ChatGPT in Education – Understanding the Bahraini Academics Perspective

Amal Alrayes[1], Tara Fryad Henari[1] and Dalal Abdulkarim Ahmed[2]
[1]College of Information Technology, University of Bahrain, Bahrain
[2]English 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

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

Cite this article: Alrayes, A., Henari, T. F. and Ahmed, D. A. 2024. ChatGPT in Education – Understanding the Bahraini Academics Perspective, *The Electronic Journal of e-Learning*, pp 112-134, https://doi.org/10.34190/ejel.22.2.3250

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 modal 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 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 UTAT 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 UTAT2 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 us 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 at 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 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 at 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 teacheroriented 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 indicted that gender plays a role in the relationship of PE and intention to use (Liu at 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 at 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.

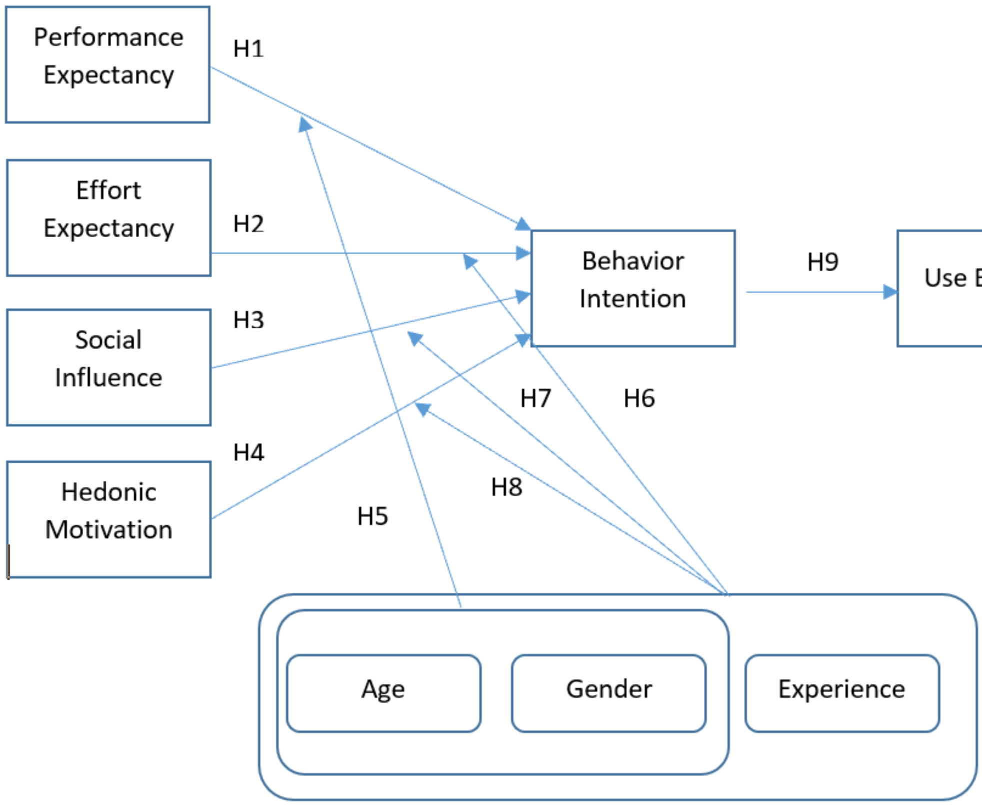

Figure 1: Research Model

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 Leven'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	-0.518
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	-0.734

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

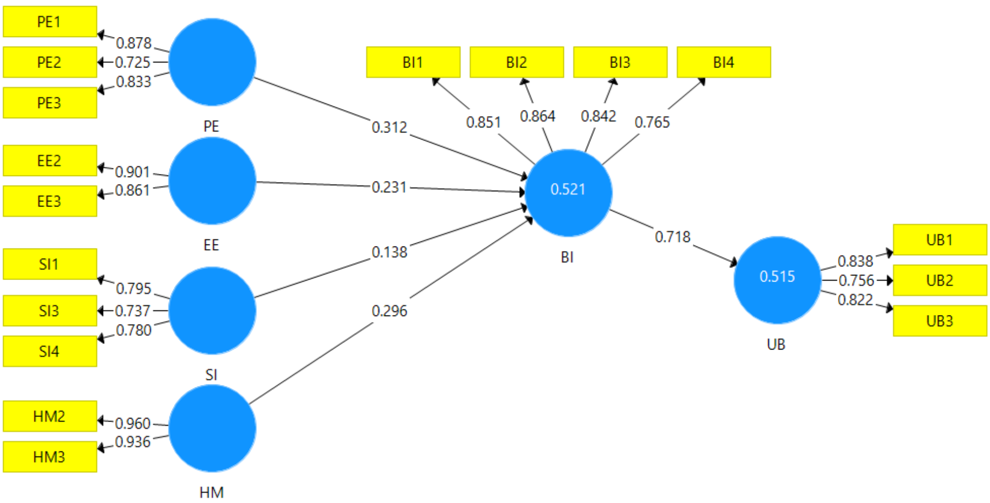

Figure 2: Factor Loadings

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 (*β = 0.313, p = 0.000*). Thus, the first hypothesis is accepted. Effort Expectancy also positively and significantly affect Behavior Intentions (*β = 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 (*β = 0.138, p = 0.257)*. Thus, the third hypothesis is rejected. In addition, Hedonic Motivation positively and significantly affects Behavior Intentions (*β = 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 (*β = 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

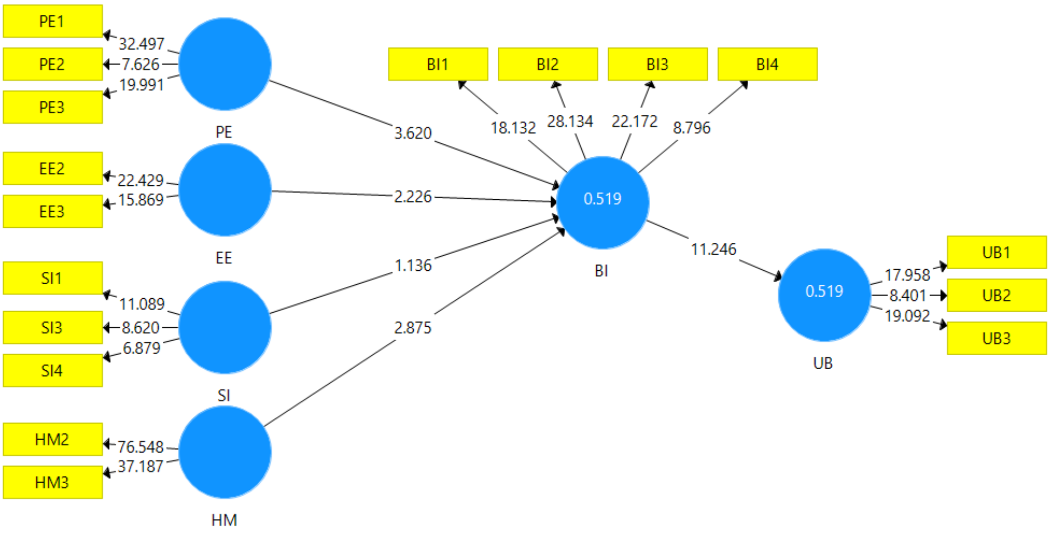

Figure 3: PLS-SEM

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.

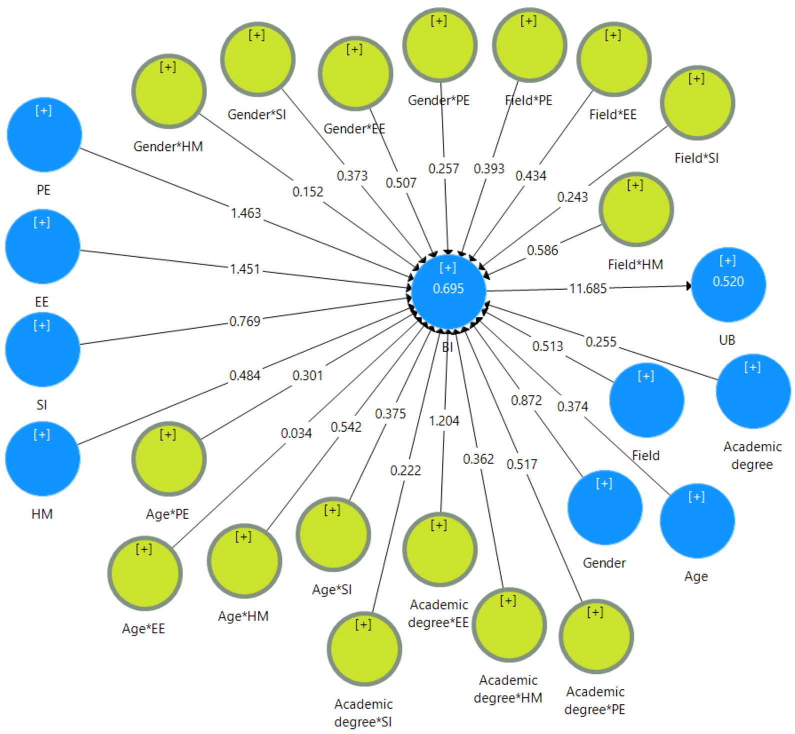

Figure 4: Structural Model

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

 ISSN 1479-4403

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.

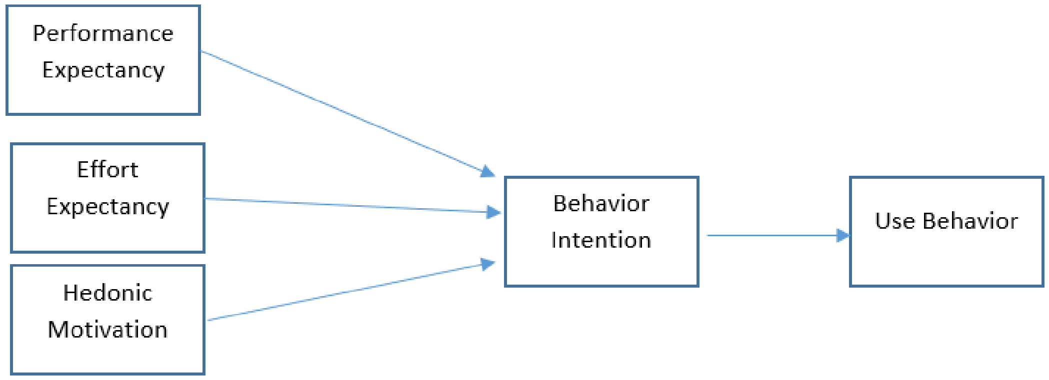

Figure 5: Updated model: Key factors to influence Behavior Intention and Use Behavior for ChatGPT

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 ans 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., Plößl, 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

