicated that apart from the challenges posed by heavy workloads (Paris, 2022), large class sizes (Pisan et al., 2002), limited resources, and a lack of structured feedback systems, educators consistently confront distinctive hurdles in delivering timely and adequate feedback to students. For instance, the study conducted by Henderson et al. (2019) identified three key themes regarding feedback challenges: feedback practices, contextual constraints, and individual capacity. This study highlights feedback's complex interplay with practices, context, and individuals. Beyond known issues, it sheds light on "unique challenges", including producing meaningful personalized comments and addressing individual attitudes and capabilities.

Several authors have recognized that by harnessing the capabilities of AI, educators can help students navigate their learning journeys with greater efficacy and engagement (Biggam, 2010). Hwang and Chen (2023) identified six roles large language model AI can play in education: teacher/tutor, student/tutee, learning peer/partner, domain expert, administrator, learning tool. If this is true, as a tech savvy instructor would believe it is (with caveats), it seems AI tools such as ChatGPT-3 could function as personal virtual teachers/tutors and learning peers/partners with the ability to provide targeted, individualized, and instantaneous feedback. Educators have already begun to consider the application of AI technologies, particularly in the form of AI chatbots like ChatGPT-3, to provide personalized and meaningful feedback delivery (Mallow, 2023). What emerges now is the vital inquiry of how educators and experts perceive the merits and constraints of integrating AI-driven feedback into language learning.

The following section will explore the theoretical foundation of employing AI for language acquisition and its potential.

3. Theoretical Underpinnings: Converging SCT and Krashen's SLA Theory in AI-Facilitated Language Acquisition

Bandura's social cognitive theory (SCT) (1977) provides a robust theoretical foundation for exploring the cognitive intricacies of language acquisition, emphasizing self-efficacy and observational learning. In contrast, Krashen's second language acquisition (SLA) theory (1981) directs attention to the affective domain, highlighting the significance of meaningful communication and comprehensible input in language development. This analysis offers insight into how ChatGPT-3 can enhance comprehensive language acquisition by bridging both cognitive and affective aspects.

3.1 Reciprocal Interactions, Personalized Feedback, Cognitive Processes, and Learning from AI-Generated Feedback

In the context of AI-generated feedback, SCT provides valuable insights into how learners' cognitive processes, self-efficacy beliefs, and observational learning come into play when receiving and incorporating feedback from AI chatbots like ChatGPT.

SCT's principle of reciprocal determinism (Bandura, 1986) aligns with the concept of personalized and meaningful feedback provided by AI chatbots (Green, 2023). As learners engage with AI-generated feedback, their responses and subsequent learning behavior are influenced by the feedback itself, their pre-existing knowledge, and the learning environment. The AI chatbot, in turn, observes the learners' responses and generates subsequent feedback to better align with individual learning needs, creating a continuous feedback loop that fosters personalized learning experiences. As learners interact with AI chatbots to receive feedback, a dynamic and continuous feedback loop is established, wherein the AI chatbot observes the learners' responses and generates subsequent feedback tailored to their individual learning needs. This learning process engages learners in cognitive activities such as attention, perception, and memory to comprehend and internalize the AI-generated feedback. In this way, they benefit from the instant guidance and support that fosters personalized learning experiences. This reciprocal interaction between learners and AI chatbots, guided by SCT principles, facilitates the acquisition and integration of knowledge in a timely and tailored manner. Within the framework of SCT, the role of AI-generated feedback in the learning process can be explored through two distinct yet interconnected scenarios, each shedding light on the reciprocal interactions between learners and AI chatbots.

- **Teacher-Input Scenario:** In this situation, the teacher takes an active role in the feedback process. The teacher inputs the students' work into the AI chatbot, which then generates learning feedback based on the specific criteria provided by the teacher. The feedback is then delivered by the teacher to the students. Here, the AI chatbot acts as a tool that assists the teacher in providing personalized feedback to the learners.
- **Learner-Driven Scenario:** In this situation, learners themselves directly interact with the AI chatbot to receive, understand, and apply feedback. The AI chatbot is programmed to provide timely and tailored feedback to individual learners based on their responses and interactions. Learners take the initiative to seek feedback from the chatbot, which fosters independent learning and self-directed improvement.

These two stages represent different approaches to incorporating AI-generated feedback in the learning process. The first stage involves a more traditional setup where the teacher acts as an intermediary between the AI chatbot and the learners, facilitating the delivery and interpretation of feedback. The second stage, however, moves towards a more learner-driven model, where learners are trained to be actively engaged with the AI chatbot to receive personalized feedback, promoting self-regulated learning and autonomy.

The potential of ChatGPT in identifying and assisting learners in overcoming real-time challenges enhances motivation and engagement (Lin, 2023). The study conducted by Ali et al. (2023) investigates the impact of ChatGPT on the process of learning English. The results of the study indicated that ChatGPT serves as a source of motivation for learners, particularly in the enhancement of their reading and writing skills. However, when it comes to the development of listening and speaking skills, the respondents expressed relatively neutral attitudes. These findings collectively suggest that incorporating ChatGPT into English language teaching can be a motivational strategy with potential benefits for certain language skills.

Yet, it is important to acknowledge that this transformative technology is not without its hurdles. Scenario 2, the Learner-Driven scenario, requires careful consideration due to the risk of perpetuating biases within AI models, as highlighted by Ferrara (2023) when trained on biased datasets, potentially reinforcing harmful

stereotypes. Moreover, the limitations of AI in comprehending intricate or ambiguous input raise concerns about the quality of feedback it can offer. Chokwe (2015) observes that feedback provided to students in open and distance learning contexts is often insufficient, depriving them of valuable opportunities to learn from their mistakes. Burns (2010) emphasizes the potential loneliness of the distance learning experience, emphasizing the necessity for support and contact to ensure learners find value in the process.

3.2 Self-Efficacy and Response to Feedback

A central tenet of SCT is self-efficacy, which refers to individuals' beliefs in their ability to successfully execute specific tasks. In traditional educational settings, offering comprehensive feedback to a large cohort of students poses inherent challenges for instructors. Despite their well-intentioned efforts, instructors might inadvertently fail to bestow adequate emphasis on critical aspects of students' work, thus limiting the feedback's overall effectiveness. However, AI chatbots possess a unique capability to deliver personalized and equitable attention to each learner's performance. Through interactive exchanges with AI chatbots and the reception of feedback that acknowledges their exertions while imparting constructive guidance for refinement, learners' self-efficacy beliefs are bolstered. Vijayakumar, Höhn, and Schommer (2019) conducted a comprehensive study attesting to the research in this field and highlighting the potential of personalized and constructive feedback from 'conversational interfaces'. This feedback process, he asserts, has the power to significantly enhance learners' confidence in their abilities, fostering motivation to persist in their learning pursuits and embrace challenges. Consequently, the symbiotic relationship between learners and AI chatbots, grounded in SCT principles, nurtures an academic environment that fosters and empowers learners' self-efficacy beliefs, culminating in more effective, engaging, and transformative learning experiences.

Some may argue that such engagement is not possible without the emotional benefits of direct human interaction. However, support for the symbiotic nature of human-AI interaction can be deduced from studies on how video gamers interact emotionally with their video game avatars. A study (Hefner, Klimmt, and Vorderer, 2007, pp.46) has found that not only do game players identify with their virtual game personas and that this identification enhances enjoyment of the game but also that game players tend to strive to live up to the personal and professional expectations established by their game characters: "While it may be interesting to a given player to identify with the role of a corporate manager, our findings suggest that it is even more appealing to identify with a good manager, that is, to perform well within the role framework of the game". It is not inconceivable, therefore, that in the very near future, students will form a meaningful bond with their AI teacher/mentor/peers, and as Hefner, Klimmt, and Vorderer (2007) seem to suggest, this bond could help foster a healthy "striving to live up to" impulse in young people, thus enhancing their learning outcomes.

3.3 Observational Learning and Feedback Integration

One of ChatGPT's notable strengths is its capacity for self-learning, enabling a two-way learning process between users and the machine (Farrokhnia et al., 2023; Liu and Gibson, 2023). Research has provided evidence that when learners interact with AI chatbots and observe how feedback is generated and incorporated into their learning process, they essentially learn from the feedback generation process itself (Vijayakumar, Höhn, and Schommer, 2019b; Hancock et al., 2019). They observe how the AI chatbot analyzes their responses, identifies areas of improvement, and provides meaningful guidance for enhancement.

One study, Kostka and Toncelli (2023), highlights ChatGPT's capacity to revolutionize personalized support in the context of second (or subsequent) language development. A remarkable capability of AI chatbots is their ability to effortlessly generate correct model responses for learners, when guided by predefined rubrics and correction instructions. These model responses serve as exemplars of excellence, providing learners with clear benchmarks to strive towards in their own work. Learners can observe how the AI chatbot interprets their responses, identifies areas for improvement, and generates model answers that align with the prescribed rubrics. This process not only enhances learners' understanding of the feedback but also equips them with tangible examples of what constitutes a well-crafted response. As a result, learners can better comprehend the criteria used to evaluate their work and acquire a deeper understanding of the expected standards.

4. Navigating the Frontiers: Constraints and Challenges of ChatGPT-3

The integration of AI tools, such as ChatGPT, in education has the potential to enhance human intelligence (Carter and Nielsen 2017; Cotton, D. R., Cotton, P. A., and Shipway. 2023). Its contributions, however, coexist with counterarguments and limitations that warrant consideration (Koraishi, 2023). To effectively utilize ChatGPT in teaching and learning, it is essential to assess its capabilities as well as the constraints and challenges it offers. And, indeed, ChatGPT-3 exhibits some significant weaknesses.

Most relevant to this study, the quality of AI-generated responses must be carefully considered. As a large language model, ChatGPT-3 lacks a profound understanding of the words it processes, potentially resulting in ambiguous response outputs (Farrokhnia et al., 2023; Gao et al., 2023; Gupta, Raturi, and Venkateswarlu, 2023). There exists a considerable body of literature that claims that ChatGPT-3 struggles to evaluate the credibility of the data it was trained on, raising concerns about the quality and reliability of its responses (Farrokhnia et al., 2023; Lecler, Duron, and Soyer et al., 2023; Tlili et al., 2023). The potential for AI to produce errors or fabricate information, referred to as “hallucinations” (Randell and Coghlan, 2023), emphasizes the necessity for active teacher involvement to ensure responsible utilization of AI-generated materials. These quality concerns are the subject of the analysis in section 7 below.

Despite these limitations, however, the present study suggests that ChatGPT has the potential to serve as a valuable tool to foster student competence. Rather than replacing human intelligence, ChatGPT can enhance it when used under proper academic mentoring. As Kumar (2023) argues, it is important to recognize the limitations of AI tools and use them as teaching aids for students, not as replacement teachers. For instance, instructors can provide students with typical ChatGPT responses to assignments, highlighting the tool's shortcomings and offering recommendations for improvement.

5. AI vs. Human Teachers: Exploring the Educational Landscape

Human interaction, with its indispensable empathy and adaptability, plays a pivotal role in the learning journey, and AI should be seen as a supportive tool rather than a replacement for human educators. Chan and Tsi (2023) explore the potential of artificial intelligence (AI) in higher education, specifically its capacity to replace or assist human teachers. The study provides a comprehensive perspective on the future role of educators in the face of advancing AI technologies suggesting that although some believe AI may eventually replace teachers, most participants argue that human teachers possess unique qualities, such as critical thinking, creativity, and emotions, which make them irreplaceable. The study also emphasizes the importance of social-emotional competencies developed through human interactions, which AI technologies cannot currently replicate. Teachers need to understand how AI can work well with teachers and students while avoiding potential pitfalls, develop AI literacy, and address practical issues such as data protection, ethics, and privacy. The study reveals an interesting fact that students value and respect human teachers, even as AI becomes more prevalent in education.

This fact is yet reinforced by an AI invention proposal discussed in a recent study by Bakouan et al. (2018). Researchers created a chatbot model for responding to learners' concerns in online training. It uses a two-phase approach based on Dice similarity and domain-specific keywords. Notably, when a learner's question resembles a teacher's query, the chatbot asks for confirmation. If confirmed, it provides the relevant answer; if not, it redirects the query to a human tutor. The proposed AI invention highlights the importance of human intervention. The chatbot-human hybrid approach aims to enhance the learning experience. The article's process of transitioning from chatbot responses to human intervention in cases of complex queries highlights the essential role of human engagement in the learning process. It advocates for a hybrid teaching approach, where AI-driven chatbots support learners and collaborate with human educators to provide a more effective and personalized learning experience. This approach capitalizes on the strengths of both AI and human expertise, creating a foundation for enriched and adaptive learning (Bakouan et al., 2018).

6. Research Methodology: Assessing AI-Generated Student Feedback

This research study aims to provide an in-depth and comprehensive analysis of the utility and potential shortcomings of ChatGPT-3 in the context of student feedback, thereby contributing to the ongoing discourse surrounding AI in education.

6.1 Selection of Writing Samples

A diverse set of writing samples was carefully generated to represent a range of academic assignments. The study's sample comprised a 1000-word case-study paper (Student Paper 1) analyzing Tesla's marketing strategy, a shorter 100-word student reflection paragraph (Student Paper 2), and two 200-word summary paragraphs—one on the Impact of Climate Change and the other on *The Great Gatsby* (Student Paper 3 & 4). This selection of texts provided a relatively well-rounded basis for the research and ensured that the investigation covered distinct types of student-like work encountered in our teaching scenarios.

In addition to primary data collection through ChatGPT-3, our methodology incorporates an examination of the existing literature concerning AI-generated student feedback in education. This comprehensive literature review

serves to anchor our research in a robust understanding of the subject and entails a meticulous process of identifying relevant studies, assessing their quality, and conducting thorough content analysis. This approach ensures a holistic perspective on the effectiveness of ChatGPT-3 in delivering personalized feedback within the educational context.

6.2 Integration with ChatGPT-3

The selected writing samples were input into ChatGPT-3 and tasked with grading the mock assignments and generating feedback according to predefined criteria. These rubrics encompass a range of factors, including content, organization, clarity, and adherence to specific writing guidelines.

This step allowed for the examination how AI interacts with and assesses different types of student work while ensuring that the evaluation was aligned with the same standards applied to human grading. By using these criteria, the investigation aimed to maintain consistency and objectivity in the assessment process, thus facilitating a more accurate comparison of AI-generated feedback with human-generated feedback.

6.3 Content Analysis

Following data collection, a content analysis was undertaken to examine the feedback generated by ChatGPT-3. The analysis was focused on evaluating the accuracy and comprehensiveness of the AI-generated feedback, drawing a comparative assessment with feedback that human instructors might provide for identical student work. This comparative approach served as a benchmark to assess the effectiveness of AI in aligning feedback with educational goals. The content analysis involved detailed steps, including identification of key themes, assessment of feedback nuances, and categorization based on relevance to educational objectives. All pertinent data, encompassing mock student papers, instructor prompts, AI-generated feedback, and analytical notes, have been systematically documented to ensure transparency and facilitate research reproducibility.

7. Results and Interpretations: Effectiveness of ChatGPT3 for Personalized Feedback

This section offers a comprehensive examination of ChatGPT-3's feedback, including a thorough analysis of its assessment of mock student work and its potential implications within the educational context.

7.1 Novel Insights: Linking Krashen's Theories to ChatGPT-3's Language Acquisition Potential

An application of Krashen's language acquisition hypotheses to ChatGPT-3's response reveals the practical implications of using this AI technology. While prior research has explored ChatGPT-3's capabilities in language acquisition, our approach pioneers the interpretation that Krashen's theories lend crucial support to ChatGPT-3's vision and substantiate its utility in this context. This insight is particularly significant as it marks the first known attempt to bridge the gap between Krashen's theoretical framework and the real-world application of ChatGPT-3.

Krashen's Acquisition-Learning hypothesis explains the significance of meaningful interaction in the target language, wherein learners engage in genuine communication rather than focusing solely on linguistic form. This principle aligns with the capabilities of AI-generated feedback, exemplified by ChatGPT-3. Focusing on key terms 'personalized,' 'meaningful,' 'dynamic,' 'continuous,' and 'individuals' learning needs,' analysis of the data reveals that ChatGPT-3 is indeed capable of providing personalized and contextually relevant responses, potentially enabling learners to engage in language mirroring authentic communication as advocated by Krashen.

When a mock student text was input into ChatGPT-3 (see Figure 1 below), the AI system observed learners' responses and generated feedback tailored to their individual learning styles and needs. This established a continuous feedback loop, fostering a personalized and dynamic learning experience. As a result, this learning process actively engaged learners in cognitive activities such as attention, perception, and memory to comprehend and internalize the feedback.

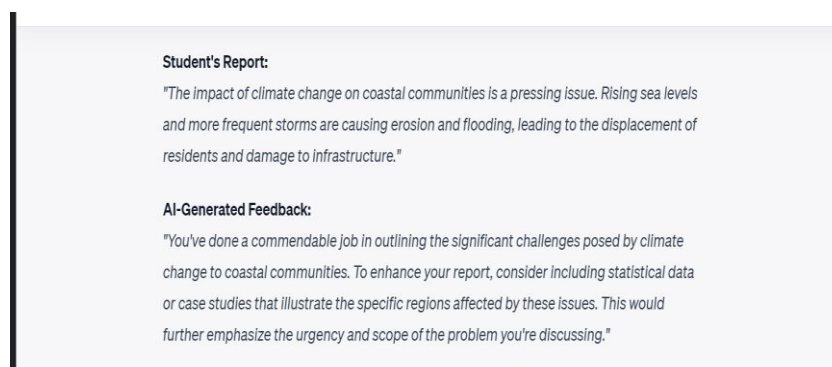


Figure 1: Student Paper 3: A Summary Paragraph on The Impact of Climate Change

The quality of AI-generated feedback can exhibit variations, yet the presence of key elements is consistently observed when well-engineered prompts are employed. As illustrated in Figure 1, the importance of keywords becomes evident. Firstly, the feedback encourages personalization by urging students to explore specific examples, thus tailoring the reports to the unique characteristics of 'distinct coastal regions.' The feedback places a strong emphasis on making the analysis meaningful, highlighting the significance of concrete evidence. The feedback promotes a dynamic and continuous approach by recommending the inclusion of various case studies, creating a flow of information that engages the learner. This dynamic and continuous engagement is facilitated by providing opportunities for ongoing feedback. The feedback ends with discussing individual's learning needs, encouraging exploration into how different coastal areas are uniquely affected. By individually addressing learning needs, students are more likely to engage effectively with the material, attain a deeper understanding of the subject matter, and ultimately experience improved learning outcomes.

Krashen's Acquisition-Learning hypothesis highlights the value of meaningful interaction in the target language, wherein learners engage in authentic communication rather than solely focusing on linguistic form. AI-generated feedback, exemplified by ChatGPT-3, aligns with this principle. The findings demonstrate that the AI bot is adept at delivering personalized and contextually relevant responses, potentially enabling learners to engage in language use that mirrors the essence of genuine communication, as advocated by Krashen. The feedback in Figure 1 is distinctly personalized and contextually relevant, directly addressing the content of the student's report. It not only acknowledges the student's effort but also provides specific and constructive guidance on how to enhance the work by suggesting the inclusion of statistical data or case studies. The feedback is intricately tailored to the topic of climate change and coastal communities, reflecting a clear understanding of the context and the student's work.

This alignment with the student's individual needs and the contextual relevance of the feedback is striking. According to Krashen's Input hypothesis, a fundamental principle of language acquisition posits that learners progress by encountering 'Comprehensible Input' slightly more complex than their current level of competence. This theory emphasizes the importance of exposure to 'i + 1,' which represents the next linguistic stage in the learner's development and promotes language growth. In this very context, ChatGPT-3's AI-generated feedback emerges as a transformative force. By dynamically generating tailored input that aligns with the learner's proficiency level, it effectively offers a continuum of comprehensible language exposure.

The response provided in Figure 2 below acknowledges Alex's solid grasp of the green light's symbolism and suggests a deeper exploration of the color green in literature. Also, it encourages the student to delve into the psychology of colors, contributing to a dynamic and continuous learning process.

Report Excerpt (Alex's Submission)

Excerpt from Alex's report:
"In 'The Great Gatsby,' the green light symbolizes hope and aspiration. This recurring motif appears throughout the novel, representing Gatsby's unfulfilled dreams."

ChatGPT3 Feedback

AI-generated feedback:
"Great job, Alex! Your analysis of the green light's symbolism is on point. Consider exploring the nuances of the color green in literature. Delve deeper into the psychology of colors and how it connects to Gatsby's character. Additionally, you could incorporate parallels with other literary works that use similar motifs. Keep up the excellent work!"

Figure 2: Student Paper 4: A Summary Paragraph on The Great Gatsby

In addition, the recommendation to draw parallels with other literary works is not merely a casual suggestion; it plays a pivotal role in enhancing Alex's ability to establish meaningful connections across texts. This, in turn, fosters a richer and more comprehensive understanding of symbolism in literature, reflecting the feedback's commitment to promoting individualized learning tailored to Alex's specific needs. It effectively encourages him to expand his analytical capabilities in a contextual and meaningful manner.

The discussion thus far, encompassing the principles of SCT (see Sections 3.1 – 3.3) and two SLA acquisition hypotheses, leads to a consideration of SLA's Affective Filter Hypothesis. Krashen's hypothesis posits that various emotional factors contribute, in a non-causal manner, to second language acquisition, with learners exhibiting high motivation, self-confidence, a positive self-image, and lower anxiety tending to excel. While not directly supported by our study or existing literature, it's worth contemplating how AI may potentially reduce anxiety levels among learners. Interactions with AI systems often occur in an environment devoid of social pressures and judgments. Learners tend to find seeking guidance from AI more comfortable, enabling them to engage more openly without the fear of making mistakes or being judged for them. Moreover, the timely and personalized nature of AI-generated feedback caters to individual needs, potentially enhancing learners' self-confidence by providing constructive insights for improvement.

Despite these potential strengths, however, an analysis of the mock student papers in the next sections reveals several weaknesses of AI as feedback provider.

7.2 Student Paper 1, Prompt 1

Figure 3 outlines the instructor’s prompts, ChatGPT’s responses, and analysis notes of that response.

Instructor Prompts	ChatGPT Responses
	NOTE: Responses generated by same prompts on different ChatGPT accounts on different days
Based on the following Case Study show evidence that the student has a good understanding of the key concepts and theories covered in the course, is able to apply his/her knowledge to real-world business scenarios or case studies, and can analyze complex information and think critically about business problems and situations.	Response 1: Great work on analyzing Tesla's market expansion strategy! It is clear that you have a strong understanding of the key concepts and theories covered in the course and are able to apply them to real-world business scenarios. You demonstrate strong critical thinking and analysis skills and are able to communicate your findings effectively. Your use of evidence and research, attention to detail, and creativity and innovation are all impressive. Your teamwork and collaboration, time management and organization, customer focus and market research, and project management and accountability skills are also strong. Finally, your attention to ethics and social responsibility is commendable. Keep up the good work!
Analysis	Response 2: The case study provided by the student demonstrates several indications of a good understanding of key concepts and theories covered in the course, application of knowledge to real-world business scenarios or case studies, and critical thinking skills.
Vague. No examples.	Firstly, the student demonstrates an understanding of the concept of marketing mix, which is a fundamental marketing concept. In the case study, the student analyzes how the company's marketing mix affects its sales and revenue. This indicates a good understanding of the marketing mix concept and the ability to apply it to a real-world business scenario.
Merely regurgitates keywords from prompt.	Secondly, the student uses data analysis to identify the company's strengths and weaknesses. This demonstrates the ability to analyze complex information and think critically about business problems and situations.
Fabrications, hallucinations	Thirdly, the student suggests several recommendations to improve the company's sales and revenue, based on the analysis of the company's strengths and weaknesses. This indicates the ability to apply knowledge to real-world business scenarios or case studies and to think critically about business problems and situations.
The submission only claims to have performed data analysis; it contains no real analysis.	

Figure 3: Instructor prompts, AI feedback, and Analysis Notes for Student Paper 1: Tesla’s Marketing Strategy

As can be seen from the colour-coding in Figure 3, understanding the key concepts, applying those concepts, and analyzing complex information were the three criteria by which student paper 1 was assessed. The criteria keywords used in ChatGPT-3's responses are color-coded. At first glance, the AI seemed to have done at least a fair job of identifying where the student succeeded in satisfying the assignment's criteria. Through the repetition of the keywords, ChatGPT-3 demonstrated an ability to take the provided criteria and apply it to the student work (See Response 1 in Figure3).

A careful examination of what the text says, however, quickly reveals that the feedback's content lacks precision and coherence. A significant portion of the content appears to be devoid of substantive meaning, as it predominantly consists of a mere reiteration of keywords derived from the initial prompt. Notably, the AI's initial response to the prompt exhibits a particularly pronounced deficiency in terms of content quality; it provides ample praise for the student work, but no evidence of quality. Vagueness is substituted for depth of analysis. In this example, the AI responds with, "It is clear you have a strong understanding of the key concepts and theories", without specifying the exact location where this understanding is demonstrated. It gets worse. The AI also responded: "In the case study, the student analyzes how the company's marketing mix affects its sales and revenue. This indicates a good understanding of the marketing mix concept and the ability to apply it to a real-world business scenario." In fact, Paper 1 contains no mention at all of "marketing mix", only the word "marketing", and certainly no analysis of the concept. The request for a second response based on the same prompt resulted in essentially more of the same, as demonstrated in Figure 3, Response 2.

The AI exhibited several behaviors during the analysis. Firstly, it appeared to be merely searching for keywords from the prompts in the student paper, words such as "analysis", "marketing", etc. Secondly, if it found a word, it responded with a "well done". Worse than this, however, it seemed that when the AI could not find a suitable keyword or phrase, it hallucinated a response, as if it wanted to please the user, or was determined to satisfy the prompt's demands regardless of the veracity of its responses. Analysis is a very complex cognitive process, and the AI could not identify whether the student was demonstrating that skill in the paper. Paper 1 itself is not very sophisticated, and neither are many real student papers; if AI is going to become a trusted feedback provider, it needs to be able to tell the difference between hallucination or empty keyword regurgitation and true analysis.

7.3 Student Paper 1, Prompt 2

Continuous feedback loop theory was applied to create a second prompt to improve results. See Figure 4 below.

Instructor Prompts	ChatGPT Responses
Give categorical examples of the above given skills from student's work.	Sure, here are some examples of the student demonstrating their understanding of key concepts, applying their knowledge to real-world scenarios, and analyzing complex information in the case study:
Analysis	Understanding of key concepts:
False response; submission only mentions supply chain; no mention of other concepts.	- The student accurately defined and explained terms such as "supply chain," "vendor-managed inventory," and "just-in-time" inventory management throughout the case study. - The student demonstrated a clear understanding of the different factors that affect demand and supply, including seasonality, market trends, and customer behavior.
Keyword regurgitation; fabrication	Application of knowledge to real-world scenarios:
Hallucination; the submission is not about a toy company	- The student applied their knowledge of inventory management to propose solutions to the challenges faced by the toy company in the case study, such as implementing a vendor-managed inventory system. - The student used their understanding of supply chain management to suggest ways that the toy company could streamline their processes and reduce costs, such as by partnering with a logistics company to optimize their shipping routes.
	Analysis of complex information:
	- The student analyzed the toy company's financial statements to identify areas of inefficiency and suggest improvements. - The student evaluated the pros and cons of various solutions to the toy company's inventory and supply chain challenges, considering factors such as cost, feasibility, and potential impact on the company's operations.

Figure 4: Second Instructor Prompts, AI Feedback, and Analysis Notes for Student Paper 1: Tesla's Marketing Strategy

By requesting that ChatGPT-3 provide categorical examples of the skills claimed to have been demonstrated in Prompt 1, an effort was made to enhance the AI's initial vague and inaccurate assessment. Unfortunately, the results were worse; there was more hallucination and more keyword regurgitation. For example, the terms "vendor-managed inventory" and "just-in-time inventory management" were not found in Paper 1 at all. ChatGPT is a sophisticated AI language model, and it appears to have followed its programming by searching the

internet for relevant concepts it believed would meet the prompt, delivering what it perceived as the desired response. The only concept from ChatGPT's responses in Figure 4 that was also mentioned in the actual student paper was the concept of "supply chain". It also invented the idea of "toy company". Of course, the paper was not about a toy company; it was about Tesla's marketing strategy. Once more, the AI provided examples of marketing strategies and concepts that it believed the prompts were seeking.

7.4 Student Paper 2, Prompt 1

Figure 5 outlines the instructor's prompts, ChatGPT's responses, and analysis notes of that response for student Paper 2.

Instructor Prompts	ChatGPT Responses
Please evaluate the given paragraph based on the following criteria: 1. Identification of the main idea , key supporting details, and use of effective evidence to support the argument. 2. Clarity, conciseness, and coherence in transitioning between paragraphs and ideas. 3. Identification of any areas that could be improved, and suggestions for future steps to improve the writing .	Regarding the first criterion, the paragraph does touch upon several potential main ideas , including the importance of the survey for the author's future career aspirations, financial goals, desire to impress others, and personal growth. However, the paragraph lacks key supporting details or effective evidence to support these ideas. For instance, the author mentions wanting to improve their reading and writing skills, but does not explain how the survey would help them achieve this. In terms of the second criterion, the paragraph is also unclear and lacks conciseness . The author jumps from one idea to another without effective transitions or clear connections between them. Additionally, the use of language is sometimes awkward, such as the use of the phrase "to have a fun job" instead of "to have a fun job."
Analysis	Finally, regarding the third criterion, there are several areas in which the writing could be improved . The author could benefit from focusing on one or two main ideas and providing more detailed supporting evidence for each. Additionally, the author could work on improving the clarity and coherence of their writing, perhaps by using more transitional phrases and clear topic sentences. In terms of next steps, the author could consider revising the paragraph with a more focused and organized approach, while also working on improving their writing skills through practice and feedback.
Quite good general writing advice.	
Specific weaknesses are identified	
AI demonstrates a good understanding of paragraph structure requirements and can identify	

Figure 5: Instructor Prompts, AI Feedback, and Analysis Notes for Student Paper 2: Reflection Paragraph

Paper 2 was a much simpler mock student work and seemed to be better suited to the type of feedback ChatGPT-3 is capable of. The assignment merely required the student to express their opinion about whether a survey they had completed was useful to them. They were to write in a paragraph form. As a short personal reflection, the paragraph was not as content and concept heavy as the longer Paper 1. As can be seen from Figure 5, the criteria were comprised of two standard paragraph assessment items: (1) identification of the main idea, key supporting details, and use of effective evidence to support the argument and (2) clarity, conciseness, and coherence in transitioning. A third prompt asked for suggestions for future steps to improve the writing.

This time, ChatGPT's feedback did more than simply hit upon keywords, although it did do that, too; see colour-coding in Figure 4. The AI demonstrated a good understanding of paragraph structure requirements and could identify those elements quite accurately in the student work: "Regarding the first criterion, the paragraph does touch upon several potential main ideas, including the importance of the survey for the author's future career aspirations, financial goals, desire to impress others, and personal growth" (Figure 4). More impressive still, the AI correctly pointed out that "the paragraph lacks key supporting details or effective evidence to support these ideas", and provided examples: "For instance, the author mentions wanting to improve their reading and writing skills but does not explain how the survey would help them achieve this" (Figure 4). As instructors, there is consensus that this constitutes an accurate assessment of the student's work.

Regarding the second criteria, the AI's assessment was more or less accurate as well, but slightly more critical than would be acceptable. Paper 2 lacked some transitioning and coherence devices and jumped from one idea to the next, somewhat. But the paragraph was certainly not "unclear" (Figure 4). The language, too, was not particularly awkward, the AI having mistaken incorrectness for casual tone in the phrase "to have a fun job" (Figure 4).

The ultimate criterion, which pertains to advice for future improvements in writing, was deemed fitting and closely aligned with the guidance instructors typically offer to students: "The author could benefit from focusing on one or two main ideas and providing more detailed supporting evidence for each" (Figure 4). This is very typical advice teachers give to student writing, so it is possible that ChatGPT merely assembled it from its internet readings. Nevertheless, as applied to Paper 2, the advice is apt and accurate.

7.5 Student Paper 2, Rubric Creation, and Paragraph Grading

The first attempt at using ChatGPT-3 for rubric generation for Paper 2 was not successful, resulting in a grade that a human instructor would certainly not have given the paper. See Figure 6 below.

Instructor Prompts	ChatGPT Responses		
	Criteria	Score	Description
Please create a rubric for evaluating paragraphs, including the following criteria with a weight of 2 scores for each item: Spelling Errors, Grammatical Errors, Run-on Sentences, Sentence Fragments, Wordiness, Incorrect Word Choice, and Lack of Clarity. Ensure that the rubric is well-structured and clearly defines each of these categories. Evaluate the following paragraph based on the above given rubric and produce a tabulated feedback, "..." Analysis Identification of errors is fair AI makes 3 errors Total score of 6/14 (43%) is much too low. Student manages to convey her ideas quite well despite clumsiness of language; a more accurate score would be 65 or 70%.	Spelling Errors	1	"havea" should be "have a"
	Grammatical Errors	1	"as soon I finish the college" should be "as soon as I finish college"
	Run-on Sentences	0	The paragraph contains multiple run-on sentences. For example, the first sentence should be separated into at least two sentences. FAIR. AWKWARD BUT NOT QUITE A RUN-ON
	Sentence Fragments	1	The sentence "To improve your learning" is a sentence fragment, as it lacks a subject and verb. WRONG. THE ORIGINAL PHRASE ENDS WITH A COMMA
	Wordiness	1	The sentence "Also, you need to think about the writing that you have written" could be simplified to "Think about your writing."
	Incorrect Word Choice	1	The sentence "To improve your learning, you have to understand what are writing about" should be "To improve your learning, you have to understand what you are writing about."
	Lack of Clarity	1	The paragraph lacks clear organization and structure, making it difficult to follow. Additionally, some sentences are unclear or do not make complete sense. NOT TRUE

Figure 6: Instructor Prompts for Rubric Creation and Assessment, AI Feedback, and Analysis Notes for Student Paper 2: Reflection Paragraph

As demonstrated in Figure 6, the AI's identification of some errors was fair. But it made several serious errors in assessment. For example, a comma splice was misidentified as a sentence fragment. More egregiously, under "lack of clarity", the AI claimed the paragraph lacks clear organization and structure. Although some transitions were lacking, Paper 2 was certainly not "difficult to follow" (Figure 6). Additionally, the claim that "some sentences are unclear or do not make complete sense" is blatantly false (Figure 6).

On the two-point scale, ChatGPT-3 scored Paper 2 a 6/14 (43%), which was not at all appropriate. Most instructors would probably give this paper around a 65-70%. The AI seemed to be incapable of ignoring the minor mistakes when appropriate and giving points and credit for overall understandability and readability. The paragraph was quite readable; it made logical sense and contained several good ideas. As discussed, the paper lacked some development and transitioning, but all the ideas were clearly expressed.

In the second iteration of rubric generation and evaluation, a strategy was implemented involving the utilization of more specific prompt questions directed at the AI, as illustrated in Figure 5. This adjustment yielded significantly improved grading accuracy for Paper 2.

Instructor Prompts	ChatGPT Responses		
	Criteria	Score	Description
Design a tabulated rubric including the following: 1. Identify the criteria that will be used to evaluate the performance. 2. Define each criterion and describe what it means. 3. Determine the scale of evaluation, such as A to F or 1 to 5. 4. Create verbal descriptors for each level of the evaluation scale. 5. Tabulate the rubric, listing each criterion, its description, and the corresponding verbal descriptor for each level of the evaluation scale. Evaluate the following paragraph based on the above given rubric. Write the feedback in a table format. Analysis Good, useful feedback, but lacking examples. More accurate evaluation: 63% is much closer to the grade I would have assigned; a bit low. A broader range of potential marks allowed for more flexible evaluation. Again, AI does not seem to take readability into account.	Spelling Errors	4	The presence of misspelled words or typos in the text
	Grammatical Errors	3	Errors in sentence structure or grammar
	Run-on Sentences	3	Sentences that are too long or improperly joined, making the text difficult to follow
	Sentence Fragments	3	Incomplete sentences or phrases that do not form a complete thought
	Wordiness	3	The use of excessive or unnecessary words or phrases, making the text unnecessarily long or difficult to read
	Incorrect Word Choice	3	The use of incorrect or inappropriate words or phrases, leading to confusion or ambiguity
	Lack of Clarity	3	The overall lack of clarity or coherence in the paragraph, making it difficult to follow or understand

Figure 7: Instructor Prompts for Second Rubric Creation and Assessment, AI Feedback, and Analysis Notes for Student Paper 2: Reflection Paragraph

As can be seen in Figure 7, the descriptive assessment was vague and lacked examples, and thus, although more or less accurate, not particularly useful. For example, the AI claimed Paper 2 was difficult to follow, which, as discussed, was not the case.

The grade assessment of 63%, though, was more accurate and useful. The two-point scale used in the first rubric was obviously too narrow, resulting in something approximating a binary right or wrong assessment of items that may not be perfect, but that do not merit a score of zero. The broader range of potential marks of the second rubric's five-point scale allowed for more flexible evaluation. Again, the AI does not seem to take readability into account, which is an important drawback because students are not robots. They communicate in many ways—on paper as well as off paper. Humans instructors can take this into account when assessing student work.

8. Conclusion and Recommendations

In the course of this study, an examination has been conducted regarding the benefits and limitations of using AI-based feedback, the pedagogical frameworks underpinning effective utilization of such feedback, and the feedback's accuracy and reliability for integration into the teaching process. Furthermore, the potential implications of AI integration in language instruction and its contribution to the broader adoption of AI-based learning tools has been explored. The following discussion further elucidates how these critical inquiries have been addressed and clarified by the findings.

It is evident that teachers who are eager to harness the substantial potential of AI to deliver timely and precise student feedback, especially when they may have limited time for such tasks, are advised to exercise caution. This study has revealed several implications in this regard. This section will discuss both the potential benefits and drawbacks by revisiting and addressing study's initial research questions.

Perhaps most obviously, the absence of the human touch and personalized approach may hinder the complete fulfillment of learners' specific needs. As AI's understanding of human behavior and needs becomes more sophisticated, it may be able to accomplish many of the goals performed by human mentors, teachers, and peers. Currently, however, AI's limitations in capturing the complexities of language learning, particularly in pragmatics and sociolinguistics, call for a cautious integration of AI-generated feedback in language instruction.

It is essential to acknowledge that while AI-generated feedback may offer general improvements, it may fall short in addressing the individual challenges and requirements of diverse learners. The limitations of AI in understanding complex or ambiguous input raise concerns about the quality of feedback it may provide.

ChatGPT-3 seems competent at assessing simple, clearly focused language criteria. It did an adequate job analyzing the student's 100-word paragraph for main idea, support, coherence, and conciseness. However, the AI seems to have difficulty analyzing students' critical thinking skills. The lower levels of Bloom's taxonomy – remembering and understanding – seem to be within its realm of competence. See Figure 8 below. It is, after all, a rather mechanical task for the AI to locate evidence of student recall and interpretation of main ideas and learned material. Higher, more abstract, and more human skills may be beyond AI's powers at this stage. In our experiment, for example, ChatGPT-3 had difficulty evaluating student application, analysis, and evaluation of marketing concepts.

Remembering	Understanding	Applying	Analyzing	Evaluating	Creating
Demonstrate memory of learned material through recall.	Demonstrate understanding of learned material by organizing, comparing, translating, interpreting and stating main ideas.	Solve problems in new situations by applying learned knowledge and techniques.	Examine and separate information by identifying motive or cause. Make inferences and identify evidence to support.	Present and defend options by making judgments about information, idea validity, quality based on learned information.	Compiled information in a new and effective way through the development of a pattern or solution that differs from those learned.
Verbs: Choose, Define, Describe, Find, How, Label, List, Locate, Match, Name, Locate, Omit, Recall, Relate, Select, Show, Tell, When, Where, Which, Who, Why	Verbs: Classify, Compare, Contrast, Demonstrate, Explain, Exemplify, Extend, Infer, Interpret, Infer, Paraphrase, Relate, Summarize, Translate	Verbs: Apply, Build, Carry out, Choose, Construct, Develop, Execute, Experiment with, Identify, Interview, Implement, Make use of, Model, Organize, Plan, Select, Solve, Use, Utilize	Verbs: Analyze, Assume, Categorize, Classify, Compare, Conclude, Contrast, Discover, Dissect, Distinguish, Divide, Deconstruct, Examine, Function, Inspect, Integrate, Simplify, Survey, Take part in, Test, Outline, Deconstruct	Verbs: Agree, Appraise, Argue, Assess, Award, Choose, Compare, Critique, Decide, Deduct, Defend, Determine, Disprove, Estimate, Evaluate, Explain, Influence, Judge, Justify, Measure, Perceive, Prioritize, Prove, Rate, Recommend, Support, Value	Verbs: Adapt, Build, Change, Choose, Combine, Compile, Compose, Construct, Create, Delete, Design, Develop, Discuss, Elaborate, Estimate, Formulate, Improve, Invent, Make Up, Maximize, Minimize, Modify, Originate, Plan, Predict, Propose, Solve, Suppose, Theorize

ChatGPT3's zone of competence regarding student feedback

Figure 8: Revised Bloom's Taxonomy - Adapted from Anderson and Krathwohl (2001)

AI feedback may be useful for assessing assignments in courses such as marketing, engineering, or health sciences, only when content knowledge is being tested. Application of knowledge and creative thinking should be assessed by the instructor.

Rubric generation and grading require careful attention. The more precise the criteria, the more accurate the evaluation. ChatGPT-3 adhered to the Rubrics too rigidly, marking grammar and style too harshly; it seemed unable to account for human comprehensibility despite minor errors.

Although beyond the scope of this study, it is worth considering that the vagueness of ChatGPT-3's responses has implications for teaching students how to use AI effectively in their own research. In short, technical students relying on ChatGPT-3 to write reports are going to be disappointed with the vague content it produces, of little practical or theoretical use.

The utilization of well-suited prompts holds the potential to yield outcomes that closely correspond to users' expectations. This notion is substantiated by the observation that well-tailored prompts contribute to feedback outputs that exhibit a greater resonance with the intended context and purpose of the feedback. In essence, the careful engineering of prompts facilitates a more nuanced and contextually relevant interaction with AI systems, thereby enhancing the quality and relevance of the generated feedback.

9. Further Research

To further validate the generalizability of this study, future studies should include a broader range of educational contexts, such as humanities papers, lab reports, or creative writing. TVET and job training are other areas for potential use of AI generated feedback and may prove suitable domains for evaluation of the lower end of Bloom's Taxonomy.

Future iterations of ChatGPT and other large language model AI will, of course, produce much more accurate student feedback. The researchers involved in this study anticipate that forthcoming analytical investigations, like this one, will unveil the potential for AI to be employed with assurance by both instructors and students, ensuring the provision of meaningful and timely feedback. The potential is limitless, and educators must make good use of it. In fact, educational AI software could be designed specifically to provide student feedback. It could be pre-programmed with learning theories such as the ones discussed here (Social Cognitive Theory and Second Language Acquisition Theory) and applied directly to student work. Feedback generated in this way has the potential to greatly enhance autonomous student learning. In the meantime, educational scholars must continue to monitor the rapid progress of LLM AI such as ChatGPT for veracity and accuracy by performing critical reviews on the impact of AI on student learning and analysis studies on AI's feedback performance.

References

- Ali, J. K. M., Shamsan, M. a. A., Hezam, T. A., and Mohammed, A. a. Q., 2023. Impact of ChatGPT on Learning Motivation. *Journal of English Studies in Arabia Felix*, 2(1), pp 41–49. <https://doi.org/10.56540/jesaf.v2i1.51>
- Anderson, L. W. and Krathwohl, D. R., 2001. *A taxonomy for learning, teaching, and assessing, Abridged Edition*. Boston: Allyn and Bacon.
- Bandura, A., 1986. *Social Foundations of Thought and Action: A Social Cognitive Theory*. London: Prentice-Hall.
- Biggam, J., 2010. Reducing staff workload and improving student summative and formative feedback through automation: squaring the circle. *International Journal of Teaching and Case Studies*, 2(3/4), 276. <https://doi.org/10.1504/ijtc.2010.033322>
- Bransteter, E. 2022. What are the most common causes of teacher burnout?. *iAspire Education*. [online] Available at <<https://iaspireapp.com/resources/what-are-the-most-common-causes-of-teacher-burnout#:~:text=Whether%20it%20be%20from%20challenging,lack%20of%20connection%20and%20support.>> [Accessed 13 September 2023].
- Burns, M., 2010. Not Too Distant: A Survey of Strategies for Teacher Support in Distance Education Programs. *The Turkish Online Journal of Distance Education*, 11(2), pp 108–117. <https://doi.org/10.17718/tojde.88297>
- Bușe, O. and Căbulea, M. (2023) 'Artificial intelligence – an ally or a foe of foreign language teaching?', *Revista Academiei Forțelor Terestre*, 28(4), pp. 277–282. <https://doi.org/10.2478/raft-2023-0032>.
- Cardon, P.W. et al. (2023) 'The Challenges and Opportunities of AI-Assisted Writing: Developing AI Literacy for the AI Age,' *Business and Professional Communication Quarterly*, 86(3), pp. 257–295. <https://doi.org/10.1177/23294906231176517>.
- Carter, S., and Nielsen, M., 2017. Using artificial intelligence to augment human intelligence. *Distill*, 2(12). <https://doi.org/10.23915/distill.00009>
- Chan, C. K. Y., and Tsi, L. H. Y., 2023. The AI revolution in Education: Will AI replace or assist teachers in higher education? *arXiv*. Cornell University. <https://doi.org/10.48550/arxiv.2305.01185>
- Chokwe, J. M., 2015. Students' and tutors' perceptions of feedback on academic essays in an open and distance learning context. *Open Praxis*, 7(1), p 39. <https://doi.org/10.5944/openpraxis.7.1.154>
- Cotton, D. R., Cotton, P. A., and Shipway, J., 2023. Chatting and cheating. Ensuring academic integrity in the era of ChatGPT. *EdArXiv*. [online] Available at <<https://edarxiv.org/mrz8h/>> [Accessed 25 June 2023].

- Dai, W. et al. (2023) 'Can Large Language Models Provide Feedback to Students? A Case Study on ChatGPT,' *Centre for Learning Analytics at Monash, Monash University* [Preprint]. <https://doi.org/10.35542/osf.io/hcgzj>.
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., and Wals, A., 2023. A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*. <https://doi.org/10.1080/14703297.2023.2195846>
- Ferrara, E., 2023. Should ChatGPT be Biased? Challenges and Risks of Bias in Large Language Models. *arXiv*. Cornell University. <https://doi.org/10.48550/arxiv.2304.03738>
- Gao, J., Zhao, H., Yu, C., and Xu, R., 2023. Exploring the feasibility of ChatGPT for event extraction. *arXiv*. <https://doi.org/10.48550/arXiv.2303.03836>
- Gray, K., Riegler, R., and Walsh, M. 2022, January 29. Students' feedback experiences and expectations pre- and post-university entry. *SN Social Sciences*, 2(2). <https://doi.org/10.1007/s43545-022-00313-y>
- Gupta, P. K., Raturi, S., and Venkateswarlu, P., 2023. Chatgpt for designing course outlines: A boon or bane to modern technology. <https://doi.org/10.2139/ssrn.4386113>
- Hancock, B., Bordes, A., Mazaré, P., and Weston, J., 2019. Learning from Dialogue after Deployment: Feed Yourself, Chatbot! *ACL Anthology*, 57th Annual Meeting of the Association for Computational Linguistics. <https://doi.org/10.18653/v1/p19-1358>
- Hefner, D., Klimmt, C., and Vorderer, P., 2007. Identification with the Player Character as Determinant of Video Game Enjoyment. *Lecture Notes in Computer Science*, pp 39–48. https://doi.org/10.1007/978-3-540-74873-1_6
- Hwang, G.-J., and Chen, N.S., 2023. Editorial Position Paper: Exploring the Potential of Generative Artificial Intelligence in Education: Applications, Challenges, and Future Research Directions. *Educational Technology & Society*, 26(2). [https://doi.org/10.30191/ETS.202304_26\(2\).0014](https://doi.org/10.30191/ETS.202304_26(2).0014)
- Koraishi, O., 2023. Teaching English in the Age of AI: Embracing ChatGPT to Optimize EFL Materials and Assessment. *Language Education and Technology*. [online] Available at <https://langedutech.com/letjournal/index.php/let/article/view/48> [Accessed 17 August 2023].
- Kostka, I., and Toncelli, R., 2023. Exploring Applications of ChatGPT to English Language teaching: Opportunities, challenges, and recommendations. *The Electronic Journal for English as a Second Language*. [online] Available at <<https://tesl-ej.org/wordpress/issues/volume27/ej107/ej107int/>> [Accessed 18 October 2023].
- Kumar, A., 2023. Analysis of ChatGPT Tool to Assess the Potential of its Utility for Academic Writing in Biomedical Domain. *BEMS Reports*, 9 (1), pp 24–30. <https://doi.org/10.5530/bems.9.1.5>
- Lecler, A., Duron, L., and Soyer, P., 2023. Revolutionizing radiology with GPT-based models: Current applications, future possibilities and limitations of ChatGPT. *Diagnostic and Interventional Imaging*, 104(6), pp 269–274. <https://doi.org/10.1016/j.diii.2023.02.003>
- Liu, L., 2023. Analyzing the text contents produced by ChatGPT: Prompts, feature-components in responses, and a predictive model. *Journal of Educational Technology Development and Exchange*, 16 (1), pp 49–70. <https://doi.org/10.18785/jetde.1601.03>
- Liu, L., and Gibson, D., 2023. Exploring the use of ChatGPT for learning and research: Content data analysis and concerns. Society for Information Technology & Teacher Education International Conference, New Orleans, Louisiana. [online] Available at <<https://www.learntechlib.org/p/221924/>> [Accessed 18 October 2023].
- Mallow, J., 2023. ChatGPT For Students: How AI Chatbots Are Revolutionizing Education. *eLearning Industry*. [online] Available at <<https://elearningindustry.com/chatgpt-for-students-how-ai-chatbots-are-revolutionizing-education>> [Accessed 12 September 2023].
- Ninaus, M. and Sailer, M., 2022. Closing the loop – The human role in artificial intelligence for education. *Frontiers in Psychology*, 13. <https://doi.org/10.3389/fpsyg.2022.956798>
- Opitz, B., Ferdinand, N.K., and Mecklinger, A., 2011. Timing matters: the impact of immediate and delayed feedback on artificial language learning. *Front Hum Neurosci*. doi:10.3389/fnhum.2011.00008
- Paris, B., 2022. Instructors' perspectives of challenges and barriers to providing effective feedback. *Teaching & Learning Inquiry: The ISSOTL Journal*, 10. <https://doi.org/10.20343/teachlearningqu.10.3>
- Pisan, Y., Sloane, A. M., Dale, R., and Richards, D., 2002. Providing timely feedback to large classes. *IEEE Xplore, International Conference on Computers in Education*. <https://doi.org/10.1109/CIE.2002.1185961>
- Randell, B., and Coghlan, B., 2023. ChatGPT's astonishing fabrications about Percy Ludgate. *IEEE Annals of the History of Computing*, 45 (2), pp 71–72. <https://doi.org/10.1109/mahc.2023.327298>
- Regional Convention on the Recognition of Studies, Diplomas and Degrees in Higher Education in Latin America and the Caribbean. 2019. UNESCO. [online] Available at < <https://www.unesco.org/en/legal-affairs/regional-convention-recognition-studies-diplomas-and-degrees-higher-education-latin-america-and-0> > [Accessed 11 July 2023].
- Tiili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., and Agyemang, B., 2023. What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10 (1), p 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Vijayakumar, B., Höhn, S., and Schommer, C., 2019. Quizbot: Exploring Formative Feedback with Conversational Interfaces. *Communications in computer and information science*, pp 102–120. https://doi.org/10.1007/978-3-030-25264-9_8
- Wang, T. et al. (2023) 'Exploring the potential impact of artificial intelligence (AI) on international students in higher education: generative AI, chatbots, analytics, and international student success,' *Applied Sciences*, 13(11), p. 6716. <https://doi.org/10.3390/app13116716>

ChatGPT in Education – Understanding the Bahraini Academics Perspective

Amal Alrayes¹, Tara Fryad Henari¹ and Dalal Abdulkarim Ahmed²

¹College of Information Technology, University of Bahrain, Bahrain

²English Language Center, University of Bahrain, Bahrain

aalrayes@uob.edu.bh

thenari@uob.edu.bh

dmukhallaf@uob.edu.bh

<https://doi.org/10.34190/ejel.22.2.3250>

An open access article under [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License](#)

Abstract: This paper provides a thorough examination of the role of Artificial Intelligence (AI), particularly ChatGPT and other AI language models, in the realm of education. Drawing insights from existing literature and a novel study on educator perspectives, the paper delves into the potential advantages, ethical dilemmas, and factors shaping educators' attitudes towards AI integration in education. AI language models have the potential to revolutionize educational content creation, personalize learning experiences, and streamline assessment and feedback processes. These capabilities hold the potential to enhance teaching and learning outcomes while catering to the diverse needs of students. However, ethical concerns loom large in the adoption of AI in education. Bias in generated content is a chief concern, as it can perpetuate societal biases and lead to unfair treatment or the dissemination of inaccurate information. The solution lies in rigorous data curation to ensure equitable educational experiences for all students. Moreover, the potential for generating inappropriate or misleading content poses a significant ethical challenge, impacting students' well-being, civic understanding, and social interactions. Safeguards must be implemented to detect and rectify biased or inappropriate content, fostering inclusive and unbiased learning environments. Transparency emerges as a crucial ethical consideration. The opacity of AI models like ChatGPT makes it difficult to comprehend their decision-making processes. Enhancing model interpretability and explainability is vital for accountability and addressing embedded ethical issues. Privacy concerns related to data collection and usage are emphasized in the literature. Clear policies and guidelines must govern data collection, use, and protection, ensuring data is solely employed for educational purposes and maintaining robust data security measures. Our study expands upon these insights by exploring socio-demographic factors, motivations, and social influences affecting educators' AI adoption in higher education. These findings inform institutions on tailoring AI integration strategies, emphasizing responsible usage through training, and assessing the impact on learning outcomes. As educational institutions increasingly embrace AI, including advanced models like GPT-4, a cautious and thoughtful approach is vital. Balancing potential benefits with ethical challenges ensures that AI enhances teaching and learning while upholding fairness, equity, and accountability. In summary, this paper illuminates the potential of AI in education, accentuates ethical concerns, and highlights the significance of understanding educators' perspectives. Collaboration between educators and policymakers is essential to navigate the complexities of AI integration, ensuring that education remains a realm of equitable, efficient, and accountable learning experiences.

Keywords: Artificial intelligence, AI, e-Learning, Open AI, ChatGPT

1. Introduction

The debate surrounding whether machines can "think" and "understand" dates back to the 1950s when John McCarthy coined the term "Artificial Intelligence" and proposed the "Turning Test" (McCarthy, 1955). The question of whether machines can truly emulate human thinking remains a topic of discussion among scientists and philosophers. Various philosophical schools, such as Consciousness, Mathematical, and Theological, have opposed McCarthy's idea that machines can imitate human thinking (Floridi, 2019). After a period of significant development and interest in AI, the field experienced what is referred to as an "AI Winter," resulting from a mismatch between expectations and actual progress. This led to a lack of investment in AI research, resulting in limited application of AI technology. However, with recent breakthroughs in deep learning, chatbots, and other AI technologies, there is renewed interest and investment in AI (Russell et al., 2021).

AI can be viewed as a tool that processes and evaluates large amounts of data to make forecasts, spot trends, and automate certain tasks. Machine learning, which uses methods to learn from data and improve over time, is frequently used in this process. AI has been applied in various fields, including healthcare, finance, retail, transportation, and education (Topol, 2019).

In 2015, a group of entrepreneurs, including Elon Musk, Sam Altman, Greg Brockman, Ilya Sutskever, John Schulman, and Wojciech Zaremba, created the OpenAI research organization. The organization's goal is to

advance AI development in a way that benefits mankind as a whole, while also ensuring that the technology is safe and poses no existential threat to humans (OpenAI, 2022). The main focus of the organization is to build powerful AI systems capable of performing a wide range of activities. The potential uses of AI in education are substantial, even though OpenAI's developments in AI technology are not primarily targeted at designing an educational environment. The advancements made by OpenAI in fields like chatbots, machine learning, and natural language processing may improve educational experiences. The ongoing rapid advancement of AI necessitates research to keep pace with its development. It is essential for understanding the efficacy, impact, and pedagogical implications of AI in education. Research sheds important light on AI's ethical implications including privacy, openness, justice, and responsibility. It assists us in comprehending the social effects of AI and in creating policies and rules to ensure its responsible usage.

This research aims to better understand educators' perspectives toward the deployment of Chat GPT in the field of education, focusing on what may influence them to use or not use ChatGPT in their classrooms. It tries to understand the source of educators' opinions, worries, and hopes surrounding the integration of Chat GPT in educational contexts by concentrating on their points of view. For AI technology to be adopted and used in education successfully, it is essential to examine educators' mindsets (Iqbal, Ahmed, & Azhar, 2023). The potential advantages and difficulties of using Chat GPT in education will become clearer once we have a better understanding of how educators view and approach its use. Along with moderators like age, gender, and experience, factors including performance expectancy, effort expectancy, social influence, and hedonic motivation are investigated. Following the UTAUT modal, the study investigates the connections between these variables, moderators, and the intention to use ChatGPT. The research also takes into account the amount of education and the specialization of the educator.

The findings of the study will provide valuable insights to educational policies, professional development initiatives, and strategies for ethically integrating AI technologies into classroom practices. The goal of this study is to advance knowledge of AI integration in educational environments by identifying the elements that influence educators' attitudes regarding the implementation of Chat GPT.

2. Literature Review

2.1 Open AI

OpenAI's breakthrough achievements include the creation of generative pre-trained transformer (GPT), which is a family of transformer-based neural language models trained using unsupervised learning on large amounts of text data. GPT has achieved cutting-edge performance on various natural language processing tasks, including text completion, question answering, and conversation response generation. It is a computer system designed to generate strings of words, codes, or other data from source input called prompts like that used to statistically predict word sequences in machine translation (Floridi & Chiriatti, 2020). The large language model uses self-attention mechanisms to capture contextual information and generate natural language text. OpenAI has released three versions of GPT, each improving on the previous one, introducing a natural language processor that allows machines to produce coherent and grammatically correct sentences, paragraphs, and even longer text passages (Brown & Venkatesh, 2020).

Version	Release Date	Parameters	Dataset (Text)	Ability
GPT-1	2018	117 million	40GB	Can generate coherent paragraphs of text but has limited ability to perform complex natural language processing tasks.
GPT-2	2019	1.5 billion	40GB	Can generate coherent paragraphs of text and perform a range of natural language processing tasks, such as translation and summarization.
GPT-3	2020	175 billion	500GB	Can generate highly coherent and fluent text in a range of styles and domains and perform a vast array of natural language processing tasks, such as question-answering, dialogue generation, and even code generation.
GPT-4	2022	1 trillion	1 terabyte	Can produce original text or graphics or solve textual issues. It has the ability for problem-solving, producing articles and programming are among the many things it is capable of. With great accuracy in 26 different languages, the GPT-4 is bilingual and includes multiple-choice questions. Along with text, it can also process photos.

2.2 ChatGPT in Education

Education, being a primary target of AI technology, has been a focus of many researchers on how a responsible and ethical application of the GPT could benefit the field. OpenAI Organization aims to advance the integration of artificial intelligence (AI) and education, from transforming how students learn and how teachers teach to developing intelligent tutoring systems that provide personalized feedback to students. The organization further highlights how new language models such as GPT can be used to generate educational content and assist with grading and assessment, reducing the workload on teachers and providing more consistent and objective grading. In their overview of "Intelligent Education," Wang et al. (2020) introduced the concept of integrating AI and education with the aim of improving teaching and learning outcomes. They highlight how AI can potentially enhance various aspects of education, such as curriculum development, teaching strategies, and student assessment. AI can provide personalized feedback and support to students, analyze large amounts of data to identify patterns and trends, and provide immediate feedback to identify areas where students may be struggling. However, they also acknowledge the challenges that come with "intelligent education," such as data standards and teacher training.

Bansal and Chugh (2021) agree with the improvements AI could bring to education, discussing the potential use of GPT in various areas of education, including language learning, content creation, student assessment, and personalized learning. The advantages of GPT in education, such as its capacity to help with personalized learning, were emphasized by many researchers, including Srinivasan and Padma (2021) and Kumar and Bhattacharya (2021). They state that GPT-3 could assist with language translation, writing assistance, and personalized learning, but they also acknowledge the limitations and ethical concerns surrounding its use in education, such as bias in language models and the potential for replacing human teachers.

Ramasubramanian's (2021) supports the use of GPT for personalized learning by proposing a framework. He believes that the framework will allow teachers to create unique material and evaluations depending on the interests, skills, and learning preferences of students. Consequently, students will more likely become more invested in the subject matter and improve their learning outcomes. The framework introduced by Ramasubramanian targets three main areas:

- Content generation and curation: The language model can be used to generate and curate educational content such as course materials and textbooks.
- Personalization of learning experiences: Chat GPT-3 has the potential to provide personalized recommendations and support to students based on their individual learning needs and preferences.
- Assessment and feedback: The language model could be used to automatically grade written assignments and provide personalized feedback to students.

After the release of Chat GPT in 2022, Atsushi Mizumoto and Masaki Eguchi examined the feasibility of using AI language models for automated essay scoring (AES). They suggest that AI language models can be effective for AES when trained on large and diverse datasets but should not be seen as a replacement for human graders. While the model used by Mizumoto held a correlation coefficient of 0.91 in comparison to human graders' scores, Bashir and Alshahrani (2021) obtained a coefficient of 0.98. However, they also identified a wide range of errors, including grammatical errors, spelling errors, and semantic errors. Like other scholars, Bashir and Alshahrani highlighted potential drawbacks in the use of GPT in assessment.

Researchers have identified several ethical concerns surrounding the use of GPT models in education. One of these concerns is the issue of bias. Since GPT models are trained on large text datasets, these datasets may contain biases related to factors such as ethnicity and gender. Therefore, it is essential to carefully select training data and assess the models for potential biases before implementing them in educational contexts (Adesope, 2020).

Another ethical issue related to the use of GPT in education is the possibility of generating inappropriate content. GPT models have been found to associate feminine characters with family and appearance, depicting them as less powerful than male characters, even when associated with high power verbs in a prompt. This reflects societal biases and preconceptions in the language generated by the model, potentially amplifying negative emotional, civic, and social tendencies (Li & Bamman, 2021; Bowdoin Science Journal, 2021). Transparency is another ethical concern raised by researchers. Education is an open and accessible process, with all information made available to teachers, students, and other stakeholders. However, with GPT models, this may not be the case. These models are becoming increasingly complex and difficult to understand, raising questions about their ability to identify and address ethical problems and prejudices. When GPT is used to create assessments, grade

assignments, or provide personalized learning experiences, it may not be clear to the user (whether teacher or student) how the model arrived at its predictions or outputs (Buruk & Oğuz 'Oz, 2023).

Privacy is one of the key issues with using AI in education. The ability of AI technology to gather and analyze vast volumes of student data is expanding as it develops and becomes more sophisticated. Sensitive information such as student personal data, academic achievement, and behavioral trends may be included in this data.

Researchers have concerns that this information may be utilized in ways that are not beneficial for both students, teachers and educational institutes. Large datasets that frequently include private information about individuals, such as personal information and secret school records, are used to train the models. "The use of these models can potentially compromise the privacy of students and educators if not properly regulated." (Adesope, 2020, p. 138). The information may be used, for instance, to target students with individualized advertisements or to decide what they should study based on criteria other than academic quality. These sheds light again upon the importance of human control on AI in order to withhold accountability and openness in the application of GPT or any form of AI in education. There is a need for clear and explicit policies for data collection, usage, and protection as well as making sure that these guidelines are being followed. Advocators like the Electronic Frontier Foundation, the National PTA, and the Consortium for School Networking are stressing the importance of giving students and their families more control over their personal data, including the opportunity to access, review, and remove information as they see fit.

As the world is getting ready to embrace GPT-4, educational institutes must have a clear stand on both the benefits and limitations of the use of GPT in the classroom. Unlike the former GPT models, GPT4 is "a supervised learning model", meaning that labeled data was used to train it. This enables GPT-4 to produce more precise and realistic findings including images and texts. With all this into perspective, universities have already started laying, yet not rolling out the red carpet towards the implementation of GPT in Education. Universities like the University of Bristol and Science Po have launched a review into the use of AI language models, including ethical considerations around their potential impact on bias and discrimination. The former university considered it as "cheating", the latter required students to be transparent on the use of Chat GPT and include it in the reference. Similarly, more than 6,000 teachers from Harvard University, Yale University, the University of Rhode Island, and other institutions have signed up to use GPTZero, a program that claims to quickly identify plagiarism in student work by comparing it with a database of previously submitted assignments (New York Times, 2023). The University of Honk Kong has allowed only its staff to use Chat GPT in Teaching and Learning (T&L) for a provisional era concurrently with a lot of restrictions on the use. Whereas the University of Sydney has allowed students to use chat GPT in T&L by allowing students to compose essays in medical school "There are different types of knowing – one of the basic types is memorizing and reproducing information or collating information – that is a stock-in-trade for universities, but ChatGPT does that," (Harris,2023). They believed that chat GPT could produce an essay as a first step and then request students to use higher level skills like that of critical thinking and judgment to evaluate and edit them.

There is also very limited research that has been done to tackle the issue of using chat GPT in education. Educational institutes have been caught off guard and there is not much time left to discuss the role AI may play in the classroom to design well-grounded scientific responses. ChatGPT is here and free to use for now.

3. Research Objectives

The primary objectives of this study are to:

- Investigate how socio-demographic factors influence the adoption of ChatGPT among academics, focusing on the impact and value of Gen-AI in learning and teaching.
- Identify the motivations that drive educators to use ChatGPT and assess the perceived impact of Gen-AI on teaching practices.
- Evaluate the social influences affecting academics' use of Gen-AI, emphasizing the broader implications and value of Gen-AI in educational contexts.

Therefore, the research questions will be as follow:

- How do socio-demographic factors influence the adoption of ChatGPT in learning and teaching?
- What motivates educators to use ChatGPT and how does it impact teaching practices?
- How do social influences affect academics in using GPT, and what is their collective impact on integrating Generative Artificial Intelligence in learning environments?

4. Research Methodology

The research utilizes the UTAUT (Unified Theory of Acceptance and Use of Technology) model developed by Venkatesh et al. (2003). The model was developed through an extensive review of existing models, including the TAM (Technology Acceptance Model), the TRA (Theory of Reasoned Action), the TPB (Theory of Planned Behavior) and DOI (Diffusions of Innovations Theory). These models aim to understand and explain individuals' attitudes, beliefs and behaviors related to technology adoption and use. However, the TAM focuses more on individuals' perceptions, the TRA and TPB focus on social influences and subjective norms and the DOI focuses on the attributes of the innovation itself and how it effects the social system (Davis, 1989; Fishbein & Ajzen, 1975; Azjen 1991; Rogers, 1962). In addition to these models, the UTAUT also takes into consideration the demographics factors of age and gender and there are two other moderating factors the model considers: experience and voluntariness. The former refers to the experience an individual has had with a similar technology, and the latter refers to the extent the individual believes they have a choice in using the technology (Venkatesh et al., 2003). Moreover, the UTAUT model has four key factors that influence technology use and acceptance: performance expectancy (PE), effort expectancy (EE), social influence (SI), and facilitating conditions (FC).

The UTAUT model is considered to be a comprehensive framework and although the model has been validated and used extensively in research, it is important to note that when initially developed it was intended for the acceptance of technology for employees in the workplace. Venkatesh, Thong & Xu (2012) expanded this model to take into consideration the consumer use, and developed the UTAUT2, an enhanced model of the Unified Theory of Acceptance and Use of Technology. Three extra factors were incorporated: hedonic motivation (HM), price value (PV), and Habit (H) and the moderator voluntariness was removed, as consumers have no organizational mandate and is considered to be voluntary when using the technology (Tamilmani, 2021).

This research utilizes the UTAUT2 model from the perspective of educators in higher education in using ChatGPT. Performance expectancy, effort expectancy, social influence and hedonic motivation are taken into consideration, including three moderators: age, gender, and experience. Relationships with the factors and moderators and intention to use will be the same as the UTAUT2 model. Experience will be considered from the perspective of the educator's specialization and level of education. Figure 1 summarizes the relations and the hypothesis and the research model.

H1: Performance Expectancy (PE) will positively predict the intention to use ChatGPT in higher education among educators, as they perceive the technology as improving their teaching experience.

H2: Effort Expectancy (EE) will positively predict the intention to use ChatGPT in higher education among educators, as they perceive the technology as easy to use and requiring minimal effort and resources.

PE refers to the extent to which individuals believe that using technology will improve performance, with focus on extrinsic motivation. Whilst EE refers to the degree to which users perceive that using the technology will be free of effort. Venkatesh et al. (2003) based these assumptions on the Expectancy-Value Theory of Motivation and the Technology Acceptance Model which suggests that an individual's motivation to engage in a behavior is influenced by their belief that the behavior will lead to certain positive outcomes and individuals are motivated to engage in a behavior when they expect it to be easy related to the value of the outcome. H1 explores the use of ChatGPT by educators during their teaching experience. Different aspects of teaching is explored, with main focus on whether there will be a positive relation on PE and lesson planning and assessment. H2 aims to explore the impact of the ease of the use of ChatGPT in higher education by educators, exploring the concept that the least effort required will have a positive impact on the use of ChatGPT in higher education.

H3: Social influence (SI) will positively predict the intention to use ChatGPT in higher education among educators, as they perceive their colleagues and management as supportive of the technology and its implementation.

SI refers to the degree to which an individual perceives that people who are important to them think they should use the technology. Venkatesh et al. (2003) based this factor on the subjective norms of the social factors of the Theory of Social Influence and the Technology Acceptance Model, where the behavior of individuals may change based on the perception of others about them. H3 refers to the use of ChatGPT by educators, and how their colleagues view the use of this technology, and those that use it, the more supportive they are whether it be verbally or written the more likely the individual forms a positive attitude towards using it. In education the concept behind using ChatGPT may have the question of whether the use is considered ethical or unethical in education. Some regions, like Europe have identified this and have a growing interest in developing teacher-

oriented guidelines and regulations for the ethical use of ChatGPT in education. However, as identified by Holmes & Tuomi (2021) there is currently no regulations or official guidelines proposed internationally.

H4: Hedonic motivation (HM) will positively predict the intention to use ChatGPT in higher education among educators, as they are motivated by enjoyment and satisfaction.

HM refers to the degree to which using a technology is considered to be fun or generates pleasure. This construct focuses on intrinsic motivation and was derived from the Motivation Theory (Venkatesh, Thong & Xu, 2012). HM has been identified in previous research to be one of the main predictors in consumer behavior, as well as IS research in the consumer technology use (Venkatesh, Thong & Xu, 2012; Brown & Venkatesh, 2005). H4 refers to the use of ChatGPT by educators, and how the novelty and excitement of using this new technology will positively affect the technology adaption and use.

H5: Age and gender will moderate the relationship between performance expectancy (PE) and an educator's intention to use ChatGPT in higher education.

H5 suggests that age and gender will moderate the relationship with PE and the educators' intention to use ChatGPT in education. Previous research has indicated that age can affect an individual's perception of the usefulness of technology, with younger educators having a stronger belief that performance will be improved when an educator utilizes technology (Teo, 2009; Wang, Wu, & Wang, 2011). Research has also indicated that gender plays a role in the relationship of PE and intention to use (Liu et al., 2010), although some scholars have indicated there is no significant differences (Wang, Wu, & Wang, 2011). This discrepancy may be based on the technology researched. This hypothesis determines there will be a significant difference based on age and gender when using ChatGPT in higher education.

H6: Age, gender and experience will moderate the relationship between effort expectancy (EE) and an educator's intention to use ChatGPT in higher education.

As ChatGPT is relatively new, research is limited as to the relation regarding this specific technology, however for example Wang, Wu, & Wang (2011) found that the relation between perceived ease of use and the intention of use of technology was stronger with younger females for the use of mobile learning. Regarding experience, this research refers to experience as the educator's specialty, H6 assumes that experience, i.e. educators' specialty will have a relation on the effort expectancy and ChatGPT in higher education.

H7: Age, gender and experience will moderate the relationship between perform Social Influence (SI) and an educator's intention to use ChatGPT in higher education.

H7 suggests that age, gender and experience will moderate the relationship with SI and the educator's intention to use ChatGPT in higher education. Previous research has indicated that younger people are more likely to be influenced by social factors when using new technology, (Venkatesh et al., 2003). Research regarding different technologies have inconsistencies in the role of gender and social influences, but H7 indicates there is a relationship between both. H7 also indicates that regarding the specialty of the educator, there is a relationship with social influence.

H8: Age, gender and experience will moderate the relationship between hedonic motivation (HM) and an educator's intention to use ChatGPT in higher education.

Studies have found that younger individuals are more likely to be motivated by the enjoyment of using technology, indicating a positive relationship between age and hedonic motivation (Venkatesh et al., 2003). Regarding the role of gender, research is inconclusive with some indicating male, and others indicating females being more hedonically motivated with the intention of using technology. However, H8 indicates there is a relationship between gender and intention to use, it also relates experience/ specialty as a moderator.

H9: Behavior Intention will positively predict the use behavior of ChatGPT in higher education among educators.

According to the UTAUT model (Venkatesh et al., 2003), there is a highly positive link between expressing a strong intention to utilize technology and actually using it. H9 states that ChatGPT will be utilized if educators declare their intention to use it in higher education.

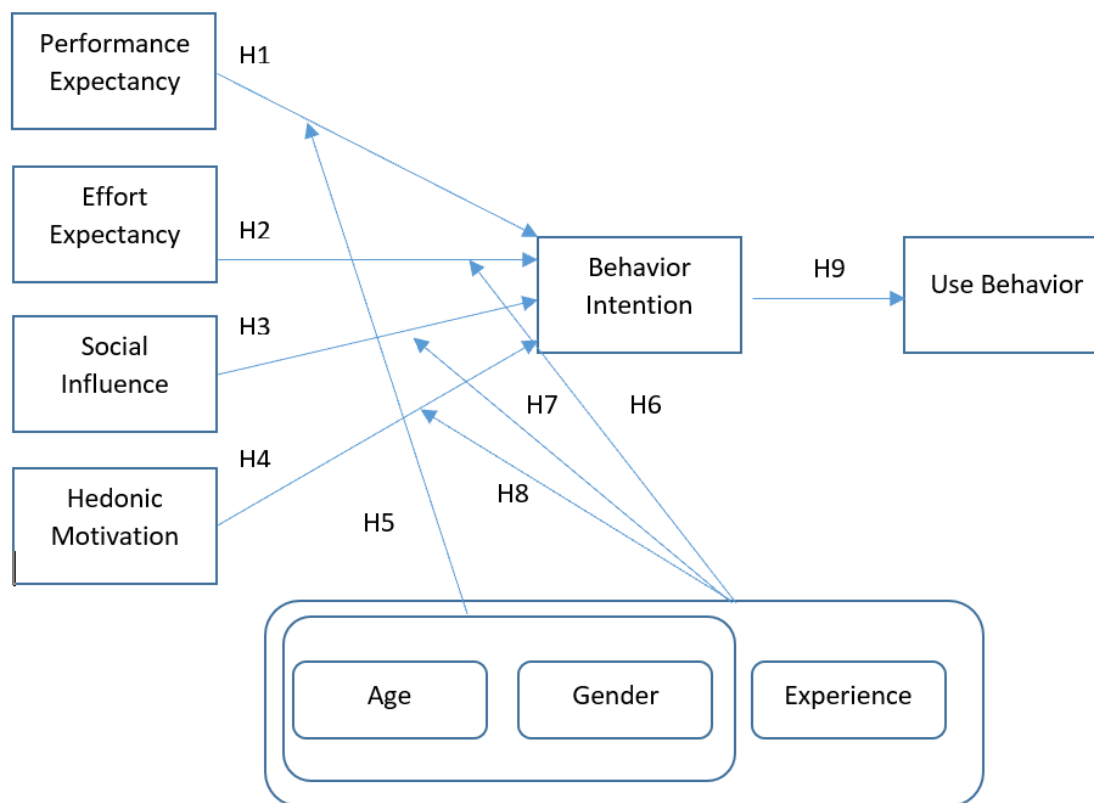


Figure 1: Research Model

In this research study, an online survey serves as the primary data collection method. The survey is designed to gather comprehensive insights into the participants' perspectives on the adoption and impact of Generative Artificial Intelligence, specifically ChatGPT, in the context of learning and teaching. To ensure the ethical collection of data, participants are approached transparently and provided with clear information about the purpose of the study, the nature of the questions, and how their responses will be utilized. Prior to participating, individuals are presented with a consent form outlining the voluntary nature of their involvement, emphasizing their right to withdraw at any point without consequences. The consent form also highlights the anonymity and confidentiality of their responses, assuring participants that their data will be aggregated and reported in a manner that protects their privacy.

5. Results

5.1 Demographic Information

This study collected data of 141 responses using an online survey. The demographic data of the participants are analyzed using frequency analysis and the results are reported in Table 1. According to the results, 39.72% of the participants were male and 60.28% were female. The age of the participants is also categorized into five categories reflecting the generations, such as Gen Z (20 to 26 years), Millennials (27 to 42 years), Gen X (43 to 58 years), Baby boomers (59 to 77 years), and the Silent generation whom above 77 years. Among the total respondents of the study, 3.55% were from Gen Z age group, 43.97% were Millennials, 46.81% were from Gen X age group, 5.67% were Baby boomers and no respondents were from the Silent generation. The results show that the majority of the respondents' age ranges from 27 to 58 years (Millennials and Gen X). In addition, the participants were asked to record their academic degrees. The results show that 12.77% of the participants were Bachelor's degree holders, 45% were Master's degree holders, and 41% were PhD degree holders. Table 1 also shows the results for the field of the respondents. Among the total participants, 6.38% were from applied studies, 5.67% were from social sciences, 36.88% were from the teaching field, 7.09 were from business, 6.38% were from engineering, 1.42% were from health and sports science, 24.82% were from information technology, 6.38% were from the law, 2.84% were from science, and 2.13% were from other fields.

Table 1: Demographic Information of Participants

	Frequency	Percent	Valid Percent	Cumulative Percent
Gender				
Male	56	39.72	39.72	39.72
Female	85	60.28	60.28	100.00
Total	141	100.00	100.00	
Age				
20-26 (Gen Z)	5	3.55	3.55	3.55
27-42 (Millennials)	62	43.97	43.97	47.52
43-58 (Gen X)	66	46.81	46.81	94.33
59-77 (Baby Boomers)	8	5.67	5.67	100.00
Total	141	100.00	100.00	
Academic Degree				
Bachelor	18	12.77	12.77	12.77
Master	64	45.39	45.39	58.16
PhD	59	41.84	41.84	100.00
Total	141	100.00	100.00	
Field				
Applied Studies	9	6.38	6.38	6.38
Social Science	8	5.67	5.67	12.06
Teaching	52	36.88	36.88	48.94
Business Administration	10	7.09	7.09	56.03
Engineering	9	6.38	6.38	62.41
Health & Sport Science	2	1.42	1.42	63.83
Information Technology	35	24.82	24.82	88.65
Law	9	6.38	6.38	95.04
Science	4	2.84	2.84	97.87
Other	3	2.13	2.13	100.00
Total	141	100.00	100.00	

The participants of the study were asked if they have already used/tried ChatGPT (Yes/No) and reasons of usage. In case of not used/tried ChatGPT then a following questions regarding the reasons for that. Among the total participants, 46.81% ($n = 66$) recorded their response as "Yes", while 53% ($n = 75$) responded as "No" (see Table 2). Table 3 shows the reasons for using ChatGPT which are recorded by the participants. According to Table 3, 37.9% ($n = 25$) of the participants used it for research, 19.7% ($n = 13$) used it for teaching and learning, 7.6% ($n = 5$) used it for assessment (creation/checking), 6.1% ($n = 4$) using it for communication, 19.7% ($n = 13$) using it for report preparation, 7.6% ($n = 5$) using it for other reasons such as curiosity, fun, or testing and experiencing the abilities of it.

Table 4 reported the reasons for not using ChatGPT. According to the results, among the total participants ($n = 75$), 44% ($n = 33$) of the participants were not interested in it, 8% ($n = 6$) found it difficult, 6.7% ($n = 5$) were responded that it is unethical, 4% ($n = 3$) not using it because of privacy, 2.7% ($n = 2$) not using it due to

inaccuracy, 34% ($n = 26$) not using it due to other reasons such as availability, accessibility, knowledge, familiarity and intention to use soon.

Table 2: Have you used/tried ChatGPT?

	Frequency	Percent	Valid Percent	Cumulative Percent
Yes	66	46.81	46.81	46.81
No	75	53.19	53.19	100.00
Total	141	100.00	100.00	

Table 3: Distribution of participants based on the use of ChatGPT

What do you use it for?				
	Frequency	Percent	Valid Percent	Cumulative Percent
Research	Frequency	Percent	Valid Percent	Cumulative Percent
Teaching and Learning	13	19.7	20.0	58.5
Assessment (creation/checking)	5	7.6	7.7	66.2
Communication	4	6.1	6.2	72.3
Report Preparation	13	19.7	20.0	92.3
Other	5	7.6	7.7	100.0
Total	65	98.5	100.0	
System	1	1.5		
Total	66	100.00		

Table 4: Distribution of data based on Reasons for not using ChatGPT

What are the reasons of not using it yet?				
	Frequency	Percent	Valid Percent	Cumulative Percent
Not interested	33	44.0	44.0	44.0
Difficult	6	8.0	8.0	52.0
Unethical	5	6.7	6.7	58.7
Privacy	3	4.0	4.0	62.7
Inaccuracy	2	2.7	2.7	65.3
Other	26	34.7	34.7	100.0
Total	75	100.0	100.0	

5.2 T-Test

T-test is used to determine the difference between the behavior intentions and use behavior regarding ChatGPT in males and females. Table 5 reported the descriptive statistics. According to the results, the mean for behavior intention of males is 3.537 with a standard deviation of 0.713, and for females is 3.679 with a standard deviation of 0.799. In addition, the mean for use behavior of males is 3.556 with a standard deviation of 0.832, and for females is 3.880 with a standard deviation of 0.789. According to the results, the means for behavior intentions and use behavior for females are greater than for males. T-test is used to check whether the difference is statistically significant or not.

Table 6 reported the results of the independent t-test. The mean difference between males' and females' behavior intentions about ChatGPT is -0.142 which is statistically insignificant at a 5% significance level because the p-value is 0.460 which is greater than 0.05. Additionally, the mean difference between males and females use behavior of ChatGPT is -0.325 which is also statistically insignificant at a 5% significance level because p-value $0.113 > 0.05$. Further, equal variances are assumed based on the results of Levene's F test, as for behavior intention $F = 1.42$ and $p = 0.237$ which is greater than a 5% significance level; thus, equal variances are assumed, and for use behavior $F = 0.165$ and p-value = 0.686 which is also greater than 0.05; hence, equal variances are assumed.

Table 5: Descriptive Statistics of Gender

Age		N	Mean	Std. Deviation
Behavior intention	Male	27	3.537	0.713
	Female	39	3.679	0.799
Use behavior	Male	27	3.556	0.832
	Female	39	3.880	0.789

Table 6: T-test Results

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
									Lower	Upper
Behavior Intention	Equal variances assumed	1.427	0.237	-0.744	64	0.460	-0.142	0.192	-0.525	0.240
	Equal variances not assumed					-0.759	59.884	0.451	-0.142	0.188
Use behavior	Equal variances assumed	0.165	0.686	-1.608	64	0.113	-0.325	0.202	-0.728	0.079
	Equal variances not assumed					-1.592	54.108	0.117	-0.325	0.204

5.3 Analysis of Variance (ANOVA)

One-way ANOVA is used to observe the difference between different age groups, academic degrees, and fields. The results are reported and discussed in subsections.

5.3.1 Age

Table 7 reported the descriptive statistics of behavior intentions and use behavior of ChatGPT between different age groups. The mean for Gen Z is 4.063 with a standard deviation of 0.657, Millennials is 3.683 with a standard

deviation of 0.859, Gen X is 3.530 with a standard deviation of 0.7323, Baby boomers is 3.5 with a standard deviation of 0.250. In addition, the mean for the use behavior of ChatGPT in Gen Z is 4 with a standard deviation of 0.816, Millennials is 3.833 with a standard deviation of 0.801, Gen X is 3.626 with a standard deviation of 0.865, and Baby boomers is 4 with a standard deviation of 0.333. ANOVA is used to test whether these differences are statistically significant or not. The results of ANOVA are presented in Table 8. According to the results, the mean difference between behavior intention is insignificant as $F = 0.674$ because the p-value is 0.571 which is greater than 0.05. This means that there is no difference between behavior intentions regarding the use of ChatGPT in higher education in different age groups. Additionally, the difference between use behaviour of different age groups is insignificant ($F = 0.550$, $p = 0.650 > 0.05$).

Table 7: Descriptive Statistics of Age

		N	Mean	Std. Dev.	Mini	Max
Behavior intention	20-26	4	4.063	0.657	3.50	5.00
	27-42	26	3.683	0.859	2.50	5.00
	43-58	33	3.530	0.723	1.75	5.00
	59-77	3	3.500	0.250	3.25	3.75
	Total	66	3.621	0.762	1.75	5.00
Use behavior	20-26	4	4.000	0.816	3.00	5.00
	27-42	26	3.833	0.801	2.33	5.00
	43-58	33	3.626	0.865	2.00	5.00
	59-77	3	4.000	0.333	3.67	4.33
	Total	66	3.747	0.817	2.00	5.00

Table 8: ANOVA (Age)

		Sum of Squares	df	Mean Square	F	Sig.
Behavior intention	Between Groups	1.194	3	.398	.674	.571
	Within Groups	36.586	62	.590		
	Total	37.780	65			
Use behavior	Between Groups	1.123	3	.374	.550	.650
	Within Groups	42.224	62	.681		
	Total	43.347	65			

5.3.2 Academic degree

Table 9 reported the descriptive statistics of behavior intentions and use behavior of ChatGPT in higher education in participants with different academic degrees such as bachelor's, master's, and Ph.D. The mean for

Bachelor's degree holders is 4.125 with a standard deviation of 0.647, master's degree holders are 3.781 with a standard deviation of 0.805, and Ph.D. degree holders are 3.431 with a standard deviation of 0.704. In addition, the mean for the use behavior of ChatGPT in bachelor's degree holders is 4.222 with a standard deviation of 0.750, master's degree holders are 3.917 with a standard deviation of 0.800, and Ph.D. degree holders is 3.556 with a standard deviation of 0.801. ANOVA is used to test whether these differences are statistically significant or not. The results of ANOVA are reported in Table 10. According to the results, the mean difference between behavior intention is statistically significant as $F = 3.162$ and the p-value is 0.049 which is less than 0.05. This means that there is a significant difference between behavior intentions of ChatGPT in higher education based on academic degrees. The difference between use behavior of ChatGPT based on academic degrees is insignificant ($F = 2.651$, $p = 0.078 > 0.05$) at a 5% significance level while significant at a 10% significance level.

Table 9: Descriptive Statistics (Academic Degree)

		N	Mean	Std. Dev.	Mini	Max
Behavior intention	Bachelor	6	4.125	0.647	3.25	5.00
	Master	24	3.781	0.805	2.00	5.00
	PhD	36	3.431	0.704	1.75	5.00
	Total	66	3.621	0.762	1.75	5.00
Use behavior	Bachelor	6	4.222	0.750	3.00	5.00
	Master	24	3.917	0.800	2.00	5.00
	PhD	36	3.556	0.801	2.00	5.00
	Total	66	3.747	0.817	2.00	5.00

Table 10: ANOVA (Academic Degree)

		Sum of Squares	df	Mean Square	F	Sig.
Behavior intention	Between Groups	3.446	2	1.723	3.162	.049
	Within Groups	34.334	63	.545		
	Total	37.780	65			
Use behavior	Between Groups	3.365	2	1.683	2.651	.078
	Within Groups	39.981	63	.635		
	Total	43.347	65			

5.3.3 Field

The difference between the behavior intentions and use behavior of individuals based on their field is examined using the ANOVA test. The results of descriptive statistics are presented in Table 11. The results demonstrate that the mean of behavior intentions for Applied studies is 3.69 with a standard deviation of 0.63, Social sciences are 3.83 with a standard deviation of 0.38, Teaching is 3.58 with a standard deviation of 0.65, Business administration is 4.09 with a standard deviation of 0.83, Engineering is 3.50 with a standard deviation of 0.66, Health and Sports science is 3.50, Information Technology is 3.61 with a standard deviation of 0.88, the Law is 2.50 with a standard deviation of 1.06, and other is 3. The results demonstrate that the mean of use behavior for Applied studies is 4.08 with a standard deviation of 0.83, Social sciences are 4.11 with a standard deviation of 0.19, Teaching is 3.72 with a standard deviation of 0.75, Business administration is 3.83 with a standard deviation of 0.94, Engineering is 3.56 with a standard deviation of 0.69, Health and Sports science is 4, Information Technology is 3.67 with a standard deviation of 0.99, the Law is 3.33 with a standard deviation of 0.47, and other is 4. The results of descriptive statistics illustrate that there are differences between the means of behavior intention and use behavior based on the field existing; thus, ANOVA is applied to test whether the

differences are statistically significant or not. The results of ANOVA are reported in Table 12 which demonstrates that there are no differences in behavior intention of individuals based on their fields as the p-value of F statistics (1.074) is 0.394 which is greater than the significance level 0.05. In addition, the results also confirm the insignificance of the difference between the use behavior of individuals based on their fields about the use of ChatGPT in higher education ($F = 0.280, p = 0.970 > 0.05$).

Table 11: Descriptive Statistics (Field)

		N	Mean	Std. Dev.	Mini	Max
Behavior intention	Applied Studies	4	3.69	0.63	2.75	4.00
	Social Science	3	3.83	0.38	3.50	4.25
	Teaching	23	3.58	0.65	2.00	4.75
	Business Administration	8	4.09	0.83	2.50	5.00
	Engineering	3	3.50	0.66	2.75	4.00
	Health & Sport Science	1	3.50		3.50	3.50
	Information Technology	21	3.61	0.88	2.50	5.00
	Law	2	2.50	1.06	1.75	3.25
	Other	1	3.00		3.00	3.00
	Total	66	3.62	0.76	1.75	5.00
Use Behavior	Applied Studies	4	4.08	0.83	3.00	5.00
	Social Science	3	4.11	0.19	4.00	4.33
	Teaching	23	3.72	0.75	2.33	5.00
	Business Administration	8	3.83	0.94	2.00	5.00
	Engineering	3	3.56	0.69	3.00	4.33
	Health & Sport Science	1	4.00		4.00	4.00
	Information Technology	21	3.67	0.99	2.00	5.00
	Law	2	3.33	0.47	3.00	3.67
	Other	1	4.00		4.00	4.00
	Total	66	3.75	0.82	2.00	5.00

Table 12: ANOVA (Field)

		Sum of Squares	df	Mean Square	F	Sig.
Behavior intention	Between Groups	4.949	8	.619	1.074	.394
	Within Groups	32.832	57	.576		
	Total	37.780	65			
Use behavior	Between Groups	1.637	8	.205	.280	.970
	Within Groups	41.710	57	.732		
	Total	43.347	65			

5.4 Partial Least Square Structural Equation Modelling (PLS-SEM)

5.4.1 Measurement model

Prior to conducting structural equation modeling (SEM), various instruments are utilized to evaluate the measurement model. The model is evaluated to ascertain its reliability and validity. The objective of this evaluation is to ascertain the constituent elements that should be incorporated within each construct, while also removing any extraneous components. Therefore, model assessment involves the utilization of factor loading, reliability tests, and validity tests. The utilization of Confirmatory Factor Analysis (CFA) is implemented to evaluate the measurement model in this study. Factor loadings are computed to assess the reliability and robustness of each item pertaining to a given construct. The factor loading is indicative of the extent to which the items utilized for assessing a given construct are indeed measuring that particular construct. According to Hair et al. (2016), the typical range for factor loading is between 0.40 and 0.60; however, a factor loading of 0.70 or higher is considered to indicate a high level of reliability and validity for the constructs. The factor loadings depicted in Figure 2 and Table 13 indicate that the items utilized for assessing the constructs are indicative of the same construct, as each item's value surpasses 0.70 and spans from 0.7 to 0.94. While four items, the first item of effort expectancy (EE1), the second and fifth items of social influence (SI2 and SI5), and the first item of hedonic motivation are removed because the factor loadings of these items were less than 0.70.

Table 13: Factor Loading

	Behavior Intention	Effort Expectancy	Hedonic motivation	Performance Expectancy	Social influence	Use Behavior
BI1	0.851					
BI2	0.864					
BI3	0.842					
BI4	0.765					
EE2		0.901				
EE3		0.861				
HM2			0.96			
HM3			0.936			
PE1				0.878		
PE2				0.725		
PE3				0.833		
SI1					0.795	
SI3					0.737	

	Behavior Intention	Effort Expectancy	Hedonic motivation	Performance Expectancy	Social influence	Use Behavior
SI4					0.78	
UB1						0.838
UB2						0.756
UB3						0.822

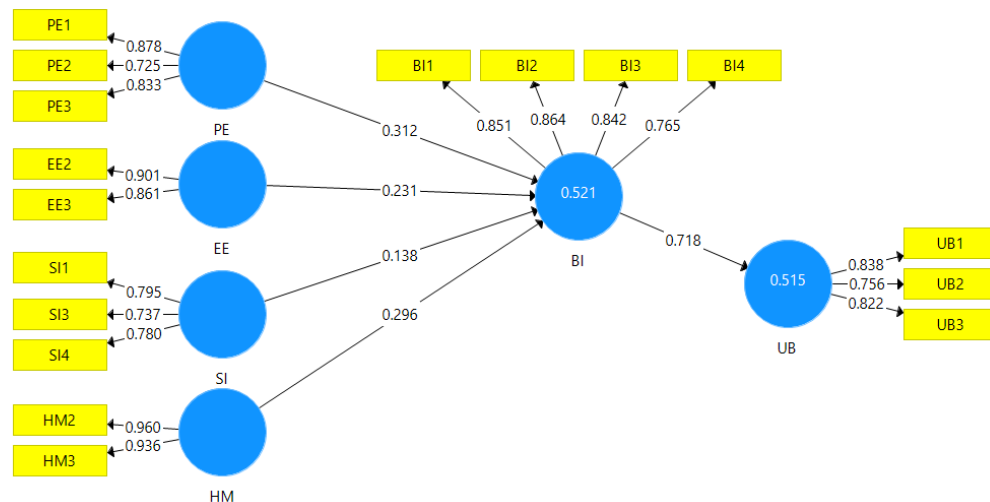


Figure 2: Factor Loadings

The assessment of the internal consistency reliability of a test pertains to the degree of consistency in the results obtained from all the constituent elements that constitute the assessment. The outcomes of the internal consistency reliability assessment demonstrate the manner in which each element of the test interacts with all other factors. According to Hair, Ringle, & Sarstedt (2011), the value of the reliability test needs to be greater than 0.60 in order to fall within the acceptable range. Reliability is tested using Cronbach alpha and composite reliability tests. The results are reported in Table 14, and according to the results, there is no issue of reliability because the values of the tests are greater than the cut-off value of 0.60.

The Average Variance Extraction (AVE) technique is frequently employed to assess convergent validity. According to Hair, Ringle, & Sarstedt (2011), the acceptable parameters for AVE require that the value be more than 0.5. The findings of the AVE reported in Table 14 demonstrate that the AVE value for each construct is greater than the cut-off value of 0.5; hence, there is no problem with the data's convergent validity.

Table 14: Reliability and Validity

	Cronbach's Alpha	Composite Reliability	Average Variance Extracted (AVE)
BI	0.851	0.899	0.691
EE	0.715	0.874	0.777
HM	0.888	0.946	0.898
PE	0.744	0.854	0.663
SI	0.66	0.815	0.594
UB	0.733	0.847	0.65

The measurement of the degree of differentiation between constructs is determined by their discriminant validity, as stated by Barclay, Higgins, & Thompson (1995, p.295). As per the measurement model, it is expected that a specific construct would exhibit a higher degree of variance with its corresponding measures as compared to the variance it shares with other constructs (Bagozzi & Yi., 1988, Hulland, 1999). The discriminant validity could be evaluated by utilizing the criteria put forth by Fornell and Larcker (1981). To adhere to the criteria outlined by Fornell and Larcker (1981), it is imperative that the square root of the Average Variance Extracted (AVE) surpasses the correlations that are present among the latent variables. Table 15 reveals that the square root of the AVE of Performance Expectancy, Effort Expectancy, Social Influence, Hedonic Motivation, Behavior Intentions, and Use Behavior exhibits a higher value than the inter-correlations among the components, thereby affirming the validation of discriminant validity.

Table 15: Discriminant Validity

	BI	EE	HM	PE	SI	UB
BI	0.831					
EE	0.53	0.882				
HM	0.52	0.337	0.948			
PE	0.55	0.417	0.266	0.814		
SI	0.531	0.505	0.455	0.456	0.771	
UB	0.718	0.479	0.563	0.397	0.541	0.806

5.4.2 Structural model

Partial Least Square Structural Equation Modelling (PLS-SEM) is applied to test the hypotheses of the study. Table 16 and Figure 3 report the results of the structural model. The results demonstrate that Performance Expectancy positively influenced the intentions to use ChatGPT in higher education which is statistically significant at a 5% significance level ($\beta = 0.313$, $p = 0.000$). Thus, the first hypothesis is accepted. Effort Expectancy also positively and significantly affect Behavior Intentions ($\beta = 0.228$, $p = 0.026$). Thus, the second hypothesis is also accepted which means that Effort Expectancy will positively predict the intention to use ChatGPT in higher education among educators, as they perceive the technology as easy to use and requires minimal effort and resources. Social influences positively but insignificantly affect Behavioral Intentions ($\beta = 0.138$, $p = 0.257$). Thus, the third hypothesis is rejected. In addition, Hedonic Motivation positively and significantly affects Behavior Intentions ($\beta = 0.296$, $p = 0.004$). Hence, the fourth hypothesis is accepted which implies that Hedonic Motivation will positively predict the intention to use ChatGPT in higher education among educators, as they are motivated by enjoyment and satisfaction ($\beta = 0.720$, $p = 0.000$). Additionally, Behavior Intentions positively and significantly affect the Use Behavior of ChatGPT. The ninth hypothesis is also accepted that Behavior Intention will positively predict the Use Behavior of ChatGPT in higher education among educators.

Table 16: Structural Model

Path	Coefficient	T Statistics	P Values	Hypothesis	Decision
PE -> BI	0.313	3.620	0.000	H1	Accepted
EE -> BI	0.228	2.226	0.026	H2	Accepted
SI -> BI	0.138	1.136	0.257	H3	Rejected
HM -> BI	0.296	2.875	0.004	H4	Accepted
BI -> UB	0.720	11.246	0.000	H9	Accepted

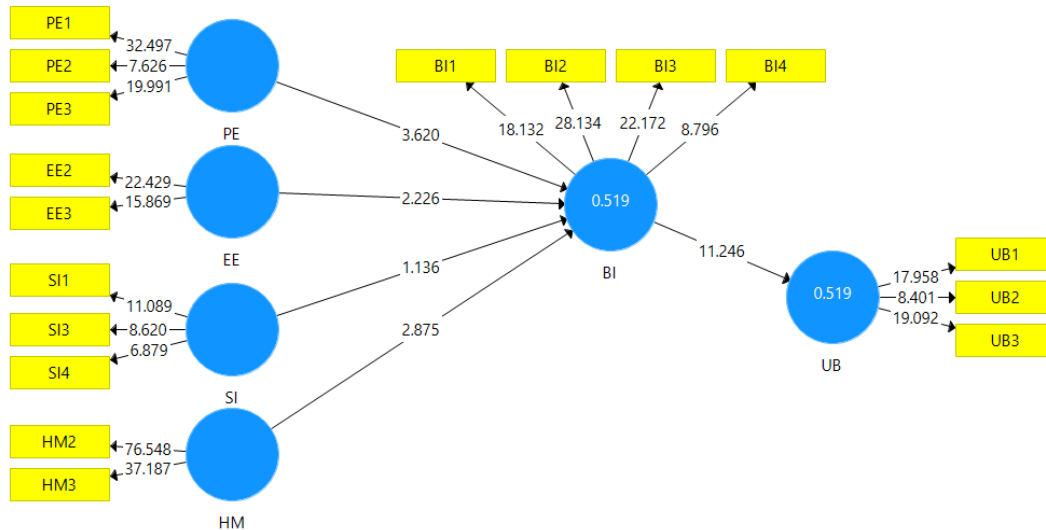


Figure 3: PLS-SEM

5.4.3 Moderation analysis

Moderation analysis is applied to observe the moderating effect of Age, Gender, Academic degree, and Field. The results are presented in Figure 4 and Table 17. According to the results, Age, Gender, Academic degree, and Field has no significant moderating effect on the relationship between Performance Expectancy and intentions to use ChatGPT in higher education. In addition, the moderating effect of Age, Gender, Academic degree, and Field on the relationship between Effort Expectancy and intentions to use ChatGPT in higher education among educators is insignificant. Further, the moderating effect of Age, Gender, Academic degree, and Field on the relationship between Social Influence and intentions to use ChatGPT in higher education among educators. Moreover, the results confirm that Age, Gender, Academic degree, and Field insignificantly moderate the relationship between Hedonic Motivation and educators' intentions to use ChatGPT in higher education. Thus, Hypotheses H5, H6, H7, and H8 are rejected.

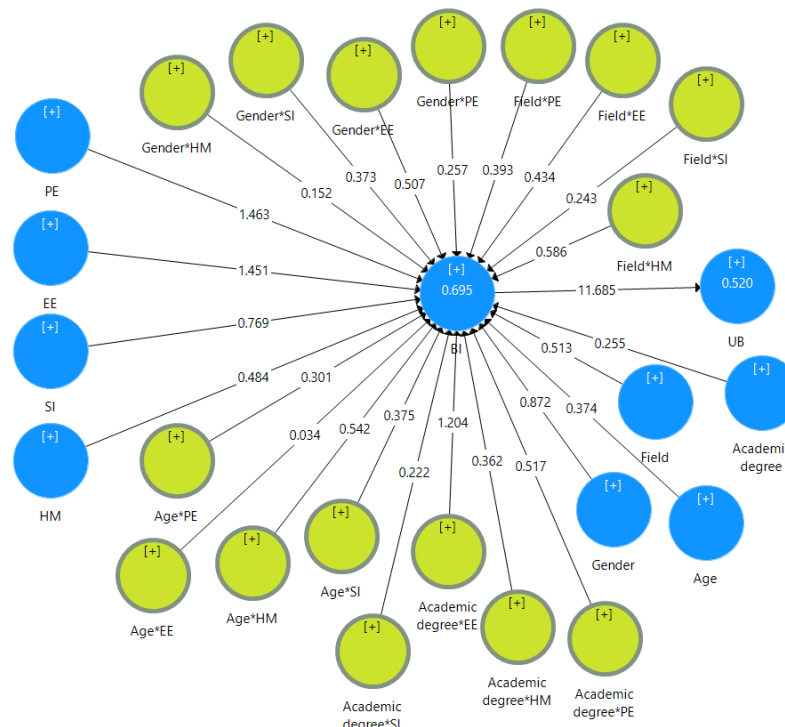


Figure 4: Structural Model

Table 17: Moderation Analysis

Path	Coefficients	T-Statistic	P Values	Hypothesis	Decision
Age*PE -> Behavior Intention	-0.082	0.301	0.763	H5	Rejected
Gender*PE -> Behavior Intention	0.062	0.257	0.797		
Academic degree*PE -> Behavior Intention	0.144	0.517	0.606		
Field*PE -> Behavior Intention	0.11	0.393	0.694		
Age*EE -> Behavior Intention	0.009	0.034	0.973	H6	Rejected
Gender*EE -> Behavior Intention	0.117	0.507	0.612		
Academic degree*EE -> Behavior Intention	-0.34	1.204	0.229		
Field*EE -> Behavior Intention	0.126	0.434	0.664		
Age*SI -> Behavior Intention	-0.116	0.375	0.708	H7	Rejected
Gender*SI -> Behavior Intention	-0.092	0.373	0.709		
Academic degree*SI -> Behavior Intention	-0.062	0.222	0.825		
Field*SI -> Behavior Intention	0.063	0.243	0.808		
Age*HM -> Behavior Intention	0.14	0.542	0.588	H8	Rejected
Gender*HM -> Behavior Intention	0.032	0.152	0.879		
Academic degree*HM -> Behavior Intention	0.121	0.362	0.718		
Field*HM -> Behavior Intention	-0.111	0.586	0.558		

6. Discussion

The research findings presented in the previous section provide valuable insights into the demographic characteristics of the users, their usage patterns, and the reasons behind their adoption or non-adoption of ChatGPT. The demographic information revealed that the majority of the participants were female (60.28%) and belonged to the age groups of Millennials (43.97%) and Gen X (46.81%). In terms of academic degrees, a significant proportion of the participants held Master's degrees (45%), followed by PhD degree holders (41%). The distribution of respondents across various fields indicated a diverse representation, with the highest percentage coming from the teaching field (36.88%), followed by information technology (24.82%).

Regarding the usage of ChatGPT, it was found that 46.81% of the participants had already used or tried it. The reasons for usage varied, with research (37.9%) and teaching and learning (19.7%) being the primary purposes. For those who had not used ChatGPT, the main reasons cited were lack of interest (44%) and finding it difficult (8%).

The t-test analysis examined the differences in behavior intentions and use behavior between males and females. The results showed that there were no statistically significant differences between genders for both behavior intentions and use behavior.

The study analyzed the ANOVA results to explore potential differences based on age groups and academic degrees. Furthermore, the study tested multiple hypotheses concerning educators' behavior intentions and use behavior concerning the adoption of ChatGPT in higher education. H1 revealed a significant positive effect of

Performance Expectancy on educators' intentions to use ChatGPT, indicating that educators who perceive ChatGPT as beneficial for their performance are more likely to have intentions to use it. H2 found a significant positive effect of Effort Expectancy on intentions, suggesting that educators who perceive ChatGPT as easy to use are more likely to have intentions to use it. However, H3 showed that Social Influence did not significantly influence intentions, indicating that the opinions and recommendations of others did not impact educators' intentions to use ChatGPT. H4 demonstrated a significant positive effect of Hedonic Motivation on intentions, indicating that educators who perceive ChatGPT as enjoyable and satisfying are more likely to have intentions to use it. H5 and H6 revealed no statistically significant differences in behavior intentions and use behavior between genders and different age groups, respectively, suggesting that gender and age do not play significant roles in shaping educators' attitudes and actions towards using ChatGPT. H7 highlighted a significant difference in behavior intentions based on participants' academic degrees, indicating that educators' academic qualifications influence their intentions to use ChatGPT in higher education. Finally, H8 showed no significant variations in behavior intentions and use behavior based on participants' field of study, indicating that educators from different academic disciplines demonstrate similar levels of behavior intentions and utilization of ChatGPT. Overall, these research results provide valuable insights into the demographic characteristics and usage patterns of participants regarding ChatGPT. The findings can be used to better understand the preferences and motivations of users, as well as to tailor the development and implementation of AI-powered chat systems like ChatGPT in higher education contexts.

Further, the research findings align with the existing literature on the use of OpenAI's GPT, in education. The literature review highlights the potential benefits of GPT in education, including content generation, personalized learning experiences, and assessment and feedback. These advantages can greatly enhance teaching and learning outcomes by providing tailored support to individual students, automating certain tasks for teachers, and improving the overall efficiency of the education system.

The literature review highlights the advantages of GPT in education, such as its ability to generate educational content, assist with grading and assessment, and provide personalized learning experiences. Wang et al. (2020) and Bansal and Chugh (2021) emphasize how AI, including GPT, can enhance teaching and learning outcomes by providing personalized feedback, analyzing data, and supporting various educational tasks. Ramasubramanian (2021) proposes a framework that harnesses GPT for content generation, personalization of learning experiences, and assessment and feedback in education.

However, the literature also acknowledges the limitations and ethical concerns associated with the use of GPT in education. These include the potential for biases in the generated content, the possibility of generating inappropriate or misleading information, the lack of transparency in how the models arrive at their predictions, and privacy concerns related to data collection and usage. Researchers such as Adesope (2020) and Li & Bamman (2021) raise concerns about bias and inappropriate content generation, while Buruk & Oğuz 'Oz (2023) highlight the need for transparency in the decision-making process of AI models. Additionally, privacy concerns surrounding student data and the responsible use of AI in education are important considerations (Adesope, 2020). However, the literature also sheds light on the limitations and ethical concerns associated with the use of GPT in education. One major concern is the issue of bias in the generated content. As GPT models are trained on large text datasets, they can inadvertently reflect the biases present in the training data. This can increase societal biases, potentially leading to unfair treatment or inaccurate information being provided to students. It is crucial to carefully select and curate training data to mitigate bias and ensure equitable educational experiences for all students. Another ethical concern highlighted in the literature is the potential for generating inappropriate or misleading content. GPT models have been found to associate certain attributes or roles with specific genders or ethnicities, reflecting societal biases. This not only reinforces stereotypes but also has the potential to negatively impact students' emotional well-being, civic understanding, and social interactions. Safeguards should be implemented to detect and mitigate the generation of such biased or inappropriate content. Transparency is also a significant ethical concern raised in the literature. The increasingly complex nature of Open AI's makes it difficult to understand how they arrive at their predictions or outputs. This lack of transparency raises questions about accountability and the ability to identify and address ethical problems or prejudices embedded in the models. Efforts should be made to enhance the interpretability and explainability of AI models, ensuring that teachers and students can understand the decisions made by the models. Privacy emerges as a crucial issue when using AI in education. The collection and analysis of large amounts of student data by AI systems raise concerns about data security and the potential misuse of personal information. Clear policies and guidelines need to be established to regulate data collection, usage, and protection, ensuring that student privacy is assured and data is used solely for educational purposes.

The literature review also discusses the current initiatives and actions taken by educational institutions in response to the integration of GPT in education. Some universities have initiated reviews and discussions around the ethical implications and potential biases of AI language models (New York Times, 2023), while others have implemented restrictions or guidelines for using GPT in teaching and learning settings (Harris, 2023). The need for collaboration between educators, software developers, researchers, and policy makers is emphasized to ensure a thoughtful and cautious integration of AI language models in education (Kolb & Kolb, 2021). Given these considerations, it is evident that the integration of ChatGPT or any Open AI in education requires a cautious and thoughtful approach. The research objectives of the current study, which aim to examine the socio-demographic factors, motivations, and social influences that affect the usage of ChatGPT among educators, are important steps in understanding the practical implications and challenges associated with the adoption of AI language models in educational settings. By addressing these factors, educational institutions can develop guidelines, provide appropriate training to teachers and students, and establish policies that promote responsible and ethical use of AI in education.

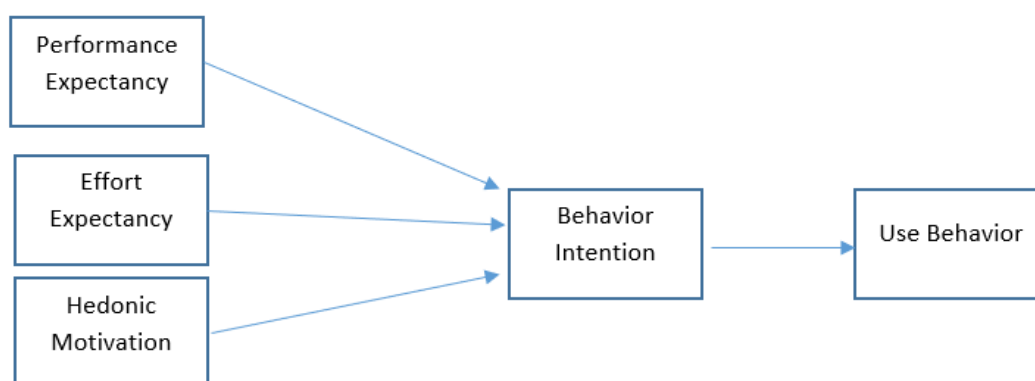


Figure 5: Updated model: Key factors to influence Behavior Intention and Use Behavior for ChatGPT

7. Conclusion

In conclusion, the integration of Large Language Models (LLMs), such as ChatGPT, into education is a multifaceted process that demands collaboration among educators, software developers, researchers, and policymakers (Kolb & Kolb, 2021). This aligns with previous studies, such as the research by Zhang, Schießl, Plößl, Hofmann, and Gläser-Zikuda (2023), which emphasize the importance of understanding educators', particularly pre-service teachers', acceptance of AI technologies for effective integration in educational settings. The evolving landscape, marked by an increasing number of LLMs, underscores the need for a comprehensive approach in shaping learning design for the future. As educational institutions embark on this transformative journey, a thoughtful and cautious strategy is imperative to leverage the potential of LLMs for enhanced learning outcomes.

Educational institutes must embark on developing a strategic roadmap that not only incorporates existing AI language models like GPT but also anticipates and adapts to the continuous advancements in this technology. The identification of areas where LLMs could be effectively employed (C. Bonk & K. Lee, 2020) becomes crucial, necessitating ongoing collaboration with educators to understand their evolving needs and the dynamic landscape of educational content.

Training becomes a linchpin in this process, not only focusing on educators but also extending to students. Effectively utilizing LLMs involves empowering both educators and students to harness the capabilities of these models for personalized learning experiences (Cordeiro, Blikstein, & Cunha, 2018). As the educational landscape evolves, a paradigm shift in learning design is essential, emphasizing adaptability and responsiveness to the diverse needs of students.

The growing presence of LLMs prompts a deeper exploration of ethical considerations, particularly concerning biases in generated content. The careful curation of training data becomes paramount to ensure that the educational experience is equitable for all students, irrespective of their background or characteristics.

Safeguards against the generation of inappropriate or misleading content must be embedded in the learning design, fostering inclusive and unbiased learning environments.

Transparency emerges as a critical factor in the design of learning experiences involving LLMs. The inherent complexity of these models requires a concerted effort to enhance interpretability and explainability. This transparency not only ensures accountability but also allows educators and students to comprehend the reasoning behind the models' outputs, fostering a more informed and collaborative educational environment.

Privacy concerns, particularly regarding data collection and usage, continue to be at the forefront of considerations. Educational institutions must establish clear policies and guidelines regulating the collection, usage, and protection of student data. Striking a balance between harnessing the power of LLMs for educational purposes and safeguarding data security is integral to the responsible integration of these technologies.

The findings from the present study and the broader literature review contribute valuable insights into the evolving landscape of LLMs in education. These insights should inform the iterative development of learning design, accounting for socio-demographic factors, motivations, and social influences affecting educators' usage of LLMs in higher education. The resulting learning designs should be dynamic, adaptable, and responsive to the evolving capabilities and challenges posed by LLMs.

As the education sector continues to witness the emergence of new LLMs, including the anticipated release of GPT-4, educational institutions must approach the integration of these models with foresight and adaptability. Learning designs should not only capitalize on the benefits offered by LLMs but also address their limitations, ensuring ethical standards, and upholding student privacy. By doing so, educational institutions can foster an environment that not only embraces technological advancements but also prioritizes fairness, equity, and accountability in education.

8. Further Research and Limitations

The study focused specifically on the use of ChatGPT in higher education, which means that the findings may not be directly applicable to other AI tools or platforms used in educational settings. It is important to recognize that different AI technologies may have unique features, functionalities, and implications for educators. Therefore, future research should aim to explore a wider range of AI technologies to gain a more comprehensive understanding of AI adoption in education and to identify the specific factors that influence educators' behavior intentions and use behavior across various AI platforms.

Additionally, it is crucial to consider the evolving nature of ChatGPT and other AI models. The study was conducted during the initial introduction of ChatGPT, and since then, significant advancements have been made in AI technology. These advancements may have influenced the capabilities, features, and user experiences associated with ChatGPT. Therefore, conducting a similar study with an updated sample and considering the current landscape of ChatGPT could provide valuable insights into the latest trends and dynamics of AI adoption in education. This will help educators, policymakers, and educational institutions make informed decisions and develop effective strategies for the integration of AI tools that align with the evolving needs and requirements of the education sector.

Moreover, the study's navigation through a relatively untapped area of research, where pre-existing literature is scarce. This limitation underscores the cutting-edge nature of the research, as it navigates a relatively unexplored area with limited pre-existing literature. Future studies can build upon this foundational research, expanding the knowledge base and exploring the evolving landscape of AI in education with a more extensive array of secondary sources. This progression will not only enrich the academic discourse but also provide practical insights for educators, policymakers, and institutions, guiding them in effectively harnessing AI technologies in educational contexts.

References

- Adesope, O. O. (2020). Ethical issues with the use of GPT models in education. In *Proceedings of the IEEE International Conference on Artificial Intelligence in Education (AIED)*, pp. 134-143.
- Ajzen, I. (1991). The Theory of Planned Behavior. *Organizational Behavior and Human Decision Processes*, 50, 179–211.
- Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, 16(1), 74–94.
- Bansal, P., & Chugh, R. (2021). GPT-3 in Education: A Comprehensive Review. *International Journal of Advanced Computer Science and Applications*, 12(3), 225-231.

- Barclay, D., Higgins, C., & Thompson, R. (1995). The partial least squares (PLS) approach to casual modeling: personal computer adoption and use as an illustration. *Technology Studies*, 2, 285–309.
- Bowdoin Science Journal. (2021). The ethical implications of using GPT models in education. Retrieved from <https://bowdoinsciencejournal.com/2021/04/06/the-ethical-implications-of-using-gpt-models-in-education/>
- Brown, S. A., and Venkatesh, V. 2005. "Model of Adoption of Technology in the Household: A Baseline Model Test and Extension Incorporating Household Life Cycle," *MIS Quarterly* (29:4), pp. 399-426.
- Buruk, O. and Oğuz 'Oz, E. (2023). Transparency issues in using GPT models for educational purposes. In Proceedings of the IEEE International Conference on Education and Technology (ICET), pp. 210-215.
- C. Bonk & K. Lee. (2020). Artificial Intelligence and Machine Learning in Education. Routledge.
- Davis, F. (1989) Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13, 319-340.
- Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention, and behavior: An introduction to theory and research. Addison-Wesley.
- Floridi, L., Chiriatti, M. GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds & Machines* 30, 681–694 (2020). <https://doi.org/10.1007/s11023-020-09548-1>
- Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research*, 18(3), 382–388.
- Hair, J. J. F., Hult, G. T. M., Ringle, C., & Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM)*. Thousand Oaks, CA: Sage Publication.
- Hair Jr, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice*, 19(2), 139–152.
- Harris, C. (2023, March 9). Medical science students were told to use ChatGPT. This is what it wrote. Sydney Morning Herald. <https://www.smh.com.au/national/nsw/medical-science-students-were-told-to-use-chatgpt-this-is-what-it-wrote-20230309-p5cqs4.html>
- Holmes W, Tuomi I (2021). State of the art and practice in AI in education. *European Journal of Education*. 2022, 57:542-70. 10.1111/ejed.12533
- Hulland, J. (1999). Use of partial least squares (PLS) in strategic management research: a review of four recent studies. *Strategic Management Journal*, 20(2), 195–204.
- Iqbal, N. Ahmed, H. & Azhar, K. (2023). Exploring Teachers' Attitudes towards Using Chat GPT. *Global Journal for Management and Administrative Sciences*. 3. 10.46568/gjmas.v3i4.163.
- Kolb, L., & Kolb, A. (2021). Using AI in education: Opportunities, risks and ethical considerations. *European Journal of Education*, 56(3), 337-349. doi: 10.1111/ejed.12424
- Kumar, V., & Bhattacharya, S. (2021). A Review on Language Models and its Applications in Education Domain. 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), 293-297.
- Li, L. and Bamman, D. (2021). Gender bias in GPT-2 prompts: Evaluation and de-biasing methods. In Proceedings of the IEEE Conference on Ethics in Natural Language Processing (ENLP), pp. 42-49.
- Liu, X., Santhanam, R., & Webster, J. (2010). Toward a better understanding of the link between gender and technology adoption: A structural equation model with moderating effects of gender differences. *Journal of the Association for Information Systems*, 11(6), 316-337.
- New York Times. (2023, January 16). Alarmed by A.I. Chatbots, Universities Start Revamping How They Teach. <https://www.nytimes.com/2023/01/16/technology/chatgpt-artificial-intelligence-universities>.
- M. Cordeiro, M. Blikstein, & R. N. Cunha. (2018). Designing for Technology-Enhanced Learning in Developing Countries: A Hybrid Approach Using Artificial Intelligence and Crowdsourcing. *International Journal of Artificial Intelligence in Education*, 28(3), 291-320.
- Mizumoto, A., & Eguchi, M. (2021). Exploring the Potential of Using an Ai Language Model for Automated Essay Scoring. Available at SSRN: <https://ssrn.com/abstract=4373111> or <http://dx.doi.org/10.2139/ssrn.4373111>.
- Ramasubramanian, R. (2021). GPT-3 Based Personalised Learning Framework for School Education. 2021 International Conference on Advanced Computational and Communication Paradigms (ICACCP), 663-666.
- Rogers, E. M. (1962). Diffusion of innovations. Glencoe, IL: Free Press.
- Srinivasan, V., & Padma, P. (2021). GPT-3 and its Applications in Education-A Survey. 2021 IEEE International Conference on Engineering Education (EDUCON), 1996-2001.
- Russell, T. W., Wu, J. T., Clifford, S., Edmunds, W. J., Kucharski, A. J., Jit, M., and Centre for the Mathematical Modelling of Infectious Diseases COVID-19 working group (2021). Effect of internationally imported cases on internal spread of COVID-19: A mathematical modelling study. *The Lancet. Public Health*, 6(1):e12–e20
- Tamilmanim, K., Ran, N.P., Wamba, S., Dwivedi, R. (2021). The extended Unified Theory of Acceptance and Use of Technology (UTAUT2): A systematic literature review and theory evaluation. *International Journal of Information Management*, Volume 57.
- Teo, T. (2009). Modelling technology acceptance in education: A study of pre-service teachers. *Computers & Education*, 52(2), 302-312.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478.
- Venkatesh, V., Thong, J. Y., & Xu, X. (2012). Consumer acceptance and use of information technology: extending the unified theory of acceptance and use of technology. *MIS Quarterly*, 36(1), 157-178.

- Wang, X., Zhou, Y., Li, S., & Li, Z. (2020). Intelligent education: An overview and the future. *Journal of Educational Technology Development and Exchange*, 13(1), 1-14
- Wang, Y. S., Wu, M. C., & Wang, H. Y. (2011). Investigating the determinants and age and gender differences in the acceptance of mobile learning. *British Journal of Educational Technology*, 42(4), 592- 605.
- Zhang, C., Schießl, J., Plöchl, L, Hoffman, F. & Zikuda, M. (2023) Acceptance of artificial intelligence among pre-service teachers: a multigroup analysis. *Int J Educ Technol High Educ* 20, 49 (2023). <https://doi.org/10.1186/s41239-023-00420-7>