Uppsala Universitet
Department of Informatics and Media

Students Acceptance and Use of ChatGPT in Academic Settings

UPPSALA
UNIVERSITET

Master Thesis
Jakob Hasselqvist Haglund
Spring 2023
Abstract

The swift progression of technology has radically reshaped our lives, becoming a big part of our daily routines and paving the way for advancements in communication, automation, and information processing. OpenAI, a company at the forefront of artificial intelligence since 2015, has made remarkable strides towards making AI accessible and beneficial for all (OpenAI, n.d.). A notable accomplishment in their journey has been the development of Chat Generative Pre-trained Transformers (ChatGPT). This study aims to identify and explore the factors influencing students' acceptance and use of ChatGPT in academic settings. Despite the rising prominence of ChatGPT across various disciplines, understanding its acceptance and utilization, particularly within the sphere of higher education, remains limited. ChatGPT holds immense potential as a valuable asset for both students and educators. Utilizing the Unified Theory of Acceptance and Use of Technology (UTAUT) and a quantitative research approach, investigating these factors. The results suggest that student acceptance and use lies in Behavioral Intention, while Behavioral Initiation is influenced by both Effort Expectancy and Performance Expectancy.

Key words:
Chat Generative Pre-trained Transformers, ChatGPT, UTAUT, Academic Settings, quantitative research, academic settings
Index of figures and tables

Figure 1. The UTAUT model by Venkatesh et al. (2003) 14
Figure 2. Monotonic- and Non-Monotonic relationships from LærdStatistics (2018) 22
Figure 3. PLS-SEM Inner- and Outer Model from Wong (2016) 23
Figure 4. PLS-SEM model from Wong (2016) 24
Figure 5. Respondents current level of education 27
Figure 6. Subject that the student are currently studying 28
Figure 7. Ways student used ChatGPT to support studies 28
Figure 8. Age and Gender description from Jamovi 29
Figure 9. Performance Expectancy bar chart from Jamovi 30
Figure 10. Effort Expectancy bar chart from Jamovi 32
Figure 11. Social Influence bar chart from Jamovi 33
Figure 12. Facilitating Conditions bar chart from Jamovi 35
Figure 13. Result from PE4, EE4 and FC1 37
Figure 14. Spearman's Rank Correlation Coefficient result for PE, EE, SI and FC 38
Figure 15. PLS-SEM redefined model 39
Table 1. Performance Expectancy response result 31
Table 2. Effort Expectancy response result 32
Table 3. Social Influence response result 34
Table 4. Facilitating Condition response result 35
Table 5. Mann-Whitney U Test for Gender 36
Table 6. Mann-Whitney U Test for Age 36
Table 7. Kruskal-Wallis H Test for Experience 37
Table 8. Kruskal-Wallis H Test for Voluntariness of Use 38
1. Introduction

This chapter starts with describing the background of OpenAI, ChatGPT technology, the effect on academia and the problematization of the study. Then existing research of the subject will be presented which leads to purpose and research questions. Lastly, the delimitations of the study will be presented.

1.1 Background

The rapid advancement of technology has profoundly transformed the way we live today, becoming an integral part of our everyday lives and ushering in a new era of progress in communication, automation, and information processing. Artificial Intelligence (AI), in particular, has revolutionized numerous industries, including the field of academia.

Natural Language Processing (NLP), an interdisciplinary subfield encompassing linguistics, computer science, and AI, serves as a powerful tool for interacting with humans by processing and analyzing vast quantities of natural language data (Chowdhury, 2003). This technology has demonstrated remarkable effectiveness in understanding context, interpreting, and generating human language (Chowdhury, 2003). Over the years, NLP has made significant strides over the years, surpassing human capabilities in many language processing tasks (Kasneci et al, 2023). Today advanced NLP models offer invaluable assistance and support in a wide range of applications, enhancing efficiency and accuracy in numerous domains (Kasneci et al, 2023).

One significant advancement which has refined the foundational mechanisms of AI, is the Transformer Model. Harnessing the power of an attention mechanism, transformers have notably enhanced the capacity to discern the relative importance of various segments within an input when formulating predictions. This enhancement has significantly improved the performance of text interpretation and handling, especially in terms of grasping relationships between words. (Kasneci et al., 2023)

A pivotal development in this context has been the evolution of pre-training processes. By training Natural Language Processing (NLP) models on vast amounts of data, followed by task-specific fine-tuning, the understanding and recognition capabilities of these models have been considerably augmented. This methodology has shown to be particularly effective across a range of tasks, such as sentence comprehension, question-answering, and entity recognition. (Kasneci et al., 2023)

Chat Generative Pre-trained Transformers (ChatGPT) is an exemplary technology that employs these refined techniques. Owing to its success, it has been widely adopted and continues to be a popular choice in the realm of advanced language models.

1.2 OpenAI and ChatGPT

OpenAI, a pioneer in artificial intelligence research and development since 2015, has made significant strides in democratizing AI for everyone's benefit (OpenAI, n.d.). One of their standout achievements has been the creation of ChatGPT, an AI model that has garnered rapid and substantial growth in its user base (Hu, 2023).
ChatGPT is developed based on a transformer architecture, trained for the task of predicting subsequent tokens in a given sequence. This training enables it to serve as an effective chatbot, equipped to handle a diverse array of queries. In essence, it is a conversational AI tool trained on a vast corpus of internet text, which allows it to comprehend and generate language patterns, leading to the production of contextually relevant and coherent responses. (OpenAI, n.d.)

OpenAI has made two versions of ChatGPT available to the public: the free-of-charge ChatGPT-3 and its more advanced, fee-based counterpart, ChatGPT-4. The latter is an improved version designed to offer users more precise and contextually fitting outputs. Despite certain limitations compared to human abilities in real-world situations, ChatGPT has shown remarkable competencies, especially evident in its performance on several professional and academic tests. Its most notable achievement is scoring in the top 10% on a simulated bar exam, underscoring its sophisticated language comprehension capabilities. (OpenAI, n.d.)

ChatGPT's effect on academia is a topic of significance in today's debate and exploration. The rapid growth and usage by students have led to an urgent call for educators and institutions for better understanding of this technology, including its limitation and occasional unpredictability (Kasnecei, 2023). The opinions on ChatGPT's role in education are divided. While some see the potential for enhanced learning and reduction of teachers' workload, others see it as a way for students to cheat and plagiarize (Kirk, 2023).

1.3 Problem

Despite the growing prominence of ChatGPT in various fields, there remains a gap in understanding its acceptance and usage, particularly in higher education. As a powerful tool with both strengths and limitations, ChatGPT has the potential to serve as a valuable resource for students and educators alike. Lundberg, a historian at Lund University, suggests that ChatGPT's analytical capabilities are highly valued in academia and could even gain approval for integration into certain university curricula (SvD, 2022). This development poses challenges for universities as they seek to leverage these advanced analytical tools in higher education settings.

Survey studies examining ChatGPT usage in academic contexts reveal varying results. One study indicates that 55% of college students utilize ChatGPT, with 22% of these students specifically using it for academic purposes (Welding, 2023). In contrast, another study reports that 89% of students have used ChatGPT in academic settings (Studie.com, n.d.). Despite the differing findings, it is clear that further research is needed to better understand the acceptance and use of ChatGPT in higher education.

Moreover, the variability in the usage and acceptance of ChatGPT among academic institutions indicates a complexity that requires thorough investigation (Kasnecei, 2023). Thus, as AI-driven tools, like ChatGPT, are getting more integrated in education, it becomes more important for policy makers and universities to gain a higher understanding of the mechanisms behind students’ acceptance and use. Therefore, a study of how students utilize ChatGPT within academic settings not only helps enrich the field of research but also helps optimize the experience and outcomes for students.

Furthermore, understanding of how ChatGPT is accepted and used could provide valuable insights to its developers, leading to enhancements that align more closely with the requirements of academic users. Given AI's transformative capabilities and the fast-paced evolution of technology, it's crucial to
comprehend its real-world ramifications, particularly in academic contexts, to foster its ethical and advantageous application.

1.4 Purpose

The objective of this study is to unravel the determinants impacting students' embrace and utilization of ChatGPT within their academic pursuits. This investigation serves three primary goals:

First, it could generate valuable insights that may aid in mitigating risks and capitalizing on the potential benefits of ChatGPT use within education contexts.

Second, comprehending the acceptance and application of ChatGPT by students may shed light on the possible risks and advantages associated with its use.

Lastly, given the paucity of existing research on this topic, this study aspires to enrich the field of Information Systems and inform decision makers and developers by offering novel insights into the ways students accept and use ChatGPT in academic settings.

1.5 Existing Research

This subchapter reviews existing research on ChatGPT's applications and implications in various domains, with a particular focus on its impact on education. There is numerous studies that focus on ChatGPT's role in generating educational materials, serving as a learning process assistant, fostering curiosity, enhancing self-directed learning, and assisting with complex tasks. Furthermore, explorations into research that assesses ChatGPT's performance in translation, mathematical capabilities, and its ease of use. We will also cover studies that examine user acceptance, technology acceptance models, and the potential challenges and limitations of ChatGPT. The research presented here highlights the diverse ways in which ChatGPT is shaping the landscape of education and other fields, offering valuable insights into its strengths, weaknesses, and potential future developments.

**Education ChatGPT research**

Language models like ChatGPT present remarkable opportunities and pose distinct challenges when implemented in educational contexts, affecting both students and teachers. While ChatGPT has the capacity to create educational materials, boost student engagement and interaction, and customize learning experiences, educators must be well-versed in the model's intricacies to fully exploit its benefits. Acquiring the requisite knowledge to grasp the technology, its limitations, and occasional unpredictability is crucial for teachers. A holistic approach to education and communication is vital for maximizing ChatGPT's potential in academic settings and optimizing its influence on the teaching profession. Moreover, addressing output bias, human supervision, and potential misuse is essential and if managed correctly, ChatGPT can offer valuable insights and opportunities in various educational areas, including reading, writing, problem-solving, and more, at both basic and advanced levels. (Kasneci, 2023)

Research also suggests that ChatGPT can serve as a learning process assistant. Dijkstra et al. (2022) demonstrated that ChatGPT-3 can generate multiple-choice questions with accurate and inaccurate answers, surpassing other quiz-generating software in performance. Although ChatGPT-3 primarily generates valid multiple-choice questions, human oversight remains necessary due to occasional inaccuracies. It also struggles with creating questions in languages other than English. However, the updated ChatGPT-4 boasts improved language capabilities and accuracy (OpenAI, 2023).
Rania Abdelghani et al. (2023) indicates that the ability to ask curiosity-driven questions is a critical skill that fuels the learning process for students. Research indicates that ChatGPT-3 supports children in generating more curious questions, making it a suitable tool for fostering curious question-asking skills. This evidence encourages the use of ChatGPT as a curiosity-stimulating approach.

Nguyen et al. (2022) emphasizes the significance of students engaging in activities and practicing their skills through assessments to enhance learning. Researchers tested ChatGPT-3’s ability to generate assessment questions based on reading materials from a data science course. The results indicated that ChatGPT-3’s output was favorable, achieving an acceptable level as evaluated by two human expert raters.

MacNeil et al. (2023) means that a significant aspect of students' education involves tackling scientific problems. Research suggests that ChatGPT can effectively assist students with complex programming tasks by providing explanations, tracing code execution, defining terms, offering hints, and delivering error feedback.

**Usage ChatGPT research**

Azcel & Wagenmakers (n.d.) discussing the application of Large Language Model (LLM), like ChatGPT, in scientific writing offers guidelines for effectively using and presenting such models. While acknowledging the limitations, the paper also highlights the potential for increased sophistication and accuracy in ChatGPT’s future outputs. In order to maximize the benefits and mitigate the effects of these limitations, the paper proposes three transparency and credibility safeguards: informing users when an LLM is being employed, attributing the LLM’s contribution, and subjecting the LLM’s output to human verification.

Firat (2023) indicates that ChatGPT shows promise as a tool for open education, as it can enhance self-directed learning and adaptability for autodidactic students while maintaining efficiency and flexibility. By providing customized guidance, mentorship, and evaluation, ChatGPT may increase motivation and engagement for those engaged in self-directed learning. This approach enables access to high-quality learning resources and materials, making education more affordable and accessible.

In the article "ChatGPT Usage and Limitations," Azaria (2022) examines the strengths and weaknesses of the AI model, primarily focusing on its limitations. The author finds that the model has difficulties processing lengthy mathematical expressions, generating biased responses, and producing varying answers due to small changes. On a positive note, the AI is capable of rectifying its responses when shown to be incorrect and can request supplementary information if needed.

ChatGPT's applicability expands beyond educational contexts, proving beneficial in diverse fields, including climate change research. In Biswas (2023), the potential impact of ChatGPT on understanding climate change and enhancing the precision of climate projections is emphasized. As a powerful tool, ChatGPT can aid researchers and policymakers in generating and analyzing various climate scenarios. The authors find ChatGPT particularly useful for data analysis and interpretation. Furthermore, ChatGPT can facilitate collaboration between researchers and policymakers by synthesizing diverse sources of knowledge, fostering interdisciplinary dialogue, and enabling more informed decision-making in climate change mitigation and adaptation strategies. However, it is essential for researchers and policymakers to consider the limitations of ChatGPT. The model's output may not always be flawless, and they should be cautious of potential biases or inaccuracies that may arise.
Acceptance model of ChatGPT research
One paper employs the Technology Acceptance Model to examine the factors influencing user acceptance and the utilization of chatbots, specifically focusing on the human-technology interaction between entrepreneurs and ChatGPT. The author, Rahaman (2023), identifies ChatGPT as a valuable resource for addressing the challenges entrepreneurs often face, such as time management, business competition, and financial concerns. Thus, ChatGPT can help mitigate these problems by automating repetitive tasks, providing relevant information, and facilitating decision-making processes. Rahaman's research suggests that ChatGPT holds considerable promise as a tool for entrepreneurs, provided it is perceived as both user-friendly and valuable for addressing their unique challenges.

Performance of ChatGPT research
Subramani et al. (2023), demonstrates that ChatGPT has the potential to be a valuable tool in medical education, particularly for medical physiology. The study highlights the model's ability to provide well-reasoned and comprehensive answers to complex questions. The researchers evaluated ChatGPT's performance by having it complete a Medical Physiology Examination for Phase I MBBS (Bachelor of Medicine, Bachelor of Surgery) students. This examination consisted of both theoretical questions and multiple-choice questions. The answers generated by ChatGPT were assessed by the faculty of the Department of Physiology. Impressively, ChatGPT completed the exam with distinction (>75%), and the answers it provided contained thorough explanations. However, the study identified one limitation: ChatGPT was unable to generate relevant diagrams alongside the answers when needed for certain questions.

Jiao et al. (2023), evaluates ChatGPT-3's translation performance by comparing it to commercial translation products like Google Translate. The study finds that ChatGPT-3 performs well with high-resource European languages but struggles with low-resource or distant languages. To address this issue, the researchers employ a pivot-promoting strategy, using ChatGPT to translate low-resource languages into high-resource languages before generating translations into other languages. This approach significantly improves translation performance. The authors also note that while commercial translation products outperform ChatGPT-3 in biomedical terminology and Reddit comments, ChatGPT demonstrates superior results for spoken language translations. The study also highlights the improved engine of ChatGPT-4, which offers a notably better performance in language translation tasks.

Frieder et al. (2023) investigate the mathematical proficiency of ChatGPT-3 by testing it on various publicly available datasets and hand-crafted ones that cover multiple aspects of elementary school mathematics. The results are compared to the performance of other specialized mathematical models. The findings indicate that ChatGPT-3 is not yet capable of delivering high-quality proofs or calculations, performing significantly below the average mathematics student. While ChatGPT can understand the questions, it often fails to provide correct answers. In contrast, the results show that specialized models perform better on tasks they are explicitly trained for; however, ChatGPT demonstrates greater flexibility and outperforms these models as a more versatile tool in the field of mathematics.

Ease of use ChatGPT research
Bitzenbauer (2023) conducts a pilot study to explore the impact of ChatGPT's integration into secondary level school academic settings. The study employs a one-group pretest-posttest design and examines two straightforward 45-minute lesson examples that feature distinct ChatGPT activities. The results of the students' first-hand experiences with ChatGPT reveal a favorable influence on their perception of the AI model, particularly in relation to its benefits and everyday usability. This study
demonstrates that the implementation of ChatGPT in educational settings can positively impact students' perceptions and potentially enhance their learning experiences, even with simple and easy-to-implement activities.

**Conclusion**

In conclusion, the research reviewed in this chapter demonstrates the vast potential of ChatGPT in various domains, particularly in education. The studies highlight the AI model's capabilities in generating educational materials, serving as a learning assistant, fostering curiosity, enhancing self-directed learning and assisting with complex tasks. Furthermore, ChatGPT's performance in translation, mathematical proficiency, and ease of use has been examined, offering insights into its strengths and areas for improvement.

However, it is essential to acknowledge the limitations and challenges associated with ChatGPT, such as biases, inaccuracies, and the need for human oversight. Researchers and practitioners must carefully consider these factors when integrating ChatGPT into their respective fields. As AI technology continues to advance, the development of more sophisticated and accurate models is expected, which will further expand the potential applications and benefits of ChatGPT.

The only study that assesses the acceptance and use of ChatGPT is Saidur Rahaman (2023) which uses a so-called Technology Acceptance Model. This indicates that it is clear that further research is needed to better understand the acceptance and use of ChatGPT in academic settings and that there is a research gap in this area.

1.6 **Research question**

The research question of the thesis is:

*What factors lead to students acceptance and use of ChatGPT in Academic settings?*

1.7 **Delimitation**

This investigation is explicitly focused on data collection from students who have engaged with ChatGPT in an academic environment. For the purpose of this study, a 'student' is characterized as an individual currently undertaking a program of study, which may include a PhD, Master's, Bachelor's degree, High School, or a professional degree. The term 'using ChatGPT in academic settings' is applied broadly, referring to any instance where ChatGPT is utilized in a manner related to a studying environment. This can be any activity ranging from seeking assistance in academic assignments to employing the tool for knowledge enhancement. This delimitation ensures a concentrated examination of the factors influencing ChatGPT acceptance and use within the confines of an educational setting.
2. Theory

This chapter lays the theoretical foundation for exploring ChatGPT's acceptance and usage in academic settings. The main theoretical framework that will be used is called the Unified Theory of Acceptance and Use of Technology (UTAUT). First, an overview of UTAUT and its origin will be presented. Subsequently, the theories that collectively form UTAUT will be discussed, shedding light on their significance to technology adoption. We'll delve deeper into the four key constructs and four moderators of UTAUT, understanding how they influence technology acceptance and use. Lastly, the application and empirical evidence of UTAUT in past research will be examined to validate its selection for this study over alternatives such as Technology acceptance model (TAM) and Unified Theory of Acceptance and Use of Technology 2 (UTAUT2). This chapter aims to provide the theoretical backbone necessary for investigating the students' acceptance and use of ChatGPT within an academic context.

2.1 Unified theory of acceptance and use of Technology

The unified theory of acceptance and use of Technology was developed in 2003 by Venkatesh, Morris, Davis and Davis (2003) to unify many different models that focused on user and innovation acceptance of technology (Williams et al., 2015). The aim of UTAUT is to measure the probability of success for new technology and understand the drivers of acceptance (Ammenwerth, 2019). UTAUT suggests that four core constructs determine Behavioral Intention and User Behavior and these constructs are moderated by four key moderators (Venkatesh et al., 2003). By analyzing the impact of these constructs in practical settings, researchers and practitioners can evaluate an individual's intention to utilize a particular system or technology. This approach enables the identification of critical factors driving technology acceptance within a specific context (Williams et al., 2015). This makes the model suitable to answer the research question of this thesis.

2.2 UTAUT based theories

The unified theory of acceptance and use of technology (UTAUT) is based on analysis and comparisons of these eight theories (Venkatesh et al., 2003):

*The Theory of Reasoned Action* (TRA) is a theory that examines the relationship between attitude and behavior in human. It is based on two concepts: *Attitude Toward Behavior*, which is described as an individual's positive or negative attitude that predicts and explains their behavior, and *Subjective Norm*, which is described as an individual's perception of how important others think they should or should not behave. (Venkatesh et al., 2003; Alomary et al., 2020)

*The Theory of Acceptance Model* (TAM) is another theory in the field of information technology that models how users accept and use technology. TAM consists of three components: the first is *Perceived Usefulness*, which is defined as the extent to which an individual believes that using a particular system would improve their performance. The second is *Perceived Ease-of-Use*, which is defined as the extent to which an individual believes that using a particular system would be effortless. The last component is *Subjective Norm*, which is defined in the same way as in TRA. (Venkatesh et al., 2003; Alomary et al., 2020)

*The Motivational Model* (MM) is a theory based on motivation theories that explains behavior. It is based on two concepts: *Extrinsic Motivation*, which refers to motivation that is reinforced by external factors that have nothing to do with the activity itself. For example, an individual may perform a
certain job because it leads to a higher salary or promotion. The second concept, *Intrinsic Motivation*, pertains to motivations driven by the satisfaction and engagement derived solely from the task at hand, absent of external rewards or outcomes. (Venkatesh et al., 2003; Alomary et al., 2020)

*The Theory of Planned Behavior (TPB)* is a psychological theory that links intention to behavior. The theory includes *Perceived Behavioral Control*, which refers to an individual's perceived ability to control and direct their own behavior. TPB has been used to examine individuals' acceptance and use of different technologies. TPB and TAM have many similarities. (Venkatesh et al., 2003; Alomary et al., 2020)

*The Combined TAM-TPB (C-TAM-TPB)* is a hybrid model where *Perceived usefulness* from TAM is combined with TPB. (Venkatesh et al., 2003; Alomary et al., 2020)

*The Model of PC Utilization (MPCU)* refers to the idea that an individual's behavior is determined by what the individual wants to do (attitude), what the individual should do (social norm), what the individual usually does (habit), and the expected consequence of the individual's behavior. This theory is identified by six basic concepts: *Suitability, Complexity, Long-term consequences, Influence on usage, Social factors* and *Facilitating conditions*. (Venkatesh et al., 2003; Alomary et al., 2020)

*Innovation Diffusion Theory (IDT)* is a theory that is considered to be the permanent theory of innovation acceptance and can be used in both individual and organizational contexts. *Innovation* refers to new ideas, processes, or objects and *Diffusion* refers to the process of innovation being socially accepted. The theory consists of seven basic concepts: *Relative advantage, ease of use, image, visibility, compatibility, observability* and *voluntariness*. (Venkatesh et al., 2003; Alomary et al., 2020)

*Social Cognitive Theory (SCT)* is a theory that deals with human behavior. It is based on the idea that people's behavior is influenced by their personal beliefs, attitudes, and expectations, as well as by the social environment they are in. The five main concepts of SCT are: *expected outcomes, personal expected outcomes, self-efficacy, influence, and anxiety*. (Venkatesh et al., 2003; Alomary et al., 2020)
2.3 UTAUT in detail

In this subchapter we will go more in depth of the Unified Theory of Acceptance and Use of Technology (UTAUT). UTAUT is, as earlier mentioned, based on the four constructs: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Facilitating Conditions (FC). These constructs have a significant role in determining the User Behavior and acceptance of technology. The three constructs PE, EE and SI can be categorized as Behavioral Intention (BI). These four constructs are affected by four key moderators (Venkatesh et al., 2003)

2.3.1 Key moderators

The key moderators are:

- Gender
  - Based on cultural expectations and should not be associated with biological differences.
- Age
- Experience
- Voluntariness of use

These four key moderators affect the four constructs in different combinations and in different amounts.

2.3.2 Behavioral Intention and Use behavior

Behavioral Intention (BI) is affected by Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI). The three constructs sum up the Behavioral Intention that an individual has on accepting and using technology. (Venkatesh et al., 2003)
Use behavior (UB), is affected by BI and Facilitating Conditions (FC) and can be measured with how often an individual uses the technology. This is the output for the acceptance and use of technology. (Venkatesh et al., 2003)

2.3.3 Four constructs

Performance Expectancy (PE), is defined by Venkatesh et al (2003) as “the degree to which an individual believes that using the system will help him or her to attain gains in job performance”. This is affected by the key moderators gender and age. There are indications that this construct differs in age and gender where younger men has a stronger expectancy on performance.

Effort Expectancy (EE), is defined by Venkatesh et al (2003) as “the degree of ease associated with the use of the system”. This is being affected by the key moderators gender, age and experience. This construct is most significant in the early stages when using the technology.

Social Influence (SI), is defined by Venkatesh et al (2003) as “the degree of which an individual perceives that important others believe he or she should use the new system”, with system it can also mean technology. This is being affected by the key moderators gender, age, experience and voluntariness of use. This construct has a bigger effect on women, especially older women with low experience.

Facilitating Conditions (FC), is defined by Venkatesh et al (2003) as “the degree to which an individual believes that an organization and technical infrastructure exists to support use of the system”. This is being affected by the key moderators age and experience. More experienced users make higher demands on FC because they rely on its usage. Thus, they are less affected Behavioral Intention because they can rely on their experience. Older users may require more support, resources or help to use the technology and are more influenced by FC.

2.4 UTAUT Advantages, Relevance and Empirical Evidence

Unified Theory of Acceptance and Use of Technology (UTAUT) have been broadly used in many studies. The original UTAUT publication by Venkatesh and Davis in 2003 shows that UTAUT was able to explain 70% of the variance of Behavioral Intention, which also makes this model a great predictive tool.

Several technology acceptance models, such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology 2 (UTAUT2), have been widely applied in various studies (King & He, 2006; Alalwana, 2018). However in this study, the UTAUT model is chosen over these alternatives for several reasons. Firstly, while TAM primarily focuses on perceived usefulness and perceived ease of use, it does not consider the social aspect as extensively as UTAUT does. Given that ChatGPT has generated significant hype in society, it is crucial to incorporate the social factor into the analysis. UTAUT2, which was developed based on UTAUT, is similar to its predecessor but includes three additional constructs: Hedonic motivation, Price value, and Habit (Venkatesh, 2012). However, these extra constructs are deemed less relevant for this study, as ChatGPT is a relatively new technology with limited data available to address these three aspects.

Williams (2014) provides a literature review on 174 existing articles between 2004 to 2015 shows that UATUT is being used in various settings and technologies and has a high explanatory power when predicting technology acceptance and use (around 40-70%), which is higher than the eight models it is based on. Thus, the article's weight analysis shows that Performance Expectancy and Behavioral Intention shows extra good results in predictors of around 81 %. The article also shows that the most
explored limitation (key weakness) across the UTAUT studies is when studies are single subject focused because it can lead to bias samples.

While other relevant models could be employed for this research, UTAUT is deemed more suitable due to ChatGPT's relatively short public availability and the time constraints of this thesis. UTAUT can assess and explain factors that influence acceptance and use and by utilizing UTAUT, this study aims to provide a comprehensive analysis of ChatGPT's acceptance in the academic setting.
3. Methodology

This chapter details the methods and procedures employed in this study to examine the factors influencing students' acceptance and use of ChatGPT in academic settings. The chapter begins by explaining the research strategy, which incorporates a quantitative approach and the use of questionnaires. It then delves into the process of data collection, survey operationalization, and the construction of the questionnaire. This chapter presents a discussion of our data analysis methods, comprising data preparation, univariate-, bivariate-, and multivariate analysis. These rigorous analyses utilize statistical tests, including the Mann-Whitney U Test, Kruskal-Wallis H Test, and Spearman's Rank Correlation Coefficient and they are being done through the software tool Jamovi. Furthermore, the chapter also outlines the application of Partial Least Square-Structural Equation Modeling (PLS-SEM) for multivariate analysis and introduces the software tool SmartPLS used for this statistical analysis. It concludes by addressing the selection of respondents, the limitations of the study, and ethical considerations in conducting the research.

3.1 Research strategy

The objective of this study is to identify the factors that contribute to students' acceptance and usage of ChatGPT in academic settings. To analyze acceptance and usage, the Unified Theory of Acceptance and Use of Technology (UTAUT) is employed as the guiding theoretical framework. The ultimate aim is to generate valuable insights that can inform future implementation of ChatGPT within the academic sphere. In order to assess the constructs influencing students' acceptance and use of ChatGPT in academic settings, a quantitative research strategy will be undertaken. This approach enables the collection and analysis of large amounts of data from a substantial group of participants, facilitating the identification of patterns (Oates, 2006).

A survey method will be utilized to obtain standardized data in a systematic manner, as it is both time and resource-efficient, and well-suited for studies requiring extensive data collection and analysis (Oates, 2006). Surveys are also a popular and effective research method in the field of information systems (Oates, 2006). By employing this method, the study aims to provide a comprehensive understanding of the factors influencing students' acceptance and use of ChatGPT in academic contexts. For this study, a minimum of 60 respondents is required to conduct the Partial Least Squares Structural Equation Modeling (PLS-SEM), as discussed further in section 3.3.4.

3.2 Data gathering method

To conduct a survey study, questionnaires are being used for collecting data. Questionnaires consist of a series of questions or statements arranged in a predefined order. By using questionnaires the study saves time for data collection, the study saves time as multiple respondents can participate simultaneously. (Oates, 2006) It is crucial to exercise diligence in creating the questionnaire, as its quality directly impacts the overall study results (Peterson, 2000). Furthermore, during a survey study, there is no opportunity to revise the questionnaire since it is only distributed once (Oates, 2006).

3.2.1 Survey operationalization

The first step in the operationalization of the survey involved translating the abstract concept of students' perceptions of ChatGPT in academic settings into concrete, measurable elements (Bayer, 2021). This was achieved through a thorough review of previous research that employed
questionnaires and the UTAUT model. Drawing from successful methodologies in earlier studies can provide valuable insights and proven methods for our research. Although there were no previous studies focusing on students, ChatGPT, and academic settings concurrently, a study by Khatun et al. (2017) examining the acceptance of MyOnlineClinic, an internet-based hospital service, with UTAUT as the base for analysis. Additionally, the original UTAUT paper by Venkatesh et al. (2003) provided a blueprint for structuring questions that effectively gather data on each construct as well as Behavioral Intention (BI) and User Behavior (UB).

The next phase involved contextualizing each construct, BI, and UB, to better understand their relevance within the current educational context and their application with ChatGPT. This process involved comparing the questionnaires used by Khatun et al. (2017) and Venkatesh et al. (2003) to identify potential areas of alignment. It was crucial to adjust for the temporal context: for instance, the statement “I have the resources necessary to use the system” for the construct Facilitating Conditions in Venkatesh et al. (2003) study reflects a time when computer access was not as ubiquitous as it is now. This point will be elaborated in chapter 3.2.2, during the formulation of our questions and statements.

Thirdly, the limitations were taken into account of existing research, given their significant impact on the outcomes. Drawing from the analyses of Rudolph (2023), Azaria (2022), Aczel & Wagenmakers (n.d.) and ChatGPT's own definition of its limitations, we identified those constraints relevant to academic settings and incorporated them into the construct illustrated in 8.1.6 (Part 5: Constructs) in the Appendices.

Lastly, the method of data collection strictly follows the MRS Guidelines for Questionnaire Design (2014), which are established by a leading global association for individuals and organizations involved in market research, social research, opinion polling, business analytics, and customer insights. Compliance with these guidelines guarantees that the survey is executed in a manner that is ethical, dependable, and professional, while ensuring the highest level of quality and safeguarding the rights and privacy of the participants.

The MRS Guidelines stress the significance of voluntary informed consent, honesty, transparency, confidentiality, and the welfare of participants. By adhering to these principles, the study seeks to uphold the integrity of the research process, preserve the reputation of the profession, and confirm that the data collection procedure is conducted by individuals possessing suitable training, qualifications, and expertise. This approach guarantees the accurate design and implementation of the survey, reducing the likelihood of errors and bias and enhancing the validity and dependability of the results.

3.2.2 Questionnaire

The questionnaire begins by providing respondents with relevant information about the study's objectives, potential contributions, and subject matter to ensure they can give informed consent (MRS, 2014). It is also highlighted that responses will remain anonymous (Appendices 8.1.1).

The questionnaire's structure aims to collect the appropriate types of data for subsequent quantitative analysis, as various analyses require different data types. The collected data comprise:

- Nominal data to categorize demographic factors such as gender (Oates, 2006).
- Ordinal data to establish a range of responses (e.g., Disagree = 1, Neutral = 2, Agree = 3) (Oates, 2006).
- Ratio data to measure interrelated integers with a true zero point, like age (Oates, 2006).
Google Forms is being utilized to create the questionnaire due to its user-friendly interface, comprehensive features, and easy distribution to respondents. The questionnaire consists of six distinct sections, each with a randomized order of questions to mitigate response bias (MRS, 2014). Each section serves a unique purpose, presents different questions, and collects a specific type of data (details in Appendices 8.1):

1. Part one aims to identify the respondents' status, to make sure only students who have used ChatGPT in an academic setting get to answer the questionnaire. Data collected here distinguish between valid and invalid respondents. In case of a invalid respondent then they will be sent to part six directly.
2. Part two gathers key moderating data such as gender, age, experience, and voluntariness of use.
3. Part three collects additional information, including 'level of education' and 'prime subject area,' to map student usage patterns. Though not directly tied to UTAUT, these questions provide context-relevant insights. E.g., the question, 'Where did you first hear about ChatGPT?' provides understanding about respondents' exposure to ChatGPT.
4. Part four comprises questions relevant to Behavioral Intention (BI) and User Behavior, mapping these using both nominal and ordinal data. While Khatun et al. (2017) and Venkatesh et al. (2003) provided inspiration, new questions were crafted to suit this study's unique context.
5. Part five collects ordinal data and statements about the constructs, each of which has four statements, including one about ChatGPT's limitations. Most of these statements have been modified from those used by Khatun et al. (2017) and Venkatesh et al. (2003) to better fit the present context.
6. Part six is an open question and invites the respondents to provide any additional information.

The design and development of the questionnaire was reviewed by the thesis supervisor. Thus, to ensure the collection of high-quality data, a pilot test was conducted on the questionnaire (MRS, 2014). The test involved 10 respondents, eight of whom were part of the target audience, while two were not and their responses were deemed invalid. Feedback from the pilot test participants, including constructive criticism and their overall perception of the questionnaire, contributed to enhancing the questionnaire's quality.

3.2.3 Selections of respondents

Considering the time constraints and the undefined population parameter inherent in this research, the target is set at approximately 100 respondents. Achieving this target necessitates a more effective strategy for selecting respondents. To resolve this challenge, a non-probability sampling approach have been chosen, which includes:

1. Purposive Sampling, also known as Judgmental Sampling: Leveraging the principles of this sampling methodology, the research questionnaire was disseminated to a set of participants. These individuals were hand-picked owing to certain unique attributes and experiences they possessed, thus aligning them precisely with the research focus (Nikolopoulou, 2022). The survey was systematically circulated across numerous social media platforms, including but not limited to Facebook, Reddit, and LinkedIn, to groups that demonstrated direct relevance to the areas of ChatGPT, academic pursuits, and scholarly environments.

Participants of online groups such as "ChatGPT & OpenAI", "ChatGPT for Business & Life", "Study Techniques And Motivation", and "ChatGPT Community" were deemed particularly
suitable for this research. These individuals, by virtue of their engagement with these groups, likely possess an informed perspective on ChatGPT, its potential applications, and its effectiveness in academic and professional settings. This knowledge and familiarity with the subject make them valuable contributors to the study. Their perspectives can provide unique insights and depth of understanding that aligns well with the research's central objectives. Furthermore, their active participation in these focused groups suggests a willingness to engage in dialogues about ChatGPT, thereby increasing the likelihood of them responding to the research survey in a meaningful way.

2. Convenience Sampling: The questionnaire was also circulated within the author's personal network on Facebook and LinkedIn, offering friends and family the opportunity to participate in the survey.

3. Snowball Sampling: Each time the survey was posted on social media, a request was made to share the survey further, employing a snowball sampling technique.

It is important to note that this approach may increase the potential for research bias as each unit of the target population does not have an equal chance of being included (Nikolopoulou, 2022). In addition, the use of convenience sampling can introduce challenges due to the potential for biased responses or inaccurate reporting (Oates, 2006). These potential issues underscore the necessity to follow the guidelines established by the Market Research Society (MRS) for questionnaire design. Adhering to these guidelines can mitigate the risk of bias, ensuring that the collected data more accurately reflects the authentic perspectives and experiences of the target population.

3.3 Data analysis method

In this chapter, the data analysis methodology is detailed, encompassing the data preparation as well as the univariate, bivariate, and multivariate analyses. The processes for data coding are outlined, statistical tests are employed to reveal correlations, and Structural Equation Modeling is leveraged to investigate relationships between variables.

3.3.1 Data preparation

To effectively utilize the data collected from the survey, it first needs to be systematically organized. Each question and statement will be assigned unique identifiers to facilitate their separation and analysis. These identifiers will be structured based on the part and question/statement they correspond to. For instance, 'PE1' is the first statement of Performance Expectancy: "I believe using ChatGPT has helped me to attain gains in educational performance."

The responses to each question/statement will also be coded, with the coding method differing based on the type of data — nominal, ordinal, or ratio. For nominal data, a numerical value will be assigned to represent each category of the demographic factor. For ordinal data, a number will be assigned to quantify the answer. For example, in the case of statement 'PE1,' "Strongly agree" will be represented by a 5, while "Strongly disagree" will be represented by a 0. In the case of statements that involve negation, such as the limitation statement 'PE4,' the coding will be reversed, this will also be mentioned in chapter 4. A more detailed depiction of this coding strategy is provided in Table 8.2.

3.3.2 Univariate Analysis

To develop an initial understanding of the data, a univariate analysis is employed, which involves examining one variable at a time (Huberty and Morris, 1989). This analysis is focusing on different constructs, Behavioral Intention, and User Behavior. Such a review allows us to identify any
anomalies within the data (Dahmström, 2005). The statistical software Jamovi will be used to conduct this analysis.

3.3.3 Bivariate Analysis

In order to establish statistically significant correlations within the research, a bivariate analysis will be undertaken. This method analyzes multiple variables simultaneously to discover potential correlations between them (Reitsma, 2005). The analytical procedures utilized will include the Mann-Whitney U Test, Kruskal-Wallis H Test, and Spearman’s Rank Correlation Coefficient, with all analyses being conducted through Jamovi.

Mann-Whitney U Test

The Mann-Whitney U Test is ideal for comparing two independent groups, for instance, gender when the only categories are men and women. The data for this test does not need to be normally distributed. The test generates a p-value ranging from -1 to 1; negative values suggest a negative correlation, while positive values indicate a positive correlation (Gunnarsson, 2002). In this study, a p-value below 0.05 will be considered a significant difference between two groups.

Kruskal-Wallis H Test

When analyzing more than two groups, then the Kruskal-Wallis H Test will be used. This test, which can be viewed as an extension of the Mann-Whitney U Test, allows for a comparison of multiple independent groups and does not require the data to be normally distributed. However, it’s important to note that this test can identify a significant difference but does not specify which group is significantly different (LærdStatistics, 2018). As in Mann-Whitney U Test, a p-value below 0.05 will be considered a significant difference.

Spearman’s Rank Correlation Coefficient

Spearman’s Rank Correlation Coefficient is employed to measure the strength and direction of association between two ranked variables. The test uses ordinal, interval, or ratio data to determine the monotonic relationship (figure 2) between the two variables (LærdStatistics, 2018). The output, called the coefficient, can range from -1 to 1. A positive result (0<result<1) indicates a direct relationship between the variables (as one increases, so does the other), while a negative result (0>result>-1) indicates an inverse relationship (as one increases, the other decreases). For a relationship to be classified as arbitrary, a value between 0.3 and 0.7 must be obtained. (Oates, 2006)
3.3.4 Multivariat variant

To simultaneously analyze multiple variables, we will apply Structural Equation Modeling (SEM). SEM is particularly valuable for visualizing the relationships between variables and can be effectively applied to UTAUT framework. Moreover, we will use Partial Least Square-Structural Equation Modeling (PLS-SEM), as it is better suited for small sample sizes and is widely used in sciences, marketing, organizational studies, management information systems, and business strategy. However, it is essential to note that PLS-SEM also has limitations. For instance, high-valued structural path coefficients are required for smaller sample sizes, and smaller samples may lead to errors and biased results. (Wong, 2016)
The SEM diagram, representing the result of this multivariate analysis, comprises two sub-models:

1. Inner model represents the relationship between independent- and dependent variables.
2. Outer model represents the relationship between latent variables and their observed indicators. The diagram also features exogenous latent variables, which have arrows pointing outwards but none leading to them, and endogenous latent variables, which have one or many arrows leading to them and represent the effects of variables. (Wong, 2016)

Bootstrapping will also be used to make sure the data is normally distributed, in other words test whether the outer weights, outer loading and path coefficients are significant (SmartPLS, n.d.). As in Mann-Whitney U Test and Kruskal-Wallis H, a p-value of 0.05 will be used as the measurement but here a value under 0.05 will be considered significant. The value will be represented in the arrow inside of parentheses.
There are three results from a PLS-SEM model which we are interested in:

1. **The coefficient of determination of the endogenous latent variables**, represented as the number in the blue circles of CXSAT and LOYAL in figure 4. This number indicates how much an endogenous variable can be explained by its latent variables (e.g. the three latent variables QUAL, CXSAT and EXPECT moderately explains 57.2% of LOYAL). (Wong, 2016) According to Hair et al. (2014) a value for a large effect is 0.35 and that will be the acceptable value for this research.

2. **The inner model coefficient size and significance** represented by the number on the arrow (Wong, 2016). The example in figure 4 shows that QUAL to LOYAL has a statistically significant 0.352 (35.2%) and EXPECT to LOYAO does not have a statistical significance because the value 0.003 (0.3%).

3. **The outer model reliability value** shows how strong the relationship is between the variable and the indicator (e.g., qual_1 and QUAL have a value of 0.881 in figure 4). In general a 0.7 score or higher is preferred but in the case a study with a smaller sample then 0.4 and higher is accepted. (Wong, 2016)

In the multivariate analysis of PLS-SEM, all components of the model are interrelated. This means that suboptimal or insignificant values can affect the entire model's results. Despite the initial UTAUT model and corresponding survey questions being formulated based on prior research, it is possible to encounter insignificant results. This situation necessitates the removal of certain indicators or variables. Criteria for removal might include a low reliability score, an Average Variance Extracted (AVE) value below 0.5, or a path coefficient (represented by the number on the arrow in a SEM diagram) that is less than 0.3 and statistically insignificant. However, it is crucial not to base these decisions solely on statistical outcomes. The theoretical relevance of each indicator or variable should be considered alongside its statistical performance to ensure that the resulting model maintains its conceptual validity. (Hair et al., 2021)

For this analysis, we will use SmartPLS, a well-regarded software application for PLS-SEM, known for its cost-effectiveness and time efficiency (Wong, 2016). In the context of this paper's theoretical
framework, UTAUT, Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions will be represented as exogenous latent variables, while Behavioral Intention and User Behavior will be presented as dependent variables. The constructs and their corresponding questions will represent the outer model, meaning that only ordinal data related to these constructs will be used in this chapter (refer to Appendix 9.1 for details). The number of respondents needed to utilize PLS-SEM is 10 respondents per construct, so in this case of UTAUT at least 60 respondents are needed and that will be the aim of least respondents. (Wong, 2016)

3.4 Method reflection

This chapter will offer a reflexive account of the methods employed in this research study, examining both their strengths and potential limitations. It will critically assess the validity, reliability, and quality of the study, as well as discuss potential sources of error and interpretive ambiguities that could potentially affect the conclusions.

Beginning with the data preparation process, assigning unique identifiers and coding responses indeed facilitated the organization and analysis of the survey data. However, there are a few potential limitations to consider. Firstly, the coding system relies heavily on the accuracy and consistency of the participant's responses. Any human error could lead to coding errors, potentially skewing the results. Also, coding assumes a linear relationship in ordinal data (e.g., the distance between "Strongly agree" and "Agree" is the same as between "Agree" and "Neutral"). This might not truly reflect the subjective perceptions of respondents, potentially introducing systematic errors.

The tests employed, namely the Mann-Whitney U Test, Kruskal-Wallis H Test, and Spearman’s Rank Correlation Coefficient, are appropriate considering the nature of the data. Nevertheless, these non-parametric tests are less powerful than their parametric counterparts when the data is normally distributed, implying a possible trade-off for their versatility (Frost, 2023).

Partial Least Square-Structural Equation Modeling (PLS-SEM) enables an understanding of complex relationships among multiple variables. Worth mentioning is that this method is correlational and identifies associations and does not establish definitive causality. Furthermore, PLS-SEM's requirement for high-valued structural path coefficients for smaller samples could lead to errors and biased results, affecting the overall reliability and validity of the model. (Hair et al., 2021) Lastly, in the case of insignificant results, removal of indicators or variables is needed. This removal enhances the statistical performance but might compromise the theoretical relevance of UTAUT (Hair et al., 2021).

In conclusion, while the methods employed in this study are rigorous and appropriate, these reflections highlight the inherent complexities of behavioral research. It is important to interpret the results with an understanding of these limitations, recognizing that the strength of this research lies in its holistic approach to examining the relationships among the variables of interest, rather than in providing absolute or causal explanations.
3.5 Ethics

Ethics in research represents a fundamental aspect to maintain the trust, integrity, and validity of the study. This research was carried out with a conscious effort to ensure adherence to ethical guidelines and consideration of potential ethical issues.

In this study, the participants were clearly informed about the purpose of the research and the type of data to be collected through an introduction at the beginning of the survey (as detailed in 8.2). This ensured that the participants were aware of the intention behind collecting the data and could make an informed decision about their participation (MRS, 2014). It is important to emphasize that their participation was completely voluntary, they had the option to skip a question and they had the right to withdraw at any point without providing a reason. Furthermore, as for privacy and confidentiality, participants were reassured that their data would be treated with utmost care and kept confidential. No identifying personal details were requested, and no data could be traced back to specific individuals. The coding and labeling of the data during the data preparation phase (refer to 3.3.1) were designed to further protect participant anonymity.

Bias represents a critical ethical issue in research. To mitigate this, an objective stance was maintained throughout data collection and analysis. However, there remains a risk of unconscious bias. For instance, during data interpretation, personal experiences and perspectives could inadvertently influence understanding. To minimize potential bias, a clear and systematic research approach was established and strictly following the MRS Guidelines for Questionnaire Design (2014).

In conclusion, throughout this research, a strong commitment has been maintained to uphold high ethical standards. Efforts have been made to be transparent, respectful, and fair in dealings with participants, data, and the application of chosen methodologies. The research has also been reviewed by the supervisor and peer student throughout all steps.
4. Results

This chapter encapsulates the results derived from the detailed methodology chapter. Firstly, it lays out the findings in a structured manner, starting with general results, including demographic information, voluntariness of use, and experience with ChatGPT. It then delves into the analysis of responses to each construct: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), and Facilitating Conditions (FC). Each construct is examined through a univariate analysis, with the responses illustrated via bar plots and tables.

Secondly, the chapter presents the bivariate analysis, deploying the Mann-Whitney U Test for analyzing gender and age differences and the Kruskal-Wallis H Test for examining experience and voluntariness of use. Additionally, this section applies Spearman's Rank Correlation Coefficient to determine the strength and direction of association between two ranked variables.

The final section of this chapter ventures into a multivariate analysis, detailing the model specification and final results with the aid of Partial Least Squares Structural Equation Modeling (PLS-SEM). Modifications in the model, the rationale behind them, and the resultant refined model are carefully expounded.

4.1 General result

The total number of 87 persons completed the questionnaire. Out of these 17 were "invalid respondents". This is because they did not study or had never used ChatGPT. Survey answers and data from the remaining 70 "valid respondents" will be used in this analysis. Since ChatGPT is being developed rapidly it is important to note that the first respondent completed the survey on 2023-03-23 and the last respondent completed the survey on 2023-05-02.

![Figure 5. Respondents current level of education](image-url)
Figures 5-7 presented provide valuable insight into the demographic and usage habits of the participating students. Figure 5 reveals that a significant proportion of respondents are studying at the bachelor's and master's level, suggesting a higher level of academic engagement. In terms of fields of study, Figure 6 indicates a majority of students are primarily pursuing information technology and data science.

Moreover, Figure 7 outlines the various ways students utilize ChatGPT to bolster their studies. Four methods emerge: research and learning, writing or editing text, creativity, inspiration or brainstorming, and problem-solving in areas such as mathematics, programming, and other scientific domains (refer to section 8.1.5 for more detail). One response, “använt”, is marked as invalid due to its lack of context or meaningful information. Taken together, these figures shed light on the nuanced profiles of the respondents and their interaction with ChatGPT in academic contexts, offering a richer understanding of the dynamics at play.
4.1.1 Key moderators result

**Age and gender**
The mean age is 25.2 years old and the age difference between the maximum- and minimum age is 36 years. To be able to use Mann-Whitney U Test the age moderator will be divided into two groups. The reason for only two groups is because the groups would become too small if it was divided further. We ended up with one group of 15-24 with 43 respondents (61%) and one group of 25-51 with 27 respondents (39%).
The result also shows that 46 men and 24 women answered the survey. The mean age of each group is 25.8 for men and 24.3 for women.

<table>
<thead>
<tr>
<th>What is your gender?</th>
<th>How old are you?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man</td>
<td>46</td>
</tr>
<tr>
<td>Women</td>
<td>24</td>
</tr>
<tr>
<td>Missing</td>
<td>0</td>
</tr>
<tr>
<td>Man</td>
<td>0</td>
</tr>
<tr>
<td>Women</td>
<td>0</td>
</tr>
</tbody>
</table>

*Figure 8. Age and Gender description from Jamovie*

**Voluntariness of use**
The result of the question, if the students feel like using ChatGPT for studies is a voluntary choice, was that none of the respondents answered that it was in some way mandatory. The answer given was:

- *Completely voluntary:* 51 respondents (73%)
- *Mostly voluntary:* 12 respondents (17%)
- *Neutral:* 7 respondents (10%)

The result indicates that students' usage of ChatGPT in academic settings is more or less voluntary.

**Experience**
With the release of ChatGPT in 30th November 2022 the result from the respondents of how long they have been using ChatGPT in their studies was:

- *One month or less*:
  - 18 respondents (25%)
- *More than one month*:
  - 12 respondents (17%)
- *More than two months*:
  - 20 respondents (29%)
- *More than three months*:
  - 20 respondents (29%)

4.2 Univariat analys

This chapter will go through the constructs and how the respondents answer to every construct. The question is written in short, e.g. PE1 stands for question number 1 in Performance Expectancy and you can find that question in appendices 8.1. All the questions with a 4 (PE4, EE4, SI4 and FC4) are about limitations and are negative. To make them relevant in this chapter I have then coded them
conversely so that they suit the other question. This chapter will use a bar plot for visualization and a table for more exact measurements.

4.2.1 Performance Expectancy (PE)

*Figure 9. Performance Expectancy bar chart from Jamovi*
Here we can see that a big majority of respondents answered 3 (partly agree) and four (strongly agree) on all statements except for PE4, where we have a more spread out answer. In PE2 and PE3 no respondents answered 0 (strongly disagree). Only one respondent per question did not know what to answer (answered: 9999) on the Performance Expectancy statements.

Table 1. Performance Expectancy response result

<table>
<thead>
<tr>
<th>Question</th>
<th>0 (Strongly disagree)</th>
<th>1 (Partly disagree)</th>
<th>2 (Neutral)</th>
<th>3 (Partly agree)</th>
<th>4 (Strongly agree)</th>
<th>9999 (do not know)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PE1</td>
<td>1</td>
<td>3</td>
<td>12</td>
<td>37</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>PE2</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>33</td>
<td>32</td>
<td>1</td>
</tr>
<tr>
<td>PE3</td>
<td>0</td>
<td>4</td>
<td>13</td>
<td>26</td>
<td>26</td>
<td>1</td>
</tr>
<tr>
<td>PE4 (reversed)</td>
<td>3</td>
<td>17</td>
<td>13</td>
<td>18</td>
<td>18</td>
<td>1</td>
</tr>
</tbody>
</table>
4.2.2 Effort Expectancy (EE)

Figure 10. Effort Expectancy bar chart from Jamovi
<table>
<thead>
<tr>
<th>Question</th>
<th>0 (Strongly disagree)</th>
<th>1 (Partly disagree)</th>
<th>2 (Neutral)</th>
<th>3 (Partly agree)</th>
<th>4 (Strongly agree)</th>
<th>9999 (do not know)</th>
</tr>
</thead>
<tbody>
<tr>
<td>EE1</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>28</td>
<td>36</td>
<td>0</td>
</tr>
<tr>
<td>EE2</td>
<td>0</td>
<td>1</td>
<td>8</td>
<td>27</td>
<td>33</td>
<td>1</td>
</tr>
<tr>
<td>EE3</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>32</td>
<td>25</td>
<td>2</td>
</tr>
<tr>
<td>EE4 (reversed)</td>
<td>7</td>
<td>21</td>
<td>18</td>
<td>13</td>
<td>9</td>
<td>2</td>
</tr>
</tbody>
</table>

*Table 2. Effort Expectancy response result*

Here we also have a big majority answering 3 (partly agree) and 4 (strongly agree) on the statements except on EE4, which is the one about the limitations. On question EE1 and EE2 no respondents answer 0 (strongly disagree).

4.2.3 Social Influence (SI)
Figure 11. Social Influence bar chart from Jamovi

<table>
<thead>
<tr>
<th>Question</th>
<th>0 (Strongly disagree)</th>
<th>1 (Partly disagree)</th>
<th>2 (Neutral)</th>
<th>3 (Partly agree)</th>
<th>4 (Strongly agree)</th>
<th>9999 (do not know)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SI1</td>
<td>7</td>
<td>7</td>
<td>24</td>
<td>16</td>
<td>5</td>
<td>11</td>
</tr>
<tr>
<td>SI2</td>
<td>7</td>
<td>8</td>
<td>18</td>
<td>21</td>
<td>16</td>
<td>0</td>
</tr>
<tr>
<td>SI3</td>
<td>5</td>
<td>8</td>
<td>24</td>
<td>24</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>SI4 (reversed)</td>
<td>3</td>
<td>14</td>
<td>16</td>
<td>18</td>
<td>15</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 3. Social Influence response result

Here the respondent mostly answered 2 (neutral) and 3 (partly agree). There is no statement that has a majority of not agreeing or agreeing to a statement. In other words the respondents do not really have an opinion about Social Influence. Also worth noticing is that 11 respondents did not know what to answer on SI1.
4.2.4 Facilitating Conditions (FC)

Figure 12. Facilitating Conditions bar chart from Jamovi

<table>
<thead>
<tr>
<th>Question</th>
<th>0 (Strongly disagree)</th>
<th>1 (Partly disagree)</th>
<th>2 (Neutral)</th>
<th>3 (Partly agree)</th>
<th>4 (Strongly agree)</th>
<th>9999 (do not know)</th>
</tr>
</thead>
<tbody>
<tr>
<td>FC1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>21</td>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td>FC2</td>
<td>1</td>
<td>6</td>
<td>14</td>
<td>28</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>FC3</td>
<td>4</td>
<td>9</td>
<td>16</td>
<td>16</td>
<td>7</td>
<td>18</td>
</tr>
<tr>
<td>FC4 (reversed)</td>
<td>3</td>
<td>17</td>
<td>18</td>
<td>14</td>
<td>16</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 4. Facilitating Condition response result

This construct has also spread answers except FC1 which a majority answered 4 (strongly agree). FC2-4 indicates that the respondents do not have a clear opinion on the Facilitating Conditions. FC3 has a big number of respondents answering that they do not know what to answer on that statement.
4.3 Bivariate analysis

This chapter employs the Mann-Whitney U test and Kruskal-Wallis H test, examining gender, age, experience, and voluntariness of use in relation to ChatGPT's usability. As chapter 2.3 shows that all key moderators do not affect all constructs. This will be dealt with in this part.

4.3.1 Mann-Whitney U Test

Gender

The result in table 5 shows that there is no significant difference between men and women in any of the construction statements that is relevant to gender.

<table>
<thead>
<tr>
<th>Statement number:</th>
<th>PE</th>
<th>EE</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.303</td>
<td>0.689</td>
<td>0.662</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>0.75</td>
<td>0.318</td>
</tr>
<tr>
<td>3</td>
<td>0.463</td>
<td>0.878</td>
<td>0.103</td>
</tr>
<tr>
<td>4 (reversed)</td>
<td>0.367</td>
<td>0.240</td>
<td>0.208</td>
</tr>
</tbody>
</table>

*Table 5. Mann-Whitney U Test for Gender*

Age

Table 6 shows that there is no significant difference except for three statements:

1. PE4: “The *limitations* of ChatGPT have negatively affected your performance when using it for studies.”
2. EE4: “The *limitations* of ChatGPT have negatively impacted its ease of use during study sessions.”
3. FC1: “I have the resources necessary to use ChatGPT.”

<table>
<thead>
<tr>
<th>Statement number:</th>
<th>PE</th>
<th>EE</th>
<th>SI</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.783</td>
<td>0.593</td>
<td>0.607</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>2</td>
<td>0.176</td>
<td>0.912</td>
<td>0.172</td>
<td>0.808</td>
</tr>
<tr>
<td>3</td>
<td>0.578</td>
<td>0.656</td>
<td>0.889</td>
<td>0.542</td>
</tr>
<tr>
<td>4 (reversed)</td>
<td>0.015</td>
<td>0.05</td>
<td>0.995</td>
<td>0.329</td>
</tr>
</tbody>
</table>

*Table 6. Mann-Whitney U Test for Age*
4.3.2 Kruskal-Wallis H Test

**Experience**

As experience consists of four different levels of experience the Kruskal-Wallis H Test was reasonable to use. The results in table 7 show that there are only three statements that have a statistically significant difference lower than 0.05. These statements are:

1. EE2: “I have felt relaxed when using ChatGPT for studies, due to its ease of use.”
2. EE3: “I felt productive when using ChatGPT for studying.”
3. SI1: “People who are important to me believe that I should use ChatGPT for educational purpose.”

<table>
<thead>
<tr>
<th>Statement number:</th>
<th>EE</th>
<th>SI</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>0.279</td>
<td>0.018</td>
<td>0.072</td>
</tr>
<tr>
<td>2:</td>
<td>0.045</td>
<td>0.971</td>
<td>0.081</td>
</tr>
<tr>
<td>3:</td>
<td>0.003</td>
<td>0.443</td>
<td>0.231</td>
</tr>
<tr>
<td>4 (reversed):</td>
<td>0.122</td>
<td>0.403</td>
<td>0.639</td>
</tr>
</tbody>
</table>

*Table 7. Kruskal-Wallis H Test for Experience*
Voluntariness of Use

This moderation key does also have several levels of voluntariness of use which makes Kruskal-Wallis H Test suitable for analyzing statistically significant differences. Voluntariness of use only affects Social Influence and none of the statements shows a statistically significant difference.

<table>
<thead>
<tr>
<th>Statement number:</th>
<th>SI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:</td>
<td>0.818</td>
</tr>
<tr>
<td>2:</td>
<td>0.923</td>
</tr>
<tr>
<td>3:</td>
<td>0.916</td>
</tr>
<tr>
<td>4 (reversed):</td>
<td>0.835</td>
</tr>
</tbody>
</table>

*Table 8. Kruskal-Wallis H Test for Voluntariness of Use*

4.3.3 Spearman's Rank Correlation Coefficient

Applying Spearman's Rank Correlation Coefficient, to determine the strength and direction of association between two ranked variables, reveals correlations among the statements. The statement number 4 in correlation to the other statements shows that it is not classified as arbitrary except between SI4 and SI1. Worth to mention is that these were the statements about limitations. All the statements from Facilitating Conditions also show a result of not being classified as having an arbitrary correlation. Also worth mentioning is the correlation between SI1 with SI2 and SI2 with SI3 has a low value under 0.3. This result will correlate with the PLS-SEM in 5.4

4.4 Multivariate Analysis

As depicted in figure 15, the final PLS-SEM model did not retain all initially proposed indicators and variables. Certain aspects of the model required adjustment due to unsatisfactory performance. For instance, the Facilitating Conditions construct did not yield significant results and thus was not considered relevant to the model. This construct met all criteria for removal—low outer model reliability value, below-threshold AVE value, and an insignificant path coefficient—indicating its exclusion was justified.
Moreover, the fourth statement related to limitations within the outer model consistently scored poorly across all constructs. Given its suboptimal performance in prior analyses and the likelihood of its lacking theoretical relevance in this specific context, this statement was removed. The indicator SI2 was also removed due to its low value.

4.4.2 Results

The final, redefined model is presented in figure 15. Despite its low coefficient significance and high bootstrapping value (0.218, which exceeds the 0.05 threshold), the Social Influence (SI) construct was retained. This decision was informed by its high AVE of 0.708 and strong relationships with SI1 and SI3. While SI may not demonstrate a significant direct impact on Behavioral Intention (BI), it retains theoretical relevance within the model. In contrast, Facilitating Conditions (FC) was removed entirely due to its lack of relevance and potentially poorly framed indicators within the context of student and academic settings.

As shown in the inner model of figure 15, both Performance Expectancy and Effort Expectancy significantly influence BI, explaining 30.2% and 35.6% of its variance, respectively. Together with SI, these constructs account for 45.8% of the variation in BI. In turn, BI has a significant effect on User Behavior (UB), explaining 62.5% of its variance and contributing to 39.1% of the total variation in UB. The outer model reliability values demonstrate robust relationships between all retained variables and indicators.

![Figure 15. PLS-SEM redefined model](image-url)
5. Analysis

This chapter dives into a comprehensive exploration of the data collected from the survey. The chapter commences with an overview of the response demographics and proceeds to investigate the influence of key moderators such as gender, age, experience, and voluntariness of use, applying the Mann-Whitney U Test and Kruskal-Wallis H Test to reveal significant differences.

The chapter further delves into a thorough analysis of the constructs - Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions - underpinning users' attitudes and behaviors towards ChatGPT. A univariate analysis of these constructs provides insight into their influence on Behavioral Intention, while Spearman's Rank Correlation Coefficient analysis is employed to determine the monotonic relationship between the constructs' individual components.

The final part of the chapter executes a Partial Least Squares Structural Equation Modeling (PLS-SEM) analysis, examining both the outer- and inner models. The outer model focuses on the relationships between the latent variables and their respective indicators, while the inner model assesses the interactions between exogenous and endogenous latent variables.

5.1 Data Overview

The survey garnered responses from 87 individuals, with 70 of these responses deemed valid for analysis. This amount of gathered respondents was enough for usage of PLS-SEM but did not meet the goal of 100 respondents. The sample was split into two gender groups comprising 46 men and 24 women. Age was also divided into two categories: 43 respondents aged 15-24 and 27 respondents aged 25-51.

5.2 Key Moderators

In this chapter, the results of the analysis from Mann-Whitney U Test and Kruskal-Wallis H Test related to key moderators will be discussed with a particular emphasis on statistically significant differences.

5.2.1 Gender

Gender was found to influence the constructs of Performance Expectancy, Effort Expectancy, and Social Influence. However, no significant differences were observed between the independent gender groups. This suggests that within the context of academic use of ChatGPT, gender does not play a significant role.

5.2.2 Age

Age affected all four constructs, with the bivariate analysis revealing statistically significant differences in three statements, each with a p-value below 0.05. These statements included PE4: “The limitations of ChatGPT have negatively affected your performance when using it for studies,” EE4: “The limitations of ChatGPT have negatively impacted its ease of use during study sessions,” and FC1: “I have the resources necessary to use ChatGPT.” (refer to figure 13 for more details). The results indicate that the younger age group is less affected by Performance and Effort Expectancy limitations than the older age group. Furthermore, the younger participants perceived themselves as
having the necessary resources to use ChatGPT. Beyond these statements, no significant age-related differences were observed across the four constructs.

5.2.3 Experience
Experience influenced Effort Expectancy, Social Influence, and Facilitating Conditions. The bivariate analysis identified statistically significant differences in three statements: EE2: “I have felt relaxed when using ChatGPT for studies, due to its ease of use,” EE3: “I felt productive when using ChatGPT for studying,” and SI1: “People who are important to me believe that I should use ChatGPT for educational purposes.” Although the Kruskal-Wallis H Test does not specify which level of experience influenced these outcomes, it can be inferred that at least one experience level from the sample differed significantly in its central tendency from the others.

5.2.4 Voluntariness of Use
Voluntariness of Use affected only Social Influence, and the results demonstrated no significant difference between the various levels of voluntariness. This suggests that all independent samples display similar central tendencies regarding this key moderator.

5.3 Constructs
This chapter will delve into an in-depth evaluation of the constructs: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), and Facilitating Conditions (FC). Univariate analysis is conducted, and Spearman's Rank Correlation Coefficient is applied to interpret their relevance and correlation within the context of students ChatGPT usage in academic settings.

5.3.1 Univariate Analysis
Performance Expectancy (PE) consists of the following four statements:

PE1: I believe using ChatGPT has helped me to attain gains in educational performance.
PE2: I have found ChatGPT useful for educational purposes.
PE3: Using ChatGPT has enabled me to acquire academic knowledge.
PE4: The limitations of ChatGPT, have negatively affected your performance when using it for studies. E.g., limitation: misinformation, not understanding context, bias, translation, inability of answer complex math or similar.

The univariate analysis reveals that PE generally exhibits a positive influence on Behavioral Intention (BI). Statements PE1-3 particularly suggest that a significant proportion of participants agree or strongly agree that ChatGPT contributes to performance gains, with no respondents strongly disagreeing with PE2 and PE3. Statement PE4 presents a more diverse range of responses. The limited "do not know" responses lend credibility to the findings, indicating that PE is a crucial factor in students' acceptance and use of ChatGPT in academic settings.

Effort Expectancy (EE) consists of the following four statements:
EE1: It has been easy to use ChatGPT for studies.
EE2: I have felt relaxed when using ChatGPT for studies, due to its ease of use.
EE3: I felt productive when using ChatGPT for studying.
EE follows a similar pattern to PE. The results show that EE generally has a positive impact on Behavioral Intention, with a majority of responses ranging from partly to strongly agree that ChatGPT is easy to use. Statement EE4 reveals that a significant number of respondents agree that the limitations of ChatGPT negatively impact EE. Notably, no respondents strongly disagreed with EE1 and EE2, and the low number of "do not know" responses suggests that EE could positively influence students' acceptance and usage of ChatGPT, though its limitations could pose a challenge.

Social Influence (SI) consists of the following four statements:

**SI1**: People who are important to me believe that I should use ChatGPT for educational purpose. E.g., important people: family, skilled student, close friends, teacher.

**SI2**: Other students have influenced me to use ChatGPT.

**SI3**: In general, I have felt influenced to use ChatGPT to acquire knowledge in my educational subject.

**SI4**: People who influence my behavior do not think that I should use ChatGPT for studying due to its limitations. E.g., limitations: inaccurate information, wrong language translation, wrong answers, bias, not useful.

The results show a range of neutral to partly agree responses to SI statements. Statement SI4 presents a diverse set of answers, indicating uncertainty regarding the social effects of ChatGPT's limitations. The notable count of "do not know" responses to SI1 suggests that some respondents struggled to answer this statement. These results could imply that the respondents are either unsure or neutral about the influence of others on their acceptance and use of ChatGPT in academic settings.

Facilitating Conditions (FC) consists of the following four statements:

**FC1**: I have the resources necessary to use ChatGPT. E.g. online resources and knowledge.

**FC2**: The more I used ChatGPT for my studies, the greater the opportunities for continuous development and acquiring new knowledge in my subject area.

**FC3**: If I encounter difficulties with ChatGPT then there is a way for me to get assistance or help. E.g., Online-support, How-to-instructions, ask ChatGPT itself, or others.

**FC4**: The technical limitations of ChatGPT have negatively impacted my studying experience. E.g., technical limitation: computing power, response time, short answers, data collection, bias, users may wait in a queue due to limited simultaneous access to ChatGPT.

The responses to FC were more varied. The majority of respondents agreed or strongly agreed with FC1, indicating they possess the necessary resources to use ChatGPT, with no disagreement noted. The responses for FC2-4 were spread out. Of note, FC3 had 18 "do not know" responses, which may undermine the relevance of this statement. Given the varied responses, it is challenging to draw a definitive conclusion about FC. The results suggest that FC might not be a crucial factor in researching student acceptance and use of ChatGPT in academic settings. This ambiguity indicates a need for further investigation into the role of Facilitating Conditions in this context.
5.3.2 Spearman's Rank Correlation Coefficient

The results from Spearman's rank correlation coefficient analysis suggest that the majority of the constructs did not demonstrate an acceptable correlation, specifically within the range of 0.3 to 0.7. This lack of correlation is most notable in the construct of Facilitating Conditions (FC), where none of its relationships yielded an acceptable result. It was observed that the questions related to limitations generally resulted in correlations below 0.3, with the exception of the relationship between SI4 and SI1. The absence of acceptable correlations may suggest that certain constructs or statements are less relevant or useful in this context, or it may indicate the influence of other, unaccounted factors. Only a select few relationships met the acceptable criteria: PE1-2, PE1-3, PE2-3, EE1-2, EE1-3, EE2-3, SI1-3, and SI1-4. These results reveal the existence of a monotonic relationship between the respective statements within these pairs.

5.4 PLS-SEM

This chapter presents a detailed analysis of the outer and inner models. The discussion will center around the relationships between latent variables and their indicators, as well as the influence of exogenous variables on endogenous ones.

5.4.1 Outer model

In the refined model, all the relationships between the latent variables and their respective indicators are acceptable. Every relationship, with the sole exception of EE1, exhibits a preferred relationship value exceeding 0.7 and a bootstrap value of 0, which is ideal. These results suggest that the remaining statements are robustly constructed, leading to acceptable relationships between each latent variable and its corresponding indicator.

5.4.2 Inner model

The inner model reveals the relationship between the exogenous and endogenous latent variables. The exogenous latent variables, namely Performance Expectancy (PE) and Effort Expectancy (EE), with respective effect sizes of 30.2% and 35.6%, have a satisfactory influence on Behavioral Intention (BI). Conversely, Social Influence (SI) with an effect size of 13.2%, falls short of the acceptable level of influence on BI. Furthermore, the influence SI does exert is not statistically significant. The variance in the endogenous variable BI can be explained to a degree of 45.8% by PE and EE, with a negligible portion accounted for by SI. BI in turn influences User Behavior (UB) to the extent of 62.5%, and 39.1% of UB's variance can be acceptably explained. The data suggests that students are motivated to use ChatGPT due to its perceived benefits in boosting academic performance. The system's user-friendly interface encourages its acceptance and utilization in educational contexts. However, students do not express a significant need for support systems while utilizing ChatGPT for their academic endeavors and are neutral to the effect of Social Influence.
6. Discussion and Conclusion

This chapter delves into the analysis and interpretation of the research findings on the factors affecting students' acceptance and use of ChatGPT within academic settings, using the Unified Theory of Acceptance and Use of Technology (UTAUT) as a theoretical framework. This chapter reviews the results of the study and compares it with the existing research. Thus, identifies the implications and recommendations of these findings for various stakeholders. Additionally, it explores potential future trends in the use of AI technologies like ChatGPT in academia. Lastly, this chapter concludes the result and acknowledges the limitations and highlights what future research could delve deeper into.

6.1 Discussion

The primary objective of this thesis was to identify factors contributing to students' acceptance and usage of ChatGPT within academic settings. To achieve this, the Unified Theory of Acceptance and Use of Technology (UTAUT) was employed to ascertain which factors exert influence over User Behavior (UB).

The findings of this research suggest that Effort Expectancy (EE) has the most significant impact, followed by Performance Expectancy (PE) on Behavioral Intention (BI). Social Influence (SI) was found to have a negligible effect on BI, and Facilitating Conditions (FC) did not influence the end result. BI was the only one affecting UB.

6.1.1 Reviewing Exciting Research with Result

These findings imply that the ease of use associated with ChatGPT plays a pivotal role in determining students' acceptance and usage of the system within an academic context. The statements associated with EE in the PLS-SEM analysis described students feeling at ease, productive, and finding ChatGPT simple to use.

Upon reviewing existing research, it was found that opinions vary regarding the impact of ChatGPT on Effort Expectancy in academic settings. Some studies suggest that the acceptance and use of ChatGPT are positively affected by Effort Expectancy, particularly in the initial stages (Bitzenbauer, 2023; Rahaman, 2023). Conversely, others suggest that potential limitations require extra effort, such as human oversight, to ensure accuracy (Kasneki, 2023; Dijkstra et al., 2022; Acel & Wagenmakers, n.d.). This study's result aligns with these findings; while respondents acknowledged that limitations negatively affect EE, they generally still felt that EE positively impacts their overall experience with ChatGPT.

Performance Expectancy (PE) was identified as the second most influential factor in determining acceptance and usage of ChatGPT. In the PLS-SEM analysis, the statements associated with PE highlighted ChatGPT's capacity to enhance educational performance, its usefulness for educational purposes, and its role in aiding students to acquire academic knowledge.

A review of the existing research about ChatGPT's Performance Expectancy, as outlined in chapter 1.4, shows a predominantly positive impact on performance within its given context. Existing studies underscore ChatGPT's capabilities, such as generating study materials (Kasneki, 2023; Dijkstra et al., 2022), providing translation services (Jiao et al. 2023), and executing advanced tasks (Subramani et
al., 2023; Rahaman, 2023). These capabilities likely contributed to the respondents' perception of PE's positive impact on Behavioral Intention (BI).

However, the existing research also draws attention to the limitations of ChatGPT that could affect PE. These include its inability to generate relevant diagrams for assignments (Subramani et al., 2023), occasional struggles with translation (Jiao et al., 2023), and challenges in solving mathematical problems (Frieder et al., 2023). In this study, there was no clear consensus among respondents regarding the impact of these limitations on PE.

Social Influence (SI) displayed a minimal effect on Behavioral Intention (BI), which could be attributed to a variety of reasons. Although the PLS-SEM analysis confirmed the validity of the SI results, it did not demonstrate a significant influence on BI. This could suggest that Social Influence might not be a pivotal factor in determining the acceptance and usage of ChatGPT. Alternatively, the neutrality of many responses could suggest more nuanced dynamics at play. The neutral responses could be indicative of a more complex scenario. Perhaps the students' significant others are unaware of their usage of ChatGPT, rendering the respondents unable to provide a definitive response to the statements related to SI or the students' significant others do not understand ChatGPT. As such, further research may be required to delve deeper into the role of Social Influence in technology acceptance.

The construct of Facilitating Conditions (FC) did not yield significant results and consequently did not show a considerable effect on User Behavior (UB). As such, it was not incorporated in the PLS-SEM model. This outcome might imply that FC may not be a crucial factor for this particular research context. The UTAUT model, which was formulated in 2003, placed a significant emphasis on FC, reflecting a different technological era where FC played a more pivotal role. However, the landscape has significantly changed since then, which could explain the diminished relevance of FC in this study. As highlighted in chapter 1.4, previous research underscored the importance of FC, particularly in terms of the technical and organizational infrastructure required for the deployment of ChatGPT in educational settings (Bitzenbauer, 2023). This emphasis seems to contrast with the findings of the current research. This can indicate that to better understand this discrepancy, further investigation might be necessary.

The limitations did not emerge as prominently as suggested by the prior research discussed in chapter 1.4. One potential explanation for this could be that respondents may not have deeply considered or encountered the limitations, resulting in a degree of uncertainty in their responses. The only discernible impact of limitations was observed in relation to Effort Expectancy. Limitations could thus be more noticeable when a tool that is expected to be user-friendly, proves not to be. Contrary to the findings of this study, previous research emphasizes various limitations and challenges associated with the use of ChatGPT across different contexts. To elaborate on these findings, it could be valuable to further investigate the respondents' experiences with ChatGPT.

6.1.2 Research Contribution

This research indicates that educators, decision-makers, and policymakers need to pay special attention to the key findings when planning and strategies are being developed. Emphasis should be placed on the significant role Effort Expectancy and Performance Expectancy impact play in student acceptance and use of ChatGPT in academic settings.
Contrarily, resources expended on supporting systems and infrastructure may not yield proportionate benefits, as these were found to have less significance on acceptance and usage. The low impact of limitations on acceptance can indicate a need for more research to discern whether students find these limitations insignificant, or if they are simply unaware of them.

The study's findings may be more representative for disciplines such as IT or Business, which were the fields of study for a majority of the respondents. Furthermore, it was found that students predominantly use ChatGPT in their academic work for research and learning, text writing and editing, brainstorming, and problem-solving in specific domains. This information could guide educators in optimally leveraging ChatGPT and help policymakers preempt and mitigate potential risks.

Simultaneously, AI developers can take cues from these insights. Understanding the areas where ChatGPT is used less frequently can guide their efforts towards improving the system to enhance its utility and acceptance further. The study's findings can help them focus on areas where development is needed the most, thereby enhancing the overall user experience.

6.1.3 Future predictions

With technological advancements progressing at an unprecedented pace, the factors driving acceptance and usage of tools like ChatGPT are likely to shift over time. ChatGPT presents itself as a valuable tool for students and it has already made a strong impact on society and, more specifically, within academic circles. Changing the ways students engage with their studies and teachers approach education, ChatGPT promises to continue reshaping the educational landscape.

With increasing AI sophistication, there might be an increase in the significance of Performance Expectancy (PE). If AI tools consistently improve their ability to deliver quality results, users may come to expect more from these technologies and that may become an important factor in what leads students to use ChatGPT. This improvement can also lead to higher expectations from educators. E.g. ChatGPT trains on a bigger amount of data and is able to offer more in-depth explanations or improve its ability of problem solving. This could lead to basic tasks less relevant and while more advanced academic tasks gain relevance for students using ChatGPT and making PE a more significant influence of Behavioral Intention.

Moreover, as previously discussed in section 2.3.3, Effort Expectancy (EE) often bears greater significance during the early stages of technology adoption. This is pertinent considering that ChatGPT, being a relatively recent development, has not yet accrued extensive user experience. Thus, AI technologies like ChatGPT becoming more embedded in daily life and educational settings, EE might diminish in importance over time. As users become more used to engaging with these technologies and as the technologies themselves become more intuitive and user-friendly, the perceived effort required to use them may decrease.

In the case of ChatGPT growing then Social Influence (SI) may also grow in significance. Because as more individuals adopt this technology and get used to it, then the influence from peers, friends and families could play a bigger role in the attitude towards using tools such as ChatGPT.

As also mentioned in 2.3.3, Facilitating Conditions (FC) typically hold greater significance for older users who often require more support to engage with new technology. This could contribute to the diminished relevance of this factor in the current study as it targets students. Thus, as this technology
develops it can become increasingly complex, necessitating particular technical infrastructure or support, the importance of FC might rise. But the development could also make FC significance limited as users get better in using this type of AI tools and the tool itself gets easier to use.

Given the uncertainty in predicting future trends or shifts, it's crucial to consistently monitor the evolving tech landscape and User Behaviors. This emphasizes the need for continuous research in this domain, both to understand the evolution of these influence factors and to distinguish how educational institutions can optimally support students in leveraging these advanced tools for academic success.

6.2 Conclusion

The purpose of this quantitative study was to examine which factors lead to students acceptance and use of ChatGPT in Academic settings. To answer this UTAUT was used. This thesis research question is:

What factors lead to students acceptance and use of ChatGPT in Academic settings?

The factor that lead to students acceptance and use, Use Behavior, is Behavioral Intention (BI) and BI is affected firstly by Effort Expectancy, followed by Performance Expectancy, Social Influence was found to have a negligible effect on BI, and Facilitating Conditions did not have any effect.

6.2.1 Further research and limitations

This study carries limitations that could potentially introduce bias to the results. For instance:

- The method of data collection could also contribute to bias. For instance, the way questions were formulated and the strategy for collection of data may have influenced the responses.
- The limited sample size may lead to biased results during the analysis, as it might not be fully representative of the broader population of students using ChatGPT.

Another key limitation is the rapidly evolving nature of the technology being studied. Significant advancements in ChatGPT have occurred during the execution of this research, which could potentially make some findings less relevant. For instance, the transition from ChatGPT-3 to ChatGPT-4 happened while this research was underway, but the study did not account for any potential differences in user experience between these versions.

This study focused solely on students' perspectives, leaving out another crucial group in academic settings: the teachers. Their views on the adoption and usage of ChatGPT could provide a more holistic understanding of the technology's impact on the educational landscape. For future research, should examine the teachers' perspective and their attitudes towards ChatGPT. Broadening the geographical scope of the study or increasing the sample size could also help provide a more representative and robust dataset.

Given that the constructs of Social Influence, Facilitating Conditions, and the impact of limitations did not yield significant results in this study, future research could delve deeper into these areas. It might be beneficial to reframe the research questions or employ a different theoretical framework to better understand these aspects.
7. References


Ammenwerth, Elske. 2019. Technology Acceptance Models in Health Informatics: TAM and UTAUT. July. DOI:10.3233/SHTI190111


Firat, Mehmet. 2023. How Chat GPT Can Transform Autodidactic Experiences and Open Education? Department of Distance Education, Open Education Faculty, Anadolu University. Volume 1. January 10.31219/osf.io/9ge8m


King, William R. He, Jun. 2006. *A meta-analysis of the technology acceptance model*. Elsevier September. DOI: 10.1016/j.im.2006.05.003


OpenAI. No Date. *About OpenAI.* Source: https://openai.com/about (Accessed 2023-04-1)


8. Appendices

8.1. Questionnaire

8.1.1 Intro text

Acceptance and Use of ChatGPT in Academic Settings

My name is Jakob Hasselqvist Haglund and I am a master student in Information Systems from Uppsala University. This questionnaire is part of my master's thesis, which focuses on understanding the acceptance and usage of ChatGPT in academic settings.

ChatGPT is an artificial intelligence (AI) technology developed by OpenAI, designed to process and generate human-like language. As a natural language processing (NLP) tool, it is trained to recognize patterns in language to generate contextually appropriate responses to text-based queries.

The aim of this study is to gather insights that will help understand the factors influencing students' adoption and acceptance of ChatGPT in their studies. Currently, there is limited research on this topic. This study will contribute to the field of Information Systems by providing valuable insights into how students use and accept ChatGPT in academic contexts.

Please note:

- The answers to this questionnaire are anonymous.
- Will take less than 10 minutes to complete.
- It contains of 6 parts, with different types of questions.

8.1.2 Part 1: Checking Questions

<table>
<thead>
<tr>
<th>Part 1: Checking Questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>● Check valid respondent</td>
</tr>
</tbody>
</table>
| ● Nominal data             | ● Yes 
|                            | ● No |
| ● Check valid respondent   | Have you used ChatGPT in an academic setting, such as during your education or as part of an academic program? |
| ● Nominal data             | |
### 8.1.3 Part 2: Key moderators

#### Part 2: Key moderators

<table>
<thead>
<tr>
<th>Questions</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Yes, No</td>
</tr>
<tr>
<td>Nominal data</td>
<td></td>
</tr>
<tr>
<td>What is your gender?</td>
<td></td>
</tr>
<tr>
<td>(Cultural gender)</td>
<td></td>
</tr>
<tr>
<td>Man</td>
<td></td>
</tr>
<tr>
<td>Women</td>
<td></td>
</tr>
<tr>
<td>Other (write own answer)</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Ratio data</td>
<td></td>
</tr>
<tr>
<td>How old are you?</td>
<td></td>
</tr>
<tr>
<td>Short written answer</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td></td>
</tr>
<tr>
<td>Ordinal data</td>
<td></td>
</tr>
<tr>
<td>How long have you been using ChatGPT in</td>
<td></td>
</tr>
<tr>
<td>your studies? (Release date: 30 November</td>
<td></td>
</tr>
<tr>
<td>2022)</td>
<td></td>
</tr>
<tr>
<td>More than three months</td>
<td></td>
</tr>
<tr>
<td>More than two months</td>
<td></td>
</tr>
<tr>
<td>More than one month</td>
<td></td>
</tr>
<tr>
<td>One month or less</td>
<td></td>
</tr>
<tr>
<td>Voluntariness of use</td>
<td></td>
</tr>
<tr>
<td>Ordinal data</td>
<td></td>
</tr>
<tr>
<td>To what extent do you feel that using</td>
<td></td>
</tr>
<tr>
<td>ChatGPT for your studies has been a</td>
<td></td>
</tr>
<tr>
<td>voluntary choice? Please select the option</td>
<td></td>
</tr>
<tr>
<td>that best describes your experience.</td>
<td></td>
</tr>
<tr>
<td>Completely voluntary: I use ChatGPT because I want to</td>
<td></td>
</tr>
<tr>
<td>Mostly voluntary: I use ChatGPT by choice, but there might be some external factors that influence my decision</td>
<td></td>
</tr>
<tr>
<td>Neutral: I use ChatGPT sometimes, but I don't feel strongly about it being voluntary or mandatory</td>
<td></td>
</tr>
<tr>
<td>Mostly mandatory: I use ChatGPT primarily because it is expected or required by my institution, teachers or peers</td>
<td></td>
</tr>
<tr>
<td>Completely mandatory: I use ChatGPT solely because it is a requirement for my studies</td>
<td></td>
</tr>
</tbody>
</table>

53
### Part 3: Extra information question for in depth analysis

<table>
<thead>
<tr>
<th>Level of education</th>
<th>What is your current level of education?</th>
</tr>
</thead>
</table>
| Nominal data       | ● PhD
|                    | ● Master
|                    | ● Bachelor
|                    | ● High School
|                    | ● Professional degree
|                    | ● No formal education
|                    | ● Others: |

<table>
<thead>
<tr>
<th>Education</th>
<th>In which primary subject area are you currently studying? Please select one that applies.</th>
</tr>
</thead>
</table>
| Nominal data       | ● Arts and Philosophy
|                    | ● Biology
|                    | ● Business, Economy or Management
|                    | ● Chemistry
|                    | ● Earth Science
|                    | ● Engineering
|                    | ● History and Archaeology
|                    | ● Information Technology and Data Science
|                    | ● Language and Literature
|                    | ● Law
|                    | ● Mathematics and Statistics
|                    | ● Medicine and Pharmacy
|                    | ● Peace and Conflict
|                    | ● Physics
|                    | ● Political and Social Sciences
|                    | ● Sustainability
|                    | ● Others: |

<table>
<thead>
<tr>
<th>Influence question but not relevant to UTAUT</th>
<th>Where did you first hear about ChatGPT? Please select the most applicable option.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal data</td>
<td>● Personal connections (e.g., friends, family, peers, teachers, professors, or academic advisors)</td>
</tr>
<tr>
<td></td>
<td>● Social media (e.g., Facebook, Twitter, Instagram)</td>
</tr>
<tr>
<td></td>
<td>● Online sources (e.g., articles, blog posts, forums, communities, search engines)</td>
</tr>
<tr>
<td></td>
<td>● Multimedia platforms (e.g., YouTube, podcasts, or online interviews)</td>
</tr>
<tr>
<td></td>
<td>● Educational institutions (e.g., university, school, or institutional newsletters or websites)</td>
</tr>
</tbody>
</table>
8.1.5 Part 4: Behavioral Intention (BI) and User Behavior (UB)

**Part 4: Behavioral Intention (BI) and User Behavior (UB)**

<table>
<thead>
<tr>
<th>Behavioral Intention</th>
<th>Do you plan to use ChatGPT for your studies in the next 6 months?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinal data</td>
<td>- Yes</td>
</tr>
<tr>
<td></td>
<td>- Do not know</td>
</tr>
<tr>
<td></td>
<td>- No</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavioral Intention</th>
<th>In which of the following ways do you plan to use ChatGPT to support your studies in the future? Please select all that apply.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal data</td>
<td>- Research or learning</td>
</tr>
<tr>
<td></td>
<td>- Write or editing text</td>
</tr>
<tr>
<td></td>
<td>- Problem solving and explanation for math, programming or other science problems</td>
</tr>
<tr>
<td></td>
<td>- Study planning and organization (stress redundant)</td>
</tr>
<tr>
<td></td>
<td>- Language learning or creation</td>
</tr>
<tr>
<td></td>
<td>- Creativity, inspiration or brainstorming</td>
</tr>
<tr>
<td></td>
<td>- Personal development (career, habits, goals)</td>
</tr>
<tr>
<td></td>
<td>- No, I do not plan to use ChatGPT in the future</td>
</tr>
<tr>
<td></td>
<td>- Others:</td>
</tr>
<tr>
<td></td>
<td>- Automation of simple tasks</td>
</tr>
<tr>
<td></td>
<td>- Research or learning, To provide a basic summary. The info it provides doesn’t feel “high level” it is a good starting point to build knowledge on.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Behavioral Intention</th>
<th>How confident are you in your ability to integrate ChatGPT into your future academic tasks, projects or studies?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ordinal data</td>
<td>- Very confident</td>
</tr>
<tr>
<td></td>
<td>- Confident</td>
</tr>
<tr>
<td></td>
<td>- Neutral</td>
</tr>
<tr>
<td></td>
<td>- Not too confident</td>
</tr>
<tr>
<td>Behavior Intention</td>
<td>Not at all confident</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------------------</td>
</tr>
<tr>
<td>Ordinal data</td>
<td>To what extent do you believe that using ChatGPT will become an essential part of your academic workflow in the future?</td>
</tr>
<tr>
<td></td>
<td>- Extremely essential</td>
</tr>
<tr>
<td></td>
<td>- Very essential</td>
</tr>
<tr>
<td></td>
<td>- Neutral</td>
</tr>
<tr>
<td></td>
<td>- Not very essential</td>
</tr>
<tr>
<td></td>
<td>- Not at all essential</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User Behavior</th>
<th>Ordinal data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>How often have you used ChatGPT in the context of your studies?</td>
</tr>
<tr>
<td></td>
<td>- Once a month</td>
</tr>
<tr>
<td></td>
<td>- Twice a month</td>
</tr>
<tr>
<td></td>
<td>- One to two times every week</td>
</tr>
<tr>
<td></td>
<td>- Three to four times every week</td>
</tr>
<tr>
<td></td>
<td>- Five or more times every week</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User Behavior</th>
<th>Ordinal data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On average, how much time do you spend using ChatGPT for your studies in each session?</td>
</tr>
<tr>
<td></td>
<td>- Less than one hour</td>
</tr>
<tr>
<td></td>
<td>- More than one hour</td>
</tr>
<tr>
<td></td>
<td>- More than two hours</td>
</tr>
<tr>
<td></td>
<td>- More than four hours</td>
</tr>
<tr>
<td></td>
<td>- More than six hours</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User Behavior</th>
<th>Nominal data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>In which of the following ways have you used ChatGPT to support your studies? Please select all that apply.</td>
</tr>
<tr>
<td></td>
<td>- Research or learning</td>
</tr>
<tr>
<td></td>
<td>- Write or editing text</td>
</tr>
<tr>
<td></td>
<td>- Problem solving and explanation for math, programming or other science problems</td>
</tr>
<tr>
<td></td>
<td>- Study planning and organization (stress redundant)</td>
</tr>
<tr>
<td></td>
<td>- Language learning or creation</td>
</tr>
<tr>
<td></td>
<td>- Creativity, inspiration or brainstorming</td>
</tr>
<tr>
<td></td>
<td>- Personal development (career, habits, goals)</td>
</tr>
<tr>
<td></td>
<td>- Other:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>User Behavior</th>
<th>Nominal data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Which of the following sources have had the most significant influence on your past decision to use ChatGPT for your studies? Please select all that apply.</td>
</tr>
<tr>
<td></td>
<td>- Personal connections (e.g., friends, family, peers, teachers, professors, or academic advisors)</td>
</tr>
</tbody>
</table>
- Social media (e.g., Facebook, Twitter, Instagram)
- Online sources (e.g., articles, blog posts, forums, communities, search engines)
- Multimedia platforms (e.g., YouTube, podcasts, or online interviews)
- Educational institutions (e.g., university, school, or institutional newsletters or websites)
- Professional events or networks (e.g., conferences, seminars, workshops, or workplace)
- News coverage
- No notable impact on my previous ChatGPT usage
- Others:

### 8.1.6 Part 5: Constructs

**Part 5: Constructs**

*In this part, you will be presented with different statements.*

*Please indicate whether you agree or disagree.*

This part only contains statements that you can answer:
- Strongly agree
- Partly agree
- Neutral
- Partly disagree
- Strongly disagree
- Do not know

Every construct got four statement each and one of those statements is about limitations.
| Performance Expectancy       | PE1: I believe using ChatGPT has helped me to attain gains in educational performance.  
|                             | PE2: I have found ChatGPT useful for educational purposes.  
|                             | PE3: Using ChatGPT has enabled me to acquire academic knowledge.  
|                             | PE4: The limitations of ChatGPT, have negatively affected your performance when using it for studies.  
|                             | ○ E.g., limitation: misinformation, not understanding context, bias, translation, inability of answer complex math or similar. |
| Effort Expectancy           | EE1: It has been easy to use ChatGPT for studies.  
|                             | EE2: I have felt relaxed when using ChatGPT for studies, due to its ease of use.  
|                             | EE3: I felt productive when using ChatGPT for studying.  
|                             | EE4: The limitations of ChatGPT have negatively impacted its ease of use during study sessions.  
|                             | ○ E.g., limitations: not understanding context, knowledge limits, language limits, wrong answers, inability to perform complex math or similar. |
| Social Influence            | SI1: People who are important to me believe that I should use ChatGPT for educational purpose.  
|                             | ○ E.g., important people: family, skilled student, close friends, teacher.  
|                             | SI2: Other students have influenced me to use ChatGPT.  
|                             | SI3: In general, I have felt influenced to use ChatGPT to acquire knowledge in my educational subject.  
|                             | SI4: People who influence my behavior do not think that I should use ChatGPT for studying due to its  


limitations.
- E.g., limitations: inaccurate information, wrong language translation, wrong answers, bias, not useful.

- Facilitating Conditions
- Ordinal data

- FC1: I have the resources necessary to use ChatGPT.
  - E.g. online resources and knowledge.
- FC2: The more I used ChatGPT for my studies, the greater the opportunities for continuous development and acquiring new knowledge in my subject area.
- FC3: If I encounter difficulties with ChatGPT then there is a way for me to get assistance or help.
  - E.g., Online-support, How-to-instructions, ask ChatGPT itself, or others.
- FC4: The technical limitations of ChatGPT have negatively impacted my studying experience.
  - E.g., technical limitation: computing power, response time, short answers, data collection, bias, users may wait in a queue due to limited simultaneous access to ChatGPT.

8.1.7 Part 6: Additional information

Part 6: Additional information

- Checking question
- Nominal data

Is there any additional information you would like to provide, or something you believe the test conductor should be aware of?
  - Open long answer

8.2 Social media post

Hello everyone!
I am conducting a study on the acceptance and usage of ChatGPT in academic settings for my master's thesis, and I need your help!

If you are a student and have used ChatGPT in academic settings, I would greatly appreciate your input by completing this short questionnaire. Your insights will contribute to understanding the factors that influence students' adoption and acceptance of this AI tool in their studies.

- Questionnaire link: https://forms.gle/RNEWiMQDR5qA2htL6

✅ Anonymity assured
✅ Takes less than 10 minutes to complete
✅ 6 sections with various question types

I'll be happy to share the results if you're interested! Your participation will help advance research in the field of Information Systems and shed light on the role of AI technologies like ChatGPT in education. Thank you for your valuable contribution!

Please feel free to share the questionnaire link with others who have used ChatGPT in academic settings!