Universiteit Leiden

Master Computer Science

Embracing AI in Education: Assessing the Acceptance and Potential of ChatGPT for Student Assistants in Computer Science Education

Name: Wout den Buurman
Student ID: s3261735

Date: 26/06/2023

Specialisation: Science Communication & Society

1st supervisor: Dr. R. Saunders
2nd supervisor: Dr. P.W.H. Van der Putten

Master's Thesis in Computer Science

Leiden Institute of Advanced Computer Science (LIACS)
Leiden University
Niels Bohrweg 1
2333 CA Leiden
The Netherlands

Acknowledgements

I would like to take this opportunity to express my appreciation and gratitude to those who have contributed to my thesis.

Firstly, I wish to extend my appreciation to Dr. R. Saunders and Dr. P.W.H. Van der Putten for their guidance and expertise throughout my thesis. Their support, valuable feedback, and dedication allowed me to shape my thesis into its final form.

I am also grateful to Xiaoling Zhang for the valuable meetings and discussions that helped me to refine the topic of my thesis and provide feedback for the decision on appropriate methods and frameworks.

I would also like to thank Jesse de Gans, a colleague-student and friend, for his collaboration, brainstorming sessions, thesis-writing sessions, and valuable advice and feedback on various aspects of my thesis. Furthermore, I wish to thank Dalia Papuc for meeting with me to discuss the work of student assistants and her participation in the pilot study of the survey. On that note, I would like to thank the other people who have participated in the pilot survey.

Additionally, I want to thank Leiden University for providing the necessary resources and academic environment that contributed to the completion of my thesis. In particular, I want to express my appreciation to the LIACS Education Office for distributing the survey among student assistants, which significantly contributed to the remarkable response rate, allowing for a more diverse study. On a similar note, I want to thank the survey participants for allowing me to carry out the survey analysis and process the results into the findings this research has uncovered.

Lastly, I want to express my gratitude to my family for their love, encouragement, and understanding throughout the journey of writing this thesis. Their constant belief in me and support have been invaluable by providing the motivation needed to overcome certain challenges.

Abstract

Background: Artificial Intelligence (AI) has been integrated into various aspects of education, including academic settings. The new and upcoming technologies in the field of generative AI bring along challenges, such as in academic education. Although discussions regarding these new technologies are happening with lecturers, student assistants, who bridge the gap between students and lecturers, are often overlooked. This research addresses this gap and explores the acceptance of ChatGPT as an assisting tool for student assistants in Computer Science Education.

Research Objective: The primary objective of this thesis is to investigate the acceptance of ChatGPT among student assistants in Computer Science Education at Leiden University. The Unified Theory of Acceptance and Use of Technology (UTAUT) framework was used to assess and measure acceptance and then was used to explore the core constructs of this framework and their influence on student assistants' willingness to use ChatGPT in their work. Additionally, the research takes a slight detour to explore the specific tasks that ChatGPT can or cannot assist with, as well as its potential suitability and applicability for different courses.

Methods: The UTAUT model, often used in survey studies to measure user expectancies, forms the foundation of this research. Based on this framework, a survey was conducted and 52 valuable responses were gathered from student assistants, which were then analysed individually alongside a relationship analysis according to the constructs of the UTAUT model. Additionally, the survey gathered qualitative data which helped this research to explore and understand the perspectives of the participants on this research topic.

Results: The survey results made it clear that there is limited support for ChatGPT in tasks that are often assigned to student assistants. This particularly had to do with ChatGPT's reliability and the tasks performed by student assistants which are integrity-sensitive. The results showed high expectations regarding ChatGPT's performance and mixed responses for expectations regarding Effort Expectancy and Social Influence. Furthermore, the participants expect support from Leiden University, although this did not influence their willingness to adopt the technology.

Conclusion: The findings of this study have emphasised the role of performance in shaping student assistants' willingness to use ChatGPT. Facilitating Conditions were found to be barely influential while Social Influence and Effort Expectancy displayed a moderate relationship with Behavioural Intention. Furthermore, the study revealed that introductory practical courses were expected to be more suitable for ChatGPT's assistance than theoretical courses. Participants did emphasise the necessity of human verification for integrity-sensitive tasks such as grading assignments and exams with the help of ChatGPT. The participants did also press the need for thorough testing and verification to ensure ChatGPT's reliability and capability. Additionally, the student assistants showed a preference for introductory, practical courses if ChatGPT were to be used as an assisting tool.

Contents

1 Introduction

In an era where students can generate assignments effortlessly with the help of ChatGPT[1], certain challenges arise regarding such technologies. Below the surface of these challenges lie plenty of unexplored opportunities and possibilities. As the field of generative Artificial Intelligence (AI) is rapidly evolving, it has the potential to transform education drastically. Tomorrow's education is full of innovation as AI boosts the capabilities of student assistants in Computer Science education and beyond. This thesis explores the acceptance of integrating ChatGPT as an assisting tool for student assistants in Computer Science education which transforms their work.

This research emerged from the various upcoming AI-related technologies, such as ChatGPT and Bard[2]. While discussions surrounding new educational methods mainly involve the perspectives of lecturers whereas student assistants, who show a certain interest towards teaching and are likely to become future educators, are often excluded from these discussions. The student assistants, who assist in teaching and often study at the same time, can provide important ideas and assist with integration into education. The current experience and sentiment of student assistants regarding ChatGPT will shape the future of AI in education going forward.

The decision to shift the focus to Computer Science Education is based on practical relevance and personal interest. Considering the connection that Computer Science students have with upcoming technologies and their interest in the inner workings, focusing on Computer Science Education aligns with the personal interest of the student assistants which is beneficial for the usefulness of this research. Additionally, as a Computer Science student, my interest lies in exploring the impact of generative AI on the many aspects of our lives.

Additionally, this research is timely. The field of generative AI is progressing rapidly which makes these technologies more suitable for use in educational settings. By researching this now, the current sentiments and perspectives regarding ChatGPT as an assisting tool among student assistants can be uncovered. The findings of these studies will serve as a snapshot which will likely adapt over the coming months and years with changing opinions regarding these tools and the general innovations for current and upcoming tools in the field of generative AI.

1.1 Structure

In the upcoming sections of this thesis, the practical uses of AI in education will be explored followed by ChatGPT's uses in and outside of academic education (Section 2). Section 3 describes the research objectives of the thesis and mentions the research questions and hypotheses. In Section 4 the thesis will discuss the methodology of this study. Then, the thesis discusses various acceptance models, of which one will be used in this study to measure acceptance among student assistants (Section 5). Afterwards, Section 6 discusses the setup of the survey that measures technology acceptance, followed by the survey results in Section 7. Lastly, the results will be discussed and a conclusion will be drawn in Sections 8 and 9, respectively.

[1] https://openai.com/blog/chatgpt
[2] https://bard.google.com/

2 Background

The background section consists of four subsections. The first subsection discusses AI in education, covering its applications and historical context. The second subsection focuses on ChatGPT, highlighting its general uses, benefits, and limitations. This subsection also explores ChatGPT in academic education, e.g. how it performs on certain exams and assessments. The third subsection explains the motivation behind the study, addressing the research gap and the personal interest of the author. The fourth subsection provides an overview of related work, summarising existing literature on ChatGPT in education.

2.1 AI in Education

The use of Artificial Intelligence (AI) in education has been a topic of interest for several decades. An analysis of the application of AI and deep learning techniques in teaching and learning reviewed more than 400 articles published over twenty years. The analysis shows the evolution of research on the use of technology in education from 2000 to 2019 and highlights the increasing focus on online education and the use of virtual reality (VR) and the increase of personalised learning through the use of big data [11]. This shows that education has evolved to meet the changing needs of students, but also educators. The increasing demand for distance education (2000-2009) has sped up the implementation and design of online education, whereas the emergence of big data has enabled the development of student profiling models and learning analytics to support personalised learning (2010-2019) [11]. This is relevant to this research, as it explores the potential of incorporating AI-related technologies, such as ChatGPT, to assist student assistants in their roles.

AI has the potential to provide personalised learning, alter assessments, and allow for interactions in online, mobile, or hybrid ways which can be used to assist educators in their roles by providing personalised support to students or by automating tasks. However, AI ethics and privacy concerns should be taken into account. [37]. This is relevant to the potential benefits of using AI in education. The capabilities of AI could be leveraged to assist student assistants in their roles, for example by providing personalised support to students or by automating certain tasks and the discussed importance of ethics and privacy concerns.

Furthermore, various (educational) institutions, such as the Harvard Graduate School of Education and UNESCO are also exploring AIEd. The Harvard Graduate School of Education published an article featuring Chris Dede who states he "is not overly worried about growing concerns over generative Artificial Intelligence, like ChatGPT, in education" [3]. Dede also suggests that educators should focus on teaching practical skills and wisdom that AI is unable to do, which would be more useful than competing with AI. The article highlights the potential benefits of using AI in education but also emphasises the need for careful consideration of how AI is integrated into teaching and learning practices. The article also compares ChatGPT to a search engine and refers back to times before search engines were popular [3]. Additionally, according to UNESCO[3], AI can address today's challenges in education and change teaching and learning practices [34].

[3]https://www.unesco.org/en

2.2 ChatGPT

ChatGPT is a state-of-the-art language model developed by OpenAI. It is designed to generate human-like text responses based on the input it receives. This makes interacting with the AI feel like a conversation, rather than a singular prompt. The AI uses advanced natural language processing techniques to understand and generate coherent and contextually relevant responses, giving the responses a conversational touch. Although the technology behind ChatGPT is exciting, it is not within the scope of this thesis to discuss this in detail. However, it is interesting to look at the potential uses of ChatGPT.

ChatGPT holds considerable potential for a wide range of applications, both inside and outside of education. In interactive learning environments, ChatGPT can simulate conversations, answer questions, and provide explanations on various topics. It could assist lecturers in developing course material and providing suggestions, while also being able to serve as a virtual tutor for students by answering questions or collaborating [17]. Additionally, ChatGPT can be used in academic education by assisting students and researchers in generating ideas, refining arguments, finding potentially relevant literature, and other tasks. ChatGPT also has a wide range of uses outside of education. It can take the role of a virtual assistant to handle customer support and it can serve as a personal assistant to help users with simple tasks such as scheduling appointments, setting reminders, or drafting an e-mail. In academia, ChatGPT has the potential to assist climate change researchers with generating and analysing various climate scenarios [4].

Like any technology, ChatGPT comes with its benefits and limitations. Within numerous domains, ChatGPT has the potential to enhance various tasks as described previously. In general, increased efficiency is often achieved, whether it be in coding, customer support, developing course material, or refining an academic paper. In some of these domains, increasing efficiency results in cost savings. This does not take away that there are limitations to the use of ChatGPT, such as model manipulation which leads to the production of incorrect or undesirable outputs [14]. Additionally, ChatGPT's responses should not replace critical thinking and fact-checking. The responses are usually fairly confident and human-like, making it important to verify the responses to prevent incorrect or unreliable information.

To gain insights into the potential integration of ChatGPT for student assistants in academic education, it is important to explore and describe the current applications of ChatGPT within educational settings. This exploration may lead to insights and can establish a context for the potential integration of ChatGPT as an assisting tool for student assistants.

2.2.1 ChatGPT in Academic Education

An article in Harvard Business Publishing Education discusses whether academia should adopt or resist ChatGPT and AI Text Generators (AITG). The article states that ChatGPT can improve engagement in online learning and increase motivation in asynchronous sessions. The tool will pique the students' interest by using it in day-to-day tasks. Additionally, AITGs can assist educators in preparing and reviewing sessions or constructing engaging educational content which improves learning experiences. It can also save educators time by performing repetitive work for educators such as preparing announcements. [36].

For ChatGPT to grade, it must do it consequently and correctly. If a correct answer is marked as incorrect by the AI, it will force the educator to correct the feedback from ChatGPT, which is undesirable and causes extra work. Therefore, ChatGPT must understand the assessment to minimise the number of mistakes. An article in ACS Publications evaluated the use of ChatGPT in answering chemistry assessment questions. The study evaluated ChatGPT-generated responses for end-of-year exam assessments for two chemistry-focused modules in a pharmaceutical science program. The study found that for questions that focused on knowledge and understanding, ChatGPT generated responses, unlike the application of knowledge and non-text information [9].

The William And Phyllis Mack Institute for Innovation Management published an article in January 2023 on how ChatGPT performed on the final exam of a typical MBA[4] core course, Operations Management. ChatGPT would correctly answer questions including the ones positioned in a certain context. However, ChatGPT would not be ready to solve simple mathematical questions and advanced process analysis questions. ChatGPT performs well at processing human hints and it would have received a B to B- grade (8.5 to 8.1 on the numerical scale[5]) on the exam [31].

A study by Md. Mostafizer Rahman and Yatuka Watanobe investigated the programming capabilities of ChatGPT for solving various problems including generating, debugging, improving, completing, and rewriting human-written code in different programming languages. The generated code was verified to evaluate its correctness and ChatGPT appeared to be suitable for daily practice, learning resources, and personalised programming support for programmers [24].

It is important to understand the capabilities of ChatGPT to understand the potential of this technology as an assisting tool for student assistants in Computer Science Education. The potential to increase students' motivation, assist educators in various tasks, take chemistry or MBA assessments, or perform coding-related tasks shows that this technology could assist student assistants.

2.3 Motivation

After conducting an extensive review of the existing scientific literature, it became apparent that there is a noticeable research gap regarding the potential of ChatGPT as a tool for student assistants in their work. While various articles have assessed ChatGPT on its performance for different exams, the potential impact on students, and the possibility for ChatGPT to assist educators in different tasks [36] [6] [8], there is an absence of studies specifically focusing on the role of ChatGPT in supporting student assistants in their work. While student assistants share similarities with lecturers in their involvement in academic tasks, such as grading and instructing students during practical sessions, their responsibilities and interactions with students differ. Unlike instructors, who primarily deliver lectures and provide course materials and assessments, student assistants often operate in a more support-oriented way, assisting students when the lecturer is unavailable, such as during practical sessions or question-and-answer sessions.

During the literature review, it was evident that the adoption and acceptance of ChatGPT or other generative AI tools within the academic context have not been thoroughly examined

[4]https://en.wikipedia.org/wiki/Master_of_Business_Administration
[5]https://www.science.smith.edu/~jorourke/Grading.html

using acceptance models such as the Technology Acceptance Model [6] or the Unified Theory of Acceptance and Use of Technology (UTAUT) [7]. While there is some research on the acceptance of ChatGPT among students [29], there is a notable gap in understanding the factors in user acceptance and user experiences specific to the use of this tool by student assistants.

By addressing this gap in the literature, this study aims to contribute to the field by exploring the acceptance of ChatGPT as an assisting tool for student assistants and bridging the gap between the existing research and the unique needs and perspectives of lecturers and students. The results of this study will likewise contribute to instructors and educators since student assistants and lecturers share similarities in their tasks. This research holds significance in providing insights into the effective integration and adoption of generative AI technologies in educational settings.

In addition, the widespread use of ChatGPT should not be underestimated, and student assistants are likely to utilise this technology in their work to some extent. As the field of AI continues to evolve, it is expected that similar technologies will emerge. Recognising the potential interest in adopting ChatGPT as an assisting tool for student assistants, it becomes important to thoughtfully approach its potential implementation, ensuring that the maximum benefits are harnessed while addressing the associated challenges and ethical considerations. Moreover, conducting an assessment of the potential interest in integrating ChatGPT as an assisting tool for student assistants makes way for further in-depth studies on the practical aspects of technology adoption and implementation.

Furthermore, the findings from this research have value beyond the specific context of ChatGPT. The results can provide valuable insights into the potential and willingness to adopt generative AI in educational settings. By examining the feasibility and acceptance of ChatGPT in the role of a student assistant, this study lays the groundwork for exploring the adaptation of other forms of Generative AI in education. Whether the findings demonstrate the potential for effective integration or reveal challenges that hinder its adaptation, the outcomes of this research can inform future decision-making processes. This ensures that resources and efforts are directed towards implementing and refining systems that improve the efficiency and effectiveness of student assistants while steering research away from pursuing unpromising technologies. Furthermore, the research also allows researchers to measure the change in views on Generative AI or more specific tools, such as ChatGPT over time.

Additionally, I find ChatGPT and the broader field of generative AI to be incredibly interesting and consider it to be full of potential. As a student assistant and someone who enjoys teaching, I am drawn to potentially utilise ChatGPT in my work. The ongoing developments in this technology have sparked my curiosity, and I am impressed by the possibilities it holds for the future. I believe that exploring the integration of ChatGPT as a tool for student assistants in academic education presents a considerable opportunity to enhance the efficiency and effectiveness of student assistants.

In conclusion, the aforementioned motivation highlights the significance of this research, focusing on a distinct and unexplored area of research. By examining the potential adoption of ChatGPT as an assisting tool for student assistants, the aim is to contribute to the growing

[6] https://en.wikipedia.org/wiki/Technology_acceptance_model
[7] https://en.wikipedia.org/wiki/Unified_theory_of_acceptance_and_use_of_technology

amount of knowledge on the integration of generative AI in educational contexts and allow for further advancements in this field.

2.4 Related Work

This section discusses the existing literature on the subject of integrating ChatGPT in education. While the specific focus on student assistants is relatively unexplored, there are valuable insights from related studies that examine the use of ChatGPT in the context of lecturers or educators. These studies explore the potential benefits and challenges of incorporating generative AI technologies in educational settings. The findings of these studies can help to understand the potential applicability and effectiveness of ChatGPT for student assistants.

Recently, an article was published in Innovations in Education and Teaching International which discusses the potential implication of ChatGPT for education using a SWOT[8] analysis. This article states that the strengths of ChatGPT include its ability to generate confident answers, its self-improving capability, and its ability to provide personalised and instant responses. This makes access to information easier, decreases teachers' workload, and allows for personalised learning [8].

On the other hand, fact-checking the responses and assessing the quality of it, a lack of higher-level thinking, and a threat of bias and discrimination limit ChatGPT's potential. Its inability to adapt to context forms a threat to the integrity of academic education. Additionally, it allows for the "democratisation of plagiarism" among other threats. The article concludes by providing actions that can be taken for education and research in the era of ChatGPT. It suggests that while ChatGPT allows for both opportunities and challenges for education, universities address these threats effectively by approaching them proactively and ethically [8].

The paper "ChatGPT for Teaching, Learning and Research: Prospects and Challenges" (Opara Emmanuel Chinonso et al., 2023) discusses the educational implications of artificial intelligence. According to the paper, ChatGPT delivers responses quickly in a conversational style. However, a lack of citations and references limits the technology which aligns with the aforementioned weakness in fact-checking responses. Additionally, the paper mentions that "expert systems and machine intelligence have the potential to revolutionise education by providing personalised learning experiences, automating repetitive tasks, and allowing teachers to focus on more important tasks such as providing one-on-one attention to students" [21] [8].

Furthermore, to assess the acceptance and adoption of ChatGPT for student assistants, it is essential to consider established acceptance frameworks, such as the UTAUT framework, which has been widely used to investigate the acceptance of various technologies in different contexts. A recent study on the use of ChatGPT in higher education tried developing a model which measures what variables determine the adoption and use of ChatGPT. The research made use of seven predictors to build this model. In this study, "Habit" came out as the most influential predictor [29].

[8]https://en.wikipedia.org/wiki/SWOT_analysis

3 Research Objectives

The research objectives of this thesis are to investigate the potential adoption of ChatGPT as an assisting tool for student assistants in Computer Science Education. The acceptance of such a tool can be measured in various ways and will be done through an acceptance model for technology, which will be discussed in Section 5. Based on the gaps in existing research and published research that aligns with this topic, four research questions have been formulated along with eight hypotheses. In the following subsections, these research questions and hypotheses will be discussed.

3.1 Research Questions

The goal of formulating research questions is to provide a clear direction and purpose for the study, guide the research process, and assist in selecting appropriate research methodologies and data collection techniques. By specifying the research questions, the data analysis and interpretation can be structured better, ensuring that the findings address the objectives of the research. Based on the goal of exploring the potential adoption of ChatGPT as an assisting tool for student assistants, the following research questions have been formulated to guide this thesis.

Research Question 1: What factors are important for student assistants when considering the adoption of ChatGPT as an assisting tool? How do factors such as performance and required effort influence the intention to adopt a technology?

The first research question aims to investigate the key factors that influence student assistants' attitudes and intentions towards adopting ChatGPT. By understanding the importance of performance and required effort, insights can be obtained regarding the personal factors that potentially influence the acceptance of adopting new technologies.

Research Question 2: To what extent do external influences, such as the opinions of others and the facilitation provided by the university, impact student assistants' willingness to adopt ChatGPT as an assisting tool?

The goal of the second research question is to explore the role of external influences that define the perception that student assistants have of ChatGPT as an assisting tool. By examining the influence of outside opinions and the expected level of facilitation provided by the university, insights into the social and organisational factors that may have an impact on the adoption of this technology can be obtained.

Research Question 3: What are the perceptions of student assistants regarding the tasks that ChatGPT can assist with, based on their typically assigned tasks?

The third research question of this thesis explores the specific tasks that are typically assigned to student assistants and their perceptions regarding the assistance of ChatGPT in those tasks. By understanding how student assistants perceive the applicability of ChatGPT to their responsibilities, insights can be gained into the potential areas of integration and the potential benefits and/or limitations of using ChatGPT as an assisting tool for student assistants.

Research Question 4: Are there specific courses in the Computer Science Bachelor Programme that student assistants believe could benefit from the integration of ChatGPT as an assisting tool?

The fourth and last research question relates to the third research question, as it connects the specific tasks of student assistants to potential courses that ChatGPT can assist with. The findings from this research question can allow further research to steer the potential integration as it provides insights into the specific areas where ChatGPT's assistance may add value.

3.2 Hypotheses

Based on the aforementioned research questions and supporting statements, the following hypotheses can be derived:

Hypothesis 1: Student assistants who expect ChatGPT to be capable of improving performance in their tasks are more likely to consider its adoption.

Hypothesis 2: Effort expectancy has a minimal impact on the likelihood of adopting ChatGPT as an assisting tool for student assistants, given its already relatively simple usage.

Hypothesis 3: Opinions from influential individuals (such as lecturers, colleagues, or supervisors) and influential authorities (such as academic institutions) influence the likelihood of student assistants adopting ChatGPT.

Hypothesis 4: Facilitation of the use of ChatGPT as an assisting tool for student assistants have minimal influence on the adoption of the technology, considering its relative ease of use.

Hypothesis 5: The survey participants have a more negative view on ChatGPT assisting with grading assignments and exams than they have with giving feedback for assignments and exams.

Hypothesis 6: The student assistants that participated in the survey believe that ChatGPT's assistance is more suitable for simple, practical courses such as introductory programming courses.

Hypothesis 7: The survey participants worry about the lack of sources that ChatGPT provides with its generated answers.

Hypothesis 8: The survey participants think that ChatGPT is unsuitable for the use of synchronous question answering (such as in class) whereas it is suitable for the use of asynchronous question answering (such as through a forum).

4 Methodology

In this thesis, the research design is centred around investigating the potential acceptance of ChatGPT as an assisting tool for student assistants in Computer Science Education at Leiden University. Given the focus on acceptance, it is crucial to employ appropriate measurement techniques. As for acceptance of technology, there are various acceptance models, such as the Technology Acceptance Model (TAM) and the Unified Theory of Acceptance and Use of Technology (UTAUT), among others. These models provide a framework with predefined variables that influence the acceptance of a particular technology. Various models will be explored and their fit to this research will be determined to ensure the best possible framework is chosen.

By adopting a mixed methods approach, the aim is to gather both qualitative and quantitative data on the variables identified within the selected acceptance model. The quantitative data will enable the study to analyze descriptive statistics, providing an overview of how student assistants perceive the potential adoption of ChatGPT as an assisting tool. On the other hand, qualitative data will allow for a deeper exploration of the thoughts and perspectives of student assistants, providing further insights into the underlying reasons behind their acceptance or rejection of the technology. Additionally, the relationships among the student assistants' perspectives will be investigated.

4.1 Data Collection

For data collection, the primary method for this study is an anonymous survey that targets student assistants involved in the Computer Science Bachelor Programme at Leiden University. Simple random sampling will be employed to ensure a fair and equal opportunity for all student assistants to participate. The LIACS Education Office will be assisting with the distribution of the survey to all student assistants, ensuring that the sample represents the population.

To maintain the integrity of the data collection process and follow Leiden University's guidelines, the survey will be designed, distributed, and analysed using the Qualtrics platform[9], which offers secure and reliable data collection functionalities. This approach ensures that the data collection aligns with the academic requirements of conducting a survey at Leiden University while safeguarding the confidentiality and privacy of the participants.

4.2 Data Analysis

For data analysis, the quantitative data collected through Qualtrics will be analysed using statistical methods and data visualisation tools provided by the platform. These built-in features allow for visualising graphs and descriptive statistics, providing an initial overview of the general perceptions and attitudes towards the adoption of ChatGPT as an assisting tool for student assistants. These graphs can then be used to generate custom graphs for this thesis. Additionally, the relationship among the data will be analysed using statistical methods.

Regarding the qualitative data, a thematic analysis approach will be employed. The qualitative responses from student assistants will be examined and coded to identify recurring

[9]https://www.qualtrics.com/

themes in their thought processes. Organising and categorising the qualitative data per category enables a deeper understanding of the underlying reasons behind their perspectives and the quantitative data. This analysis will provide valuable insights into why student assistants perceive certain factors, such as performance, effort, Social Influence and facilitation as influential in their decision-making process regarding the acceptance of ChatGPT.

By combining the quantitative and qualitative analyses, the study aims to present a complete picture of student assistants' views on the potential adoption of ChatGPT as an assisting tool. The quantitative data will offer an overview of the general trends and perceptions, the qualitative data will provide a richer and more nuanced understanding of their thought processes and motivations, and the correlations among the data will discover any relationships among the results. By cross-referencing these findings, the study aims to offer insights into the development and implementation of ChatGPT as an assisting tool in the context of student assistants.

4.3 Limitations

One limitation of this methodology is the potential bias among participants. The survey participants are exclusively student assistants at Leiden University, specifically assisting in courses within the Computer Science Bachelor Programme. It is important to acknowledge that these student assistants may already hold a particular attitude to or have a particular interest in AI, specifically ChatGPT, which could influence their perceptions and willingness to adopt the technology. The significance of the bias is minimal, as the research focuses on the potential adoption of ChatGPT in the context of Computer Science Education.

Another limitation is the lack of information regarding the personal information of the participants. As the survey is conducted anonymously for ethical reasons (Section 4.4), there is a lack of awareness of the distribution of participants across different courses among other demographic variables. It is possible that larger first-year courses may have a larger number of student assistants, leading to a potential bias in the results. This bias could arise if a greater proportion of student assistants from these courses participate in the survey compared to student assistants from other courses. Although efforts have been made to ensure an equal chance of participation for all student assistants, this limitation should be taken into account when interpreting the findings.

The survey data relies on self-reported responses from student assistants which means the survey may be subject to self-reporting bias. Participants might provide socially desirable answers or their responses may not fully reflect their behaviours or opinions. Additionally, since the study makes use of a framework that measures technology acceptance, which has predefined variables that measure acceptance, the study may not capture all the relevant factors that influence the adoption of ChatGPT as an assisting tool for student assistants. Other unmeasured variables or factors may play a (significant) role in the adoption process. The chosen acceptance model may exclude factors that could have an influence on the acceptance of ChatGPT in this context.

4.4 Ethics

To ensure the ethical conduct of this study, several considerations were taken into account. Prior to data collection, consent was obtained from the participants. The survey was distributed through an official channel of Leiden University, namely the Education Office, which helped establish the legitimacy and acceptance of the study by student assistants. By informing the student assistants about the reason for this study, the student assistants knew why this survey was being conducted. Moreover, it is important to note that participation in the survey was voluntary. The survey included an introductory text that communicated the research objectives and emphasised the anonymity of the participants.

Additionally, it is worth noting that this research did not require ethical approval through the ethics review committee of the Faculty of Science at Leiden University[10]. A checklist provided by the committee indicated that when the research does not collect personal data that can be traced back to an individual, an ethical review would not be necessary. Since the survey in this study did not gather such data, it did not require an ethical review. Nonetheless, the principles of anonymity and data protection were maintained to safeguard the participants' privacy and ensure their voluntary participation.

[10]https://www.organisatiegids.universiteitleiden.nl/en/faculties-and-institutes/science/committees/ethics-review-committee

5 Acceptance Models

This section explores various acceptance models that have been developed to understand the factors that influence user acceptance of technology. These models provide valuable frameworks for examining the adoption and use of new technologies in different contexts. By examining these models, insights are gained into the key determinants and the psychological processes that lay the foundation for the acceptance of ChatGPT as an assisting tool for student assistants.

The models that will be discussed include the Motivational Model, the Innovation Diffusion Theory, the Social Cognitive Theory, the Technology Acceptance Model (TAM), and the Unified Theory of Acceptance and Use of Technology (UTAUT). Each of these models offers a unique perspective on technology acceptance, focusing on different aspects. It is important to note that the UTAUT is a unified theory that is based on other acceptance models, combining the strengths of other acceptance frameworks, hence the name [35].

By examining the advantages and disadvantages of each of these models, the final choice to utilise the UTAUT as the most suitable framework for this research is justified. This comprehensive model considers multiple factors, including Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions and their effect on Behavioural Intention which provide a robust foundation for understanding the potential adoption of ChatGPT as an assisting tool for student assistants.

5.1 The Motivational Model

The Motivational Model emphasises the role of individual motivations in technology adoption. This model suggests that intrinsic and extrinsic motivation can influence a user's decision to adopt a technology. In this context, extrinsic motivation refers to the perception that users are driven to engage in an activity because they believe it will lead to desired outcomes that are separate from the activity itself. Intrinsic motivation refers to the perception that users are motivated to engage in an activity purely for the inherent enjoyment and fulfilment achieved from the activity itself, without any external incentives or rewards.

A study in 1989 used the model in the research on user acceptance of computer technology. In this research, Davis and Bagozzi claim Perceived Usefulness as an extrinsic motivation and Perceived Ease of Use as an intrinsic motivation [10]. A different study, which also mentions Davis and Bagozzi's work, states that Output Quality and Perceived Ease of Use impact the experienced enjoyment and experienced usefulness. The study found that the Output Quality and Perceived Ease of Use indirectly influence Behavioural Intention [30].

The Motivational Model is quite straightforward in terms of use. While motivation does play an important role in improving the utilisation of certain technologies, this model falls short when applied to the context of a work-related tool. When working on specific tasks, especially in the context of teaching, accuracy is crucial. Additionally, since the tool can be used as a time-saving measure for student assistants, the required effort to use the tool is also important. Furthermore, the limited availability of recent studies that use The Motivational Model raises a concern regarding the applicability of the model.

5.2 The Innovation Diffusion Theory

The Innovation Diffusion Theory has various versions which are used in different contexts. In this case, the model used within information systems, Moore and Benbasat (1991) [20], adapted the characteristics of innovation presented by Rogers (1995) [35]. According to the same article (Venkatesh et. al, 2003) [35] and the original article by Moore and Benbasat (1991) [20], the Innovation Diffusion Theory comprises seven core constructs that influence the adoption process. Relative Advantage refers to the perception of an innovation being better than its predecessor. Ease of Use reflects the perceived difficulty of using the technology or innovation. The image relates to how the use of innovation enhances one's social image or status. Visibility defines the extent to which others in the organization can observe the use of the innovation. Compatibility considers the consistency of the innovation with existing values and experiences whereas results demonstrate the tangibility and communicability of the innovation's results. Lastly, Voluntariness of Use examines the perception of using innovation as a voluntary act [35] [20].

A recently published article investigated the factors affecting students' Behavioural Intention to use a Massive Open Online Course (MOOC) system. This study integrates the Technology Acceptance Model, which is mentioned in Section 5.4 in combination with the Innovation Diffusion Theory. The study was performed with data collected from 1148 students using a MOOC system in Malaysia. The results showed that the use of a MOOC system can be used to improve the learning performance of students [25].

A slightly more dated article investigated the factors that affect business employees' Behavioural Intention to use e-learning systems by integrating the Innovation Diffusion Theory with the Technology Acceptance Model. The study used data from 552 participants who use an e-learning system in Taiwan. The study showed that Complexity, Relative Advantage among other constructs has a significant influence on the Perceived Ease of Use. These findings used an extended model of the Technology Acceptance Model in combination with the Innovation Diffusion Theory to measure the acceptance of the e-learning system [16].

The above-mentioned studies show that the Innovation Diffusion Theory is often used in combination with aspects of the Technology Acceptance Model. Since this research aims to investigate ChatGPT, a new technology and its potential adoption for student assistants as an assisting tool, combining the Innovation Diffusion Theory with another framework can potentially be beneficial. Considering that this research does not aim to develop a framework for technology acceptance, the aim will be to find a framework that can be used without first researching the benefits of combining certain frameworks for the best potential fit.

5.3 The Social Cognitive Theory

The Social Cognitive Theory claims that individual behaviour is influenced by mutual interaction between environmental factors, cognitive and personal factors, and behaviour itself. The theory has been applied in various domains, including a study of computer utilisation [5], a study of behavioural change [18], but also a study on its contributions to information science research with particular reference to research into information-seeking behaviour and knowledge sharing [19].

The core variables of the Social Cognitive Theory include Outcome Expectations (performance and personal), Self-Efficacy, Effect, and Anxiety. Outcome Expectations refer to the anticipated consequences of behaviour, including performance-related and personal-related outcomes. Self-Efficacy concerns an individual's belief in their ability to use a specific technology. Effect represents an individual's liking or preference for a particular behaviour, such as computer use. Anxiety refers to the emotional or anxious reactions associated with performing a behaviour, such as a change in behaviour [35].

The Social Cognitive Theory promises a comprehensive approach to understanding behaviour, considering both individual and environmental factors. It takes into account the influence of personal beliefs, environmental influences, and the behavioural outcomes that shape behaviour. The result is a perspective that acts as a whole, where the core constructs are interconnected. The fact that the Social Cognitive Theory takes self-efficacy into account is important, especially in technology acceptance because it influences individuals' beliefs in their capabilities and confidence to use technology. On the other hand, the Social Cognitive Theory also comes with its limitations, such as the complexity of assessing and measuring the various constructs. Self-efficacy and Effect can be quite challenging to objectively quantify. Furthermore, the theory's emphasis on cognitive and individual factors may not fully capture the social and contextual aspects that have an impact on technology acceptance.

An article published in 2015 by Vanessa Ratten investigated factors affecting consumers' Behavioural Intention to use cloud computing services by integrating the Technology Acceptance Model with the Social Cognitive Theory. The study uses The Social Cognitive Theory and the Technology Acceptance model to test the consumer's intention to adopt cloud computing [26].

As similarly mentioned in Section 5.2, it appears that this model is often combined with other frameworks for technology acceptance which creates a challenge for the potential use of this framework for this research. Furthermore, the aforementioned challenge in measuring self-efficacy in this model poses an additional challenge to measuring the potential of ChatGPT to assist student assistants in their work.

5.4 The Technology Acceptance Model

The Technology Acceptance Model is a widely used model for explaining user acceptance of technology. It uses Perceived Usefulness and Perceived Ease of Use to determine user acceptance of technology. One of the major strengths of the Technology Acceptance Model includes its simplicity and its ability to consistently explain the variance in user intentions and behaviours across various technologies and user populations [32]. The reason behind this is that it is relatively easy to understand and apply, considering it measures acceptance through two variables (Perceived Usefulness and Perceived Ease of Use). According to an article published by the University of Nebraska - Lincoln, this can be a limitation as it may be too simple to capture a comprehensive picture of the adoption process of technology. Additionally, the same article argues that the Technology Acceptance Model does not take into account the dynamic nature of technological changes and human behaviour, as well as social contexts that facilitate intention and usage, which can be very important in certain adoption processes [2].

The Technology Acceptance Model was first developed in 1986, while later adapted and finalised in 1989 and 1996 [15]. The initial model, as shown in Figure 1 demonstrates how potential external factors, such as demographic information (denoted by $X1$, $X2$, and $X3$)

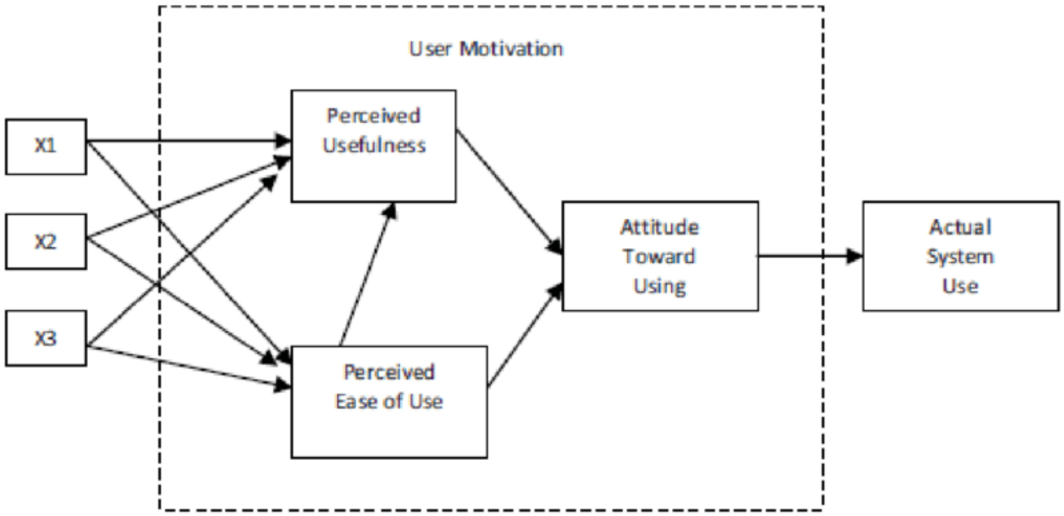

Figure 1: The Original Technology Acceptance Model (Davis, 1986)

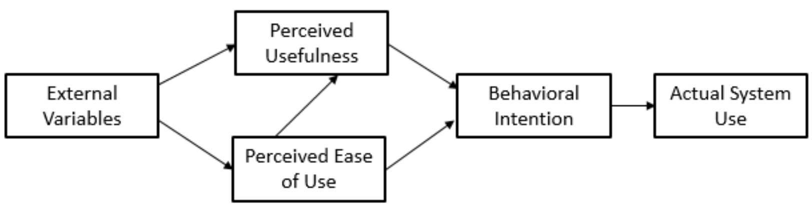

Figure 2: The Finalised Technology Acceptance Model (Venkatesh and Davis, 1996)

have an impact on Perceived Usefulness and Perceived Ease of Use. The perceived ease of use impacted perceived usefulness in some way and both Perceived Usefulness and Perceived Ease of Use influenced the attitude toward using, reflecting back to the Use Behaviour of technology. The finalised model is somewhat different, as shown in Figure 2.

Although the model is somewhat dated, it is still used in recent research. Recently, an article presented a structural equation modelling approach to explaining teachers' integration of digital technology in education. The specific study found that the data for that particular study fit the Technology Acceptance Model well and showed that Perceived Usefulness impacted Behavioural Intention [28]. Similarly, the framework was used in a study for the use of health information systems development. In this study, the Technology Acceptance Model was expanded by adding components of other frameworks, such as the Theory of Planned Behaviour [23].

5.5 The Unified Theory of Acceptance and Use of Technology

The Unified Theory of Acceptance and Use of Technology (UTAUT), shown in Figure 3, is a technology acceptance model formulated by Venkatesh and others in "User acceptance of information technology: Toward a unified view" [35]. This theory aims to explain user intentions to use information systems and usage behaviour. The theory holds that there are

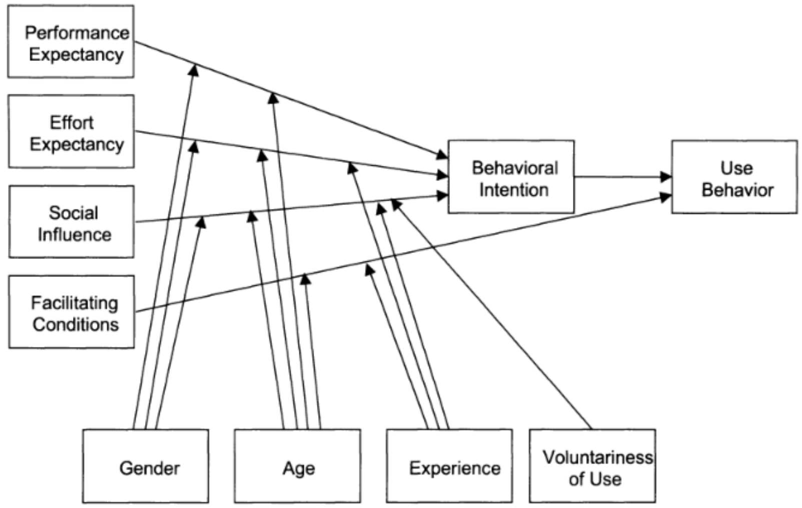

Figure 3: The Unified Theory of Acceptance and Use of Technology (Venkatesh et al., 2003)

four key constructs, namely: Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions. The first three constructs are direct determinants of whether the user ends up using a certain technology or not. The fourth construct, Facilitating Conditions influences whether or not an individual can use technology regardless of their intention [35].

In The Unified Theory of Acceptance and Use of Technology, Performance Expectancy refers to "the degree to which individuals believe that using technology will help them to improve their job performance or enhance their productivity." Effort Expectancy is considered to measure the required effort for using a certain technology. Social influence refers to social factors, such as opinions of colleagues, supervisors, and influential individuals and institutions on the user's beliefs and intentions to use the technology. The Facilitating Conditions refer to the amount of organisational and technical support that is required to use the technology to its fullest potential [35].

The framework was constructed through a review of eight models that explain information systems usage behaviour, such as the Technology Acceptance Model (Section 5.4), The Social Cognitive Theory (Section 5.3), The Innovation Diffusion Theory (Section 5.2, The Motivational Model (Section 5.1, The Theory of Reasoned Action, The Theory of Planned Behaviour, The Model of Personal Computer Use and a collaboration between the Theory of Planned Behaviour and The Technology Acceptance Model, that bring together important aspects of these models into a unified model [35].

The Unified Theory of Acceptance and Use of Technology has been widely used to evaluate technology acceptance in various contexts. For example, one study used the UTAUT model to understand students' usage of e-learning systems in developing countries [1]. Another study applied the UTAUT model to explain the students' acceptance of an AI-based early warning system at an online university. The study found that the acceptance changed over time [22].

The gathering of demographic data is a potential challenge to this research, as the research does not take into account personal data as explained in Section 4.4. However, since the target group for this research consists of similar people in terms of age group, background, and educational background, the relevance of the demographic data is minor. Furthermore, the research questions do not focus on the demographic statistics of the participants which strengthens the fact that the relevance of the demographic data is negligible.

5.6 Conclusion

The decision to use the Unified Theory of Acceptance and Use of Technology (UTAUT) model in this study was based on several factors. Firstly, the UTAUT model offers a unified view that integrates the strengths of multiple acceptance frameworks, providing a complete understanding of technology acceptance. Incorporating key constructs from various models allows for a more fulfilled and robust analysis of the factors influencing technology adoption.

Secondly, the UTAUT model aligns with the specific considerations of this study. It addresses important aspects such as Performance Expectancy, which is crucial in an educational environment where the accuracy and performance of ChatGPT as an assisting tool are of considerable significance, especially for tasks that are integrity-sensitive. Similarly, Effort Expectancy is relevant as it examines the technology's ease of use and time-saving aspects, which are essential for student assistants' workflow. Additionally, Social Influence captures the impact of influential individuals and social dynamics within an academic setting, while Facilitating Conditions explores the support and resources provided by Leiden University for technology adoption. All these core constructs measure the Behavioural Intention of the participants.

It is important to note that the demographic determinants in the UTAUT model will not be utilised, except for the participants' familiarity with ChatGPT, due to ethical considerations within Leiden University as briefly mentioned in Section 4.4 and the minor relevance due to lack of relevance to the research questions as described in Section 3.1. Furthermore, the homogeneity of the participant group, consisting of student assistants studying Computer Science at Leiden University, mitigates potential variations related to age and background, as students will fall within a similar age category and share a common educational background.

In summary, this study's decision to use the UTAUT model was driven by its unified nature, incorporating the strengths of multiple acceptance frameworks and providing an understanding of technology acceptance. The UTAUT model's focus on Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions aligns well with the research objectives, addressing the specific considerations of ChatGPT adoption in an educational environment. With the use of this framework, this thesis aims to explore the factors that influence student assistants' Behavioural Intention towards adopting ChatGPT as an assisting tool.

6 Survey Setup

This section covers the main component of this thesis: the survey setup to measure the acceptance of ChatGPT as an assisting tool for student assistants. It covers various subjects, such as the survey design (Section 6.1, the target population 6.2, and the pilot study 6.3.

6.1 Survey Design

The survey questions for this survey study were developed through a series of feedback rounds. Initially, the questions were manually formulated based on the research questions that were set up for this study. However, the questions underwent a refinement session by referring to sample questions provided in a research paper by Venkatesh on the UTAUT model [35]. Based on an intersection between the research questions and the questions provided by Venkatesh's paper [35], additional questions were added to ensure the research questions could be answered after conducting the survey. The questions that were developed for each of the core constructs of the UTAUT model and additional areas of interest, such as familiarity, can be seen in Table 1.

Aside from the questions and statements shown in Table 1, the survey contained a question that asked participants to rate the suitability of ChatGPT assisting with tasks that are often assigned to student assistants as shown in Table 2 and qualitative questions per core construct which are shown in Table 3 and their codes, respectively.

6.2 Target Population

The target population for this study consists of student assistants who currently work or have recently worked at Leiden University as student assistants. The decision to focus on student assistants at Leiden University was based on the advantages of familiarity and first-hand experience with the tasks and responsibilities that student assistants have, as I am a student assistant myself.

The target population includes students with a background in Computer Science. To be a student assistant, it is a requirement to be a student at Leiden University, which gives a rough idea of the age group of the target population, as this will not be measured through the survey. The participants may be pursuing either a Bachelor's or Master's degree in Computer Science or a strongly related field. In the analysis of the data, submissions from individuals that have been student assistants in the past will still be considered, while responses from participants who have not been student assistants will be excluded as their input is not relevant to this research.

The survey will be conducted anonymously for ethical reasons, as mentioned in Section 4.4, but also for the lack of relevance of the demographic data to the research questions and the homogeneity of the target population. The survey will be distributed through the LIACS Education Office to reach the target population effectively. This will ensure a trustworthy feeling with the survey, thereby ensuring the collection of valuable data for the study.

Variable	Code	Description
Familiarity	FM1	How often do you use ChatGPT?
	FM2	How often do you use ChatGPT for student assistant-related tasks?
Performance Expectancy	PE1	I expect that using ChatGPT will be useful.*
	PE2	I expect that using ChatGPT will allow me to accomplish tasks more quickly.*
	PE3	I expect that using ChatGPT will increase my productivity.*
	PE4	I expect that using ChatGPT will allow me to efficiently complete tasks that I consider difficult or impossible.
Effort Expectancy	EE1	I expect that becoming skilled at using ChatGPT will take significant effort.*
	EE2	I expect that becoming skilled at writing prompts for ChatGPT will take significant effort.
	EE3	I expect that becoming skilled at fact-checking answers from ChatGPT will take significant effort.
	EE4	I expect that becoming skilled at reformulating answers from ChatGPT will take significant effort.
	EE5	I expect that becoming skilled at contextualising answers from ChatGPT will take significant effort.
	EE6	I expect that keeping up with the innovations from ChatGPT will take significant effort.
Social Influence	SI1	I expect that using ChatGPT is approved by Leiden University.*
	SI2	I expect that using ChatGPT is approved by lecturers.*
	SI3	I expect that using ChatGPT is approved by other student assistants.*
	SI4	I expect that using ChatGPT is approved by students.*
	SI5	I expect that using ChatGPT is approved by people outside of academia, e.g. parents, potential employees, media, etc.*
	SI6	I expect that the opinions of others will influence my decision to use ChatGPT.*
	SI7	I expect that the acceptance of the use of ChatGPT will change overtime.
Facilitating Conditions	FC1	I expect Leiden university to provide significant support for using ChatGPT for work-related tasks.*
	FC2	I expect Leiden university to provide me with a license for ChatGPT.
	FC3	I expect Leiden university to provide me with training on the use of ChatGPT.
	FC4	I expect Leiden university to ensure ChatGPT is compatible with other university applications and systems.*
	FC5	I expect Leiden university to introduce policies and regulations regarding the use of ChatGPT.
Behavioural Intentions	BI1	I expect to use ChatGPT in the future for my work as a student assistant.*
	BI2	If I used ChatGPT as a student assistant, I expect I would be motivated to become skilled at using it.
	BI3	If I were to use ChatGPT as a student assistant, I expect I would explore different ways to use it.

Table 1: Survey statements and codes per core construct according to the UTAUT framework. Note: Statements marked with an asterisk (*) are inspired by Venkatesh et al., 2003 [35]

Code	Survey Question
TA1	Providing feedback for theoretical-based assignments.
TA2	Providing feedback for theoretical-based exams.
TA3	Providing feedback for practical-based assignments.
TA4	Providing feedback for practical-based exams.
TA5	Grading theoretical-based assignments.
TA6	Grading theoretical-based exams.
TA7	Grading practical-based assignments.
TA8	Grading practical-based exams.
TA9	Developing ungraded study material.
TA10	Answering student questions asynchronously (such as through a forum)
TA11	Answering student questions synchronously (such as in class)

Table 2: Tasks assigned to student assistants for Computer Science at Leiden University and their codes

Code	Survey Question
QTA1	Why do you think the use of ChatGPT is suitable or unsuitable for the tasks listed?
QTA2	What course in the Computer Science Bachelor programme is most suitable for the use of ChatGPT as a tool for student assistants? Why do you think so?
QPE1	In what ways do you envision integrating ChatGPT into your tasks as a student assistant?
QEE1	What specific tasks in your work as a student assistant do you think will require the most effort and why?
QSI1	How do you expect the opinions about using ChatGPT will change over time?
QFC1	If Leiden University were to provide training on how to use ChatGPT as a student assistant, what would be the most important part(s) of this training?

Table 3: Qualitative questions presented to survey participants and their codes

6.3 Pilot Study

To ensure the quality of the survey and its results, a pilot study was conducted before distributing it to the target group. The pilot study involved a small group of five participants who were strategically selected. After filling in the survey study, these participants were asked to fill out a feedback form right away which consisted of the following points:

- **Survey Ease of Understanding:**
  The majority of participants found the survey easy to understand. However, some participants suggested providing more clarity on the location of text boxes. Additionally, there was a missing title in the survey.

- **Relevance of Survey Questions:**
  All participants agreed that the survey questions were relevant to the topic of the survey.

- **Coverage of Necessary Aspects:**
  Participants confirmed that the survey covered all necessary aspects of the topic. One participant mentioned the importance of including information about personal experiences and the courses in which student assistants used the technology.

- **Difficulty/Unclear Questions:**
  Regarding the suitability of certain tasks, participants suggested allowing respondents to elaborate on their answers as the "why" aspect might provide more insightful information.

- **Additional Questions:**
  Participants expressed the desire for additional questions related to personal experiences and the difficulty/time spent in repurposing ChatGPT questions for future use in the classroom.

- **Survey Completion Time:**
  The average time to complete the survey was approximately 15 minutes.

- **Length of the Survey:**
  Participants had mixed opinions on the length of the survey. Two out of five participants felt it was too long, two out of five felt it was just right, and one out of five felt it was somewhat repetitive in terms of the core constructs of the UTAUT model.

- **Technical Issues:**
  Some participants encountered difficulties with questions that required a minimum of three answers and a maximum of three answers while the question allowed for a maximum of three answers.

- **Suggestions for Improvement:**
  Participants noted that certain questions related to Social Influence might be too personal and suggested revising them accordingly.

The participants in the pilot study were primarily Computer Science students, except one individual who is extensively involved with student assistants throughout the academic year. This participant's expertise and knowledge were valuable in identifying any inaccuracies or misconceptions within the survey. Among the Computer Science students, most are studying at Leiden University, including both teaching assistants and non-teaching assistants. The level of diversity among the participants aimed to gather various perspectives and valuable feedback to address any potential uncertainties or (technical) difficulties in the survey. Although all participants in the pilot study were of significant help, the participants who are current student assistants were the most helpful. Their input was essential as these participants relate as closely to the target group as possible.

After several rounds of feedback and the pilot study, the questions in Tables 1, 2, and 3 were finalised and presented to the survey participants.

7 Survey Results

The survey results will be discussed in two parts. Firstly, the data will be examined on a question-to-question basis, where each question will be individually addressed and its results will be briefly discussed, e.g. through distributions or analyses of figures. This analysis will also include the relevant qualitative data that was derived from the open-ended questions, where applicable. Secondly, the relationships between the core constructs will be evaluated, which is a key aspect of this research. Understanding the relationship (or the lack thereof) of the core constructs on the Behavioural Intention is one of the primary goals of the UTAUT model, and this analysis will explore these relationships.

By examining the survey results from these two perspectives, the research allows for a better understanding of the results and expectations from the participants individually but also allows for a better understanding of the underlying relationships of the core constructs according to the UTAUT model, which is particularly relevant for the UTAUT model since it will link the importance of core constructs to Behavioural Intention since those core constructs impact Behavioural Intention as shown in Figure 4.

The survey yielded a total of 61 results of which four were marked faulty as they were incomplete with a completion rate below 50%. Consequently, these results were removed leaving the total results at 57 completed survey results which could be further examined. The first question of the survey aimed to verify participants' current or past experience as student assistants, yet five individuals specified that they have not been student assistants, making their responses irrelevant to the research objectives. Consequently, these responses were excluded from the dataset. Following the filtration of the data, a final sample size of 52 responses ($n = 52$) was obtained, exceeding the expected range of 20 to 30 participants.

7.1 Individual Question Data and Qualitative Findings

This section will provide a summary of the dataset, highlighting the main findings and exploring the overall perspectives of the survey participants. As the survey was made with the UTAUT model in mind, the survey follows a clear structure. The first section asks questions regarding the familiarity that the participants have with ChatGPT, followed by a few questions regarding their thoughts on what ChatGPT can potentially assist with. Sections 2-6 ask questions regarding the core constructs of the UTAUT model: Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), Facilitating Conditions (FC) and Behavioural Intention (BI). In the following subsections, the results for each section will be discussed.

For each core construct (excluding Behavioural Intention), the participants were asked for their opinion through one open question, aiming to gain additional insights into their thought process. The data was manually categorised based on common themes and patterns. After categorisation, the results were ordered based on similar items within each category. This helped to identify common themes among the responses. It is important to note that the qualitative data analysis was based on the subset of participants who provided responses to the open-ended questions, as they were not mandatory. Nevertheless, the results are interesting and show the thought process of the participants when filling out the survey, highlighting important parts regarding the potential of adopting ChatGPT as an assisting tool for student assistants.

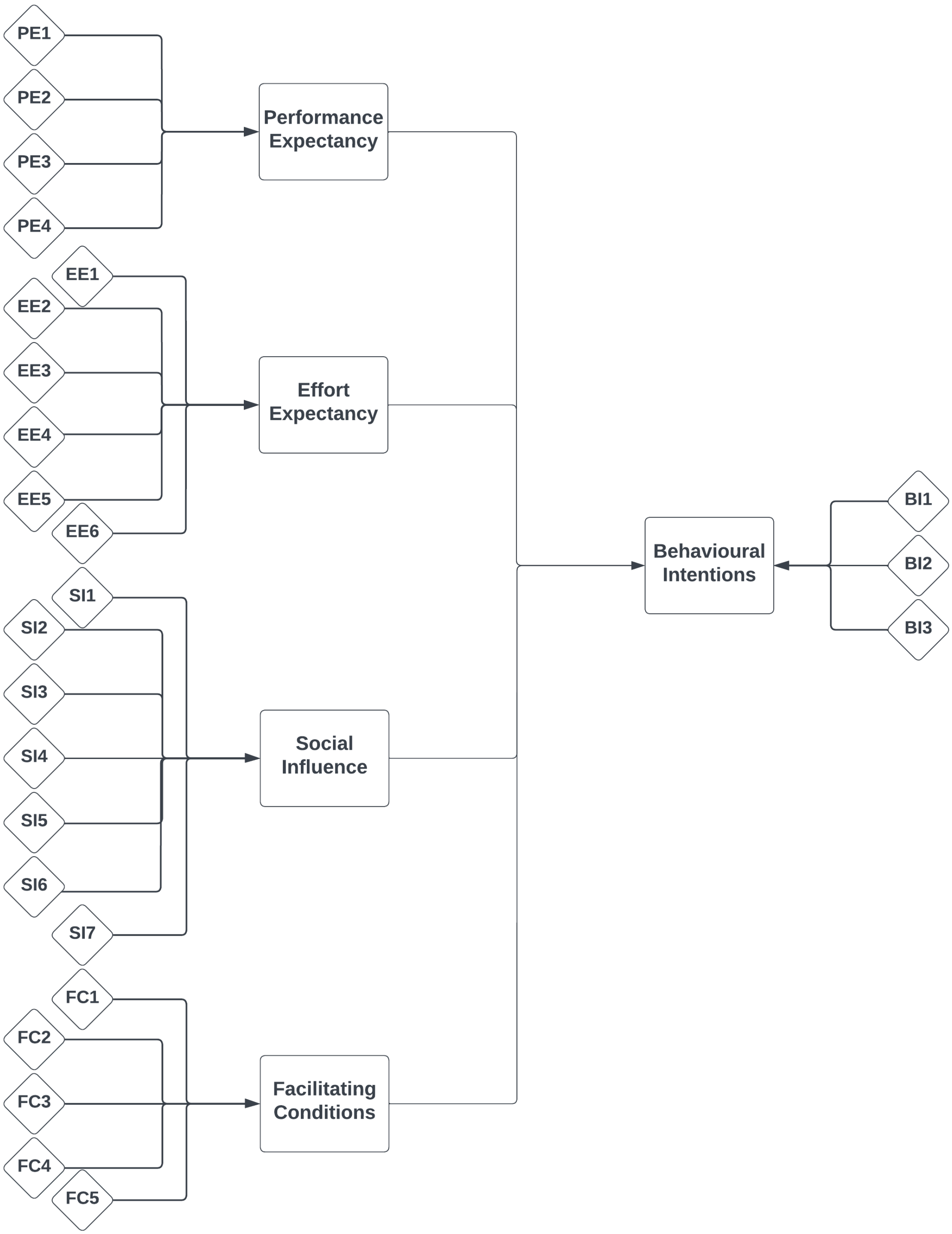

Figure 4: Visualisation of the relationships between core constructs and individual statements

7.1.1 Familiarity

Figure 5 shows the results of how often the participants use ChatGPT. In this case, using refers to how often they interact with it, regardless of the reason behind the usage. These findings have found that a considerable portion of the group has interacted with ChatGPT at least once. These results provide an overview of the participants' general familiarity with ChatGPT. The connection of this to the other aspects of the survey will be explored in Section 7.2.

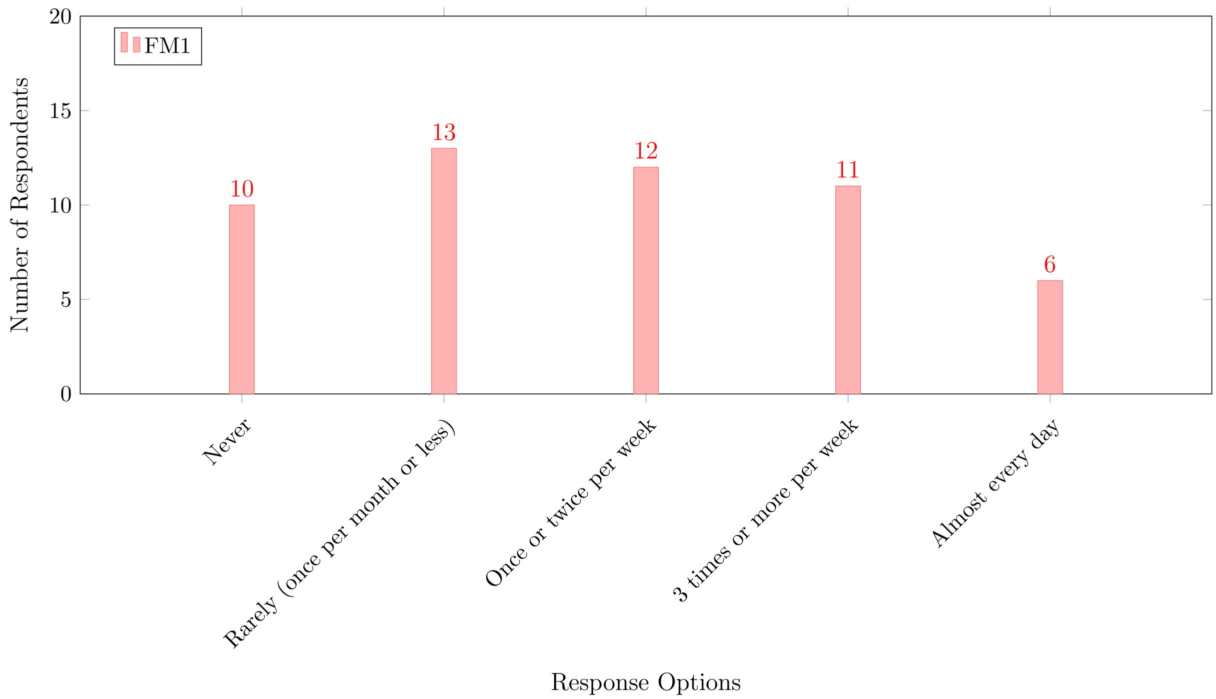

Figure 5: Results FM1: "How often do you use ChatGPT?" ($n = 52$)

According to Figure 6, a majority of the participants have not used ChatGPT for student assistant-related tasks. These findings indicate that a small portion of the group has experimented with ChatGPT for student assistant-related tasks in comparison to the general use. This could mean that these participants, who are less familiar with ChatGPT, have not thought of using it for this goal, but it could also indicate that they have had a negative experience with using it for student assistant-related tasks.

The findings regarding familiarity with ChatGPT have visualised a gap between the general use of ChatGPT and its current application as a tool for student assistant tasks. While there are varying levels of engagement with ChatGPT among the participants, a considerable number have not yet used it for work-related tasks. Further analysis in Section 7.2 will dive deeper into its relation to Behavioural Intention and each other.

7.1.2 Suitability

In a previous section, participants were asked to assess the suitability of ChatGPT for various tasks using a 5-point Likert scale[11]. The objective was to measure their opinions on the tool's

[11] https://en.wikipedia.org/wiki/Likert_scale

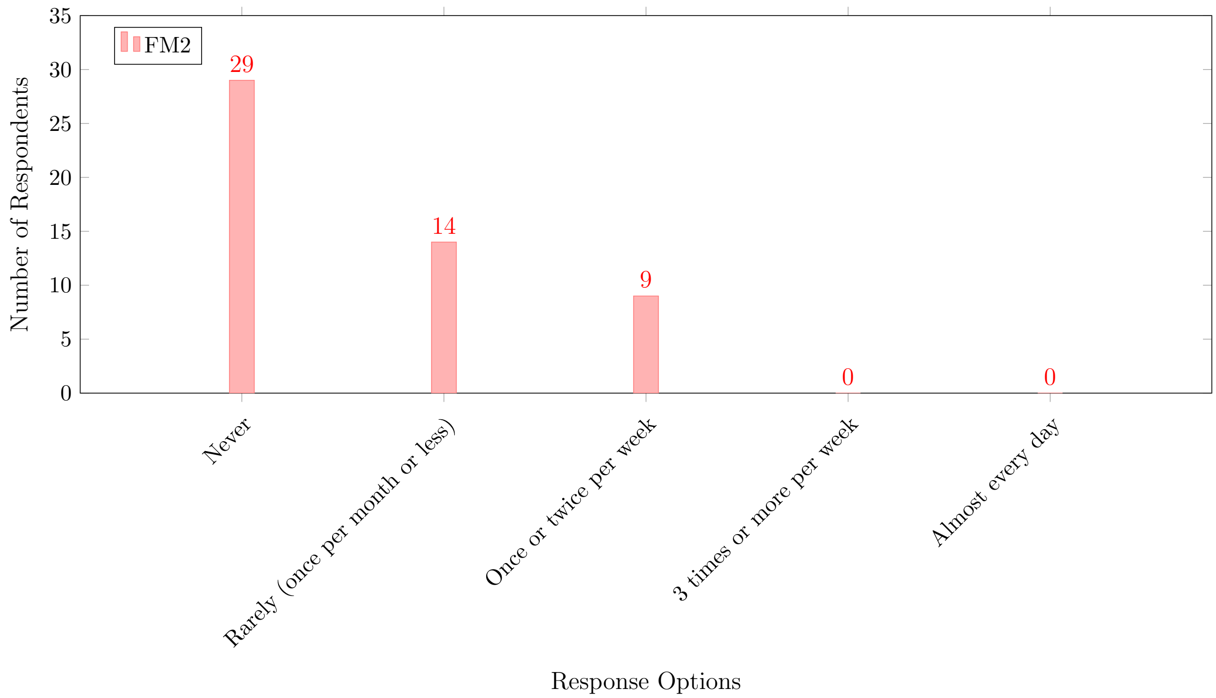

Figure 6: Results FM2: "How often have you used ChatGPT for student assistant-related tasks?" ($n = 52$)

applicability and understand their perceptions regarding specific assistance tasks. Table 4 shows the tasks that the students were asked to mark as suitable or unsuitable on a 5-point Likert Scale, including the open-ended questions, along with the codes which will be used throughout the analysis.

Individual Question Data - Suitability

Figure 7 presents the overall results regarding task suitability. A non-stacked version of the bar chart, which visualises the data in a different way, can be found in Appendix A.1. TA1, TA2, TA3, and TA4 represent the results related to providing feedback for exams and assignments, both theoretical and practical. The findings indicate that, in general, participants consider ChatGPT to be unsuitable for these tasks, especially for practical exams. The suitability scores for theoretical assignments and exams are slightly less negative than those for practical assignments and exams. Interestingly, exams are perceived as more unsuitable than assignments, potentially due to the formal nature of exams compared to assignments.

The results regarding the grading of exams and assignments (TA5, TA6, TA7, and TA8) show that grading is considered considerably more unsuitable in comparison to providing feedback. Practical assignments receive more negative responses compared to theoretical assignments, aligning with the trend observed for providing feedback. Similarly, practical -based assignments and exams seem to be less supported in terms of suitability, according to Figure 7 and Figures 20 and 21 in Appendix A.1.

Lastly, TA9, TA10 and TA11 (additionally shown in Figure 22 in Appendix A.1 present an overview of the suitability for developing ungraded study material and answering student questions, both synchronously and asynchronously. There is a notably higher number of

Code	Survey Question
TA1	Providing feedback for theoretical-based assignments.
TA2	Providing feedback for theoretical-based exams.
TA3	Providing feedback for practical-based assignments.
TA4	Providing feedback for practical-based exams.
TA5	Grading theoretical-based assignments.
TA6	Grading theoretical-based exams.
TA7	Grading practical-based assignments.
TA8	Grading practical-based exams.
TA9	Developing ungraded study material.
TA10	Answering student questions asynchronously (such as through a forum)
TA11	Answering student questions synchronously (such as in class)
QTA1	Why do you think the use of ChatGPT is suitable or unsuitable for the tasks listed?
QTA2	What course in the Computer Science Bachelor programme is most suitable for the use of ChatGPT as a tool for student assistants? Why do you think so?

Table 4: Tasks assigned to student assistants for Computer Science at Leiden University and open-ended questions and its codes

positive reactions regarding the development of ungraded study material. As for answering student questions, asynchronous responses are perceived as more suitable compared to synchronous responses. This may be due to the perception that if students can use ChatGPT themselves, synchronous assistance becomes redundant.

Qualitative Data - Task Suitability

The analysis of the open-ended questions regarding task suitability for ChatGPT's assistance revealed four themes as shown in Figure 8, some of which are more prominent than others. A total of 31 participants shared their opinions regarding this question. One participant highlighted that while ChatGPT performs well at finding theoretical knowledge online, it tends to make errors when applying this knowledge, making it unsuitable for such situations. Two participants suggested that ChatGPT should only be used as an assisting tool, which aligns with the opinions of seven participants who press the need for human verification, particularly in tasks that require integrity, such as grading and providing feedback for exams and assignments. Furthermore, the largest category of responses emphasised reliability concerns. Some of these answers warned for ChatGPT generating confident answers that are incorrect, while others noted the importance of thorough testing if it were to be used for exam grading.

Qualitative Data - Course Suitability

Aside from questioning the participants on task suitability, the survey asked them about course suitability: for which courses would the assistance of ChatGPT for student assistants

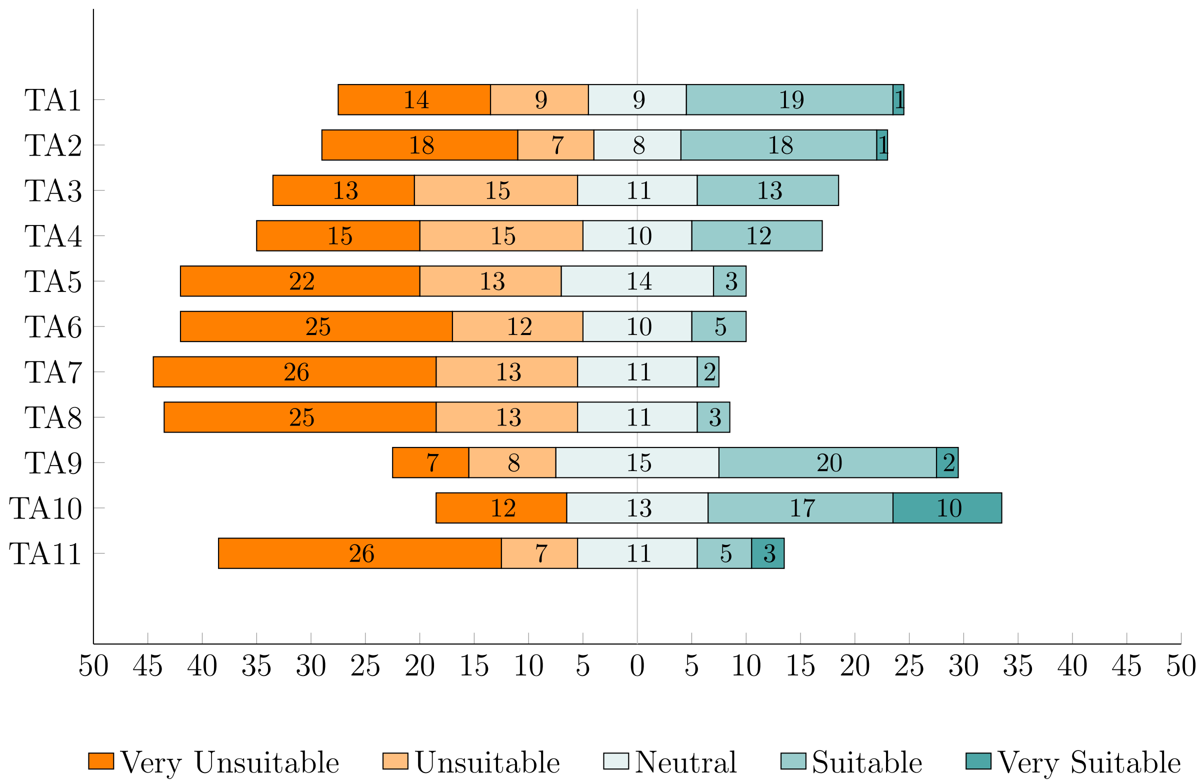

Figure 7: Stacked bar chart of survey results on tasks assigned to Computer Science student assistants at Leiden University ($n = 52$)

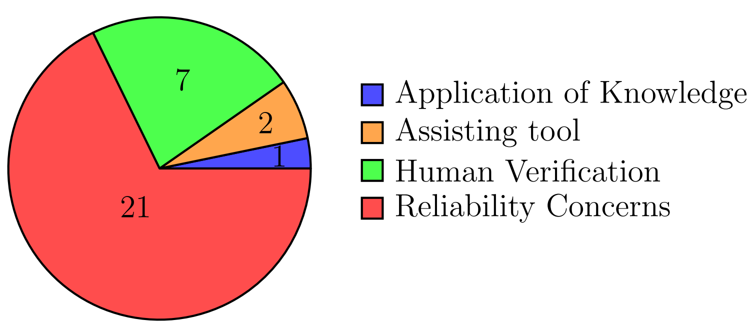

Figure 8: Pie chart of categories for student assistants' views on ChatGPT's integration in tasks

be suitable? The focus was on courses within the Computer Science Bachelor Programme at Leiden University[12]. It is important to note that these results may be influenced by personal experiences of the participants.

The results are shown in Figure 9. Among the mentioned courses, "Algoritmiek", an introductory course on algorithms, was mentioned the most, namely seven times. Additionally, the second most referenced course was "Introduction to Programming" [13], which was mentioned five times. This course teaches the programming fundamentals of Python. "Programming Methods", an introductory course in C++, and "Databases", an introductory course on databases, were both mentioned four times. These findings suggest a preference for introductory and practical courses.

Additionally, "Datastructures", "Introduction to Logic", and "Research Methods in CS" were all mentioned three times, indicating a mixture of more advanced and theoretical courses. "Studying and Presenting", "Orientation Informatics", and the mathematical courses such as "Calculus" and "Linear Algebra" were mentioned twice. Overall, these results indicate a greater interest in using ChatGPT for introductory practical courses, rather than theory-oriented courses.

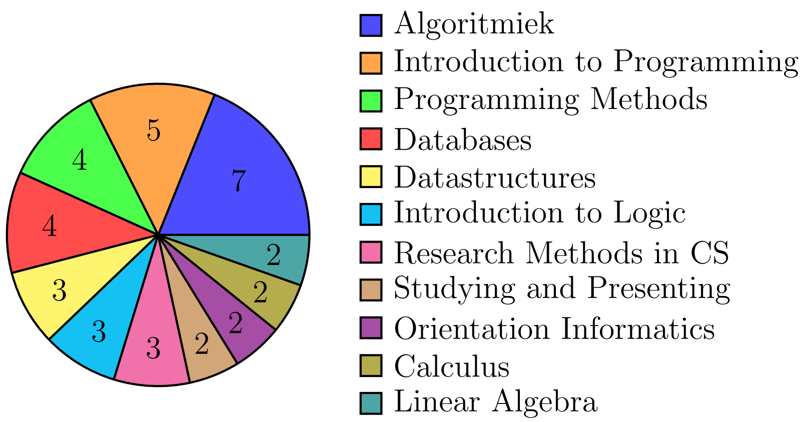

Figure 9: Pie chart of courses for student assistants' views on ChatGPT's assistance

Summary - Suitability

Overall, the reactions tend to be somewhat negative with a few exceptions. The qualitative questions revealed that the biggest pitfall regarding ChatGPT's assistance is the reliability and the pressing need for human verification if the tool were used. The qualitative results were also in favour of introductory, practical-based courses in which ChatGPT could be able to assist, rather than the more advanced courses and theoretical courses, such as Datastructures, Calculus, and Linear Algebra.

7.1.3 Performance Expectancy

In this section of the survey, the participants were questioned to what extent they agreed with the statements related to the Performance Expectancy of ChatGPT as an assisting

[12] https://studiegids.universiteitleiden.nl/en/studies/9254/computer-science#tab-1

[13] https://studiegids.universiteitleiden.nl/en/courses/116581/introduction-to-programming-bsc

tool. The goal is to determine if the participants expect ChatGPT to be useful, and improve their efficiency and productivity, and whether the student assistants expect that ChatGPT can assist with difficult or impossible tasks or not. Table 5 shows each statement that was presented to the participants with a code which will be used for data visualisation purposes.

Code	**Survey Question**
PE1	I expect that using ChatGPT will be useful.
PE2	I expect that using ChatGPT will allow me to accomplish tasks more quickly.
PE3	I expect that using ChatGPT will increase my productivity.
PE4	I expect that using ChatGPT will allow me to efficiently complete tasks that I consider difficult or impossible.
QPE1	In what ways do you envision integrating ChatGPT into your tasks as a student assistant?

Table 5: Survey statements and codes for Performance Expectancy

Individual Question Data - Performance Expectancy

Figure 10 shows the results of the statements related to Performance Expectancy. Figures 23 in Appendix A.2 show the same results in a different way. The distribution of the responses reveals certain trends. PE1, which focuses on the overall usefulness of ChatGPT, shows a left-skewed distribution, indicating that a majority of participants expect ChatGPT to be useful. Similarly, PE2, which is about accomplishing tasks more quickly, also shows a left-skewed distribution, which shows that a majority of participants expect ChatGPT to be able to assist student assistants in accomplishing tasks more quickly.

Moving on to PE3, the distribution is slightly left-skewed as well, suggesting that a slight majority of participants expect ChatGPT to enhance student assistants' productivity. PE4 mentions the capability of ChatGPT to handle difficult or impossible tasks and shows lesser left-skewed distribution which indicates that participants expect ChatGPT to be capable of assisting with challenging tasks, although to a lesser extent.

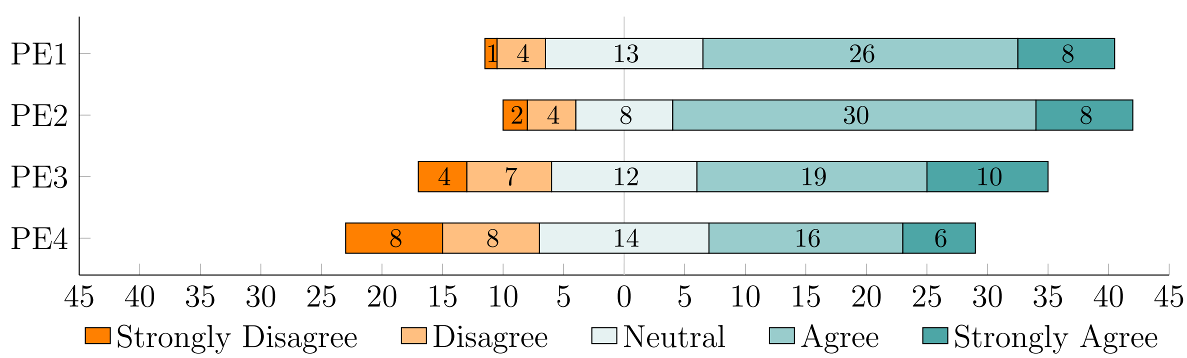

Figure 10: Stacked bar chart of survey results on Performance Expectancy statements ($n = 52$)

Qualitative Data - Performance Expectancy

Aside from the statements in Table 5, the participants were asked about their view on the ways they envision ChatGPT's integration into their tasks as a student assistant. A visualisation of the results is shown in Figure 11. Among these responses, seven individuals claimed that integration is not yet possible, due to the significant shortcomings in terms of truthfulness. This highlights concerns regarding the reliability and accuracy of the information provided by ChatGPT and potentially its lack of references and citations.

On the other hand, the vast majority of the group, totalling 22 participants, believe that ChatGPT can be integrated effectively. Within this group, twelve participants specifically mentioned its potential for assisting with writing feedback and grading. This subgroup did mention that human verification is still needed in both cases to prevent inaccuracies, but by allowing ChatGPT to do a "first round" of grading/giving feedback, it would save considerable time for the student assistants (and lecturers) to grade and give feedback.

Another four participants shared their opinion regarding the usefulness of ChatGPT in assisting with understanding course material or understanding concepts. A participant mentioned: "It can help me understand snippets of code easily", which can sometimes be difficult while grading. Another participant mentioned that it can be used for understanding syntax that is not immediately recognised or "in the cases that my own knowledge would not be sufficient enough". Another three participants mentioned that ChatGPT can help answer simple questions, especially the ones not involving complex calculations or advanced code, as "GPT-models fail very badly" on those tasks.

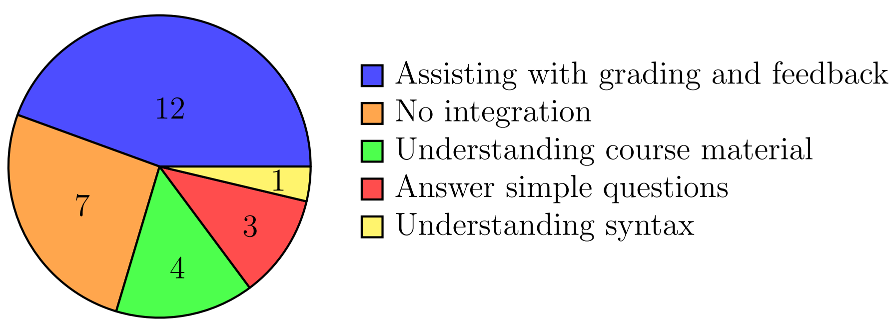

Figure 11: Pie chart of categories for student assistants' views on ChatGPT integration in their work

Summary - Performance Expectancy

Overall, the results reveal a positive sentiment regarding the statements related to Performance Expectancy. However, Figure 10 shows a lower level of support for PE3 and PE4 with a weaker left-skewed distribution compared to PE1 and PE2, which suggests that there may be doubts and uncertainties regarding those statements. Together with the analysis of the qualitative data, it shows there are concerns regarding the reliability and accuracy of ChatGPT, but the majority believes in potential integration, particularly in assisting with writing feedback and grading. Some participants highlighted ChatGPT's usefulness in understanding course material, explaining concepts, and answering simple questions while highlighting limitations in terms of applicability for more complex concepts.

7.1.4 Effort Expectancy

In this section of the survey, the participants were asked to what extent they agreed or disagreed with the statements related to ChatGPT's required effort. The goal of measuring Effort Expectancy is to determine whether the participants think that ChatGPT will take significant effort to keep up with improvements/changes or to become skilled at utilising it for certain work-related tasks. Table 6 demonstrates the codes that will be used to visualise and refer to statements.

Code	Survey Question
EE1	I expect that becoming skilled at using ChatGPT will take significant effort.
EE2	I expect that becoming skilled at writing prompts for ChatGPT will take significant effort.
EE3	I expect that becoming skilled at fact-checking answers from ChatGPT will take significant effort.
EE4	I expect that becoming skilled at reformulating answers from ChatGPT will take significant effort.
EE5	I expect that becoming skilled at contextualising answers from ChatGPT will take significant effort.
EE6	I expect that keeping up with the innovations from ChatGPT will take significant effort.
QEE1	What specific tasks in your work as a student assistant do you think will require the most effort and why?

Table 6: Survey statements and codes for Effort Expectancy

Individual Question Data - Effort Expectancy

Table 6 indicates that Effort Expectancy is measured through six statements. The results for each statement have been visualised in Figure 12. The standalone bar chart can be found in Appendix A.3. The first statement, EE1, shows a bimodal distribution with a distinguished peak at "Disagree" and a smaller, second peak at "Agree". Therefore, the majority of participants do not expect becoming skilled at using ChatGPT to take significant effort, although a considerable subgroup of the participants hold the opinion that some effort will be required. Similarly, EE2 shows a slightly bimodal distribution, with a right-skew and a less explicit bimodal pattern. Most participants disagree with the expectation that becoming skilled at writing prompts with ChatGPT requires effort, but the results also show that there is a split opinion on this matter.

EE3 displays a somewhat left-skewed distribution, which indicates that fewer participants expect fact-checking to require minimal effort. The majority of the participants either hold a neutral opinion or agree that fact-checking requires significant effort. The relatively split opinion also shows that there may be some uncertainty regarding the answers provided for EE3. Conversely, EE4 shows a slightly right-skewed distribution, which suggests no general agreement could be found with the expected effort in reformulating ChatGPT's answers.

However, this distribution also shows a split opinion, with many neutral responses indicating uncertainty again.

EE5 shows a relatively standard distribution, which implies a balanced distribution of opinions on the necessary effort to become skilled at contextualising answers. This indicates a divided opinion among the survey participants regarding this statement, which again could indicate uncertainty among the answers. On the other hand, EE6 shows a slightly right-skewed distribution, which indicates a split opinion on the required effort to keep up with ChatGPT's innovations.

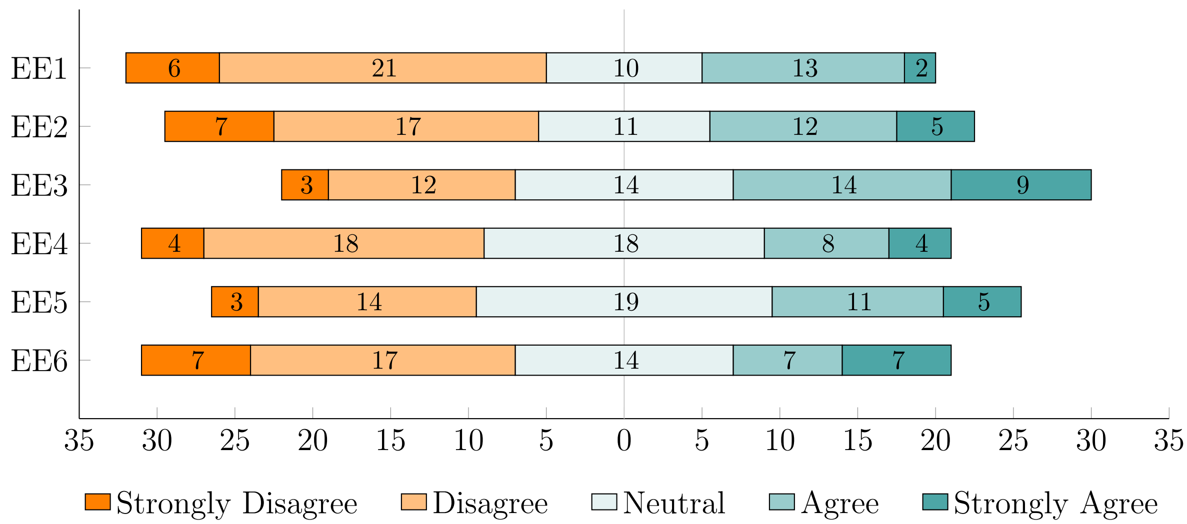

Figure 12: Stacked bar chart of survey results on Effort Expectancy statements ($n = 52$)

Qualitative Data - Effort Expectancy

Fifteen participants provided insights into their thinking process regarding the specific tasks that are perceived to require the most effort. The responses were categorised into two larger categories and four smaller categories which have been visualised in Figure 13.

The largest category identified that fact-checking ChatGPT's answers is among the tasks that are expected to require significant effort. Six participants agreed with this and have concerns about the reliability of the information provided by ChatGPT, noting that it does not provide sources and may even fabricate sources. On the other hand, another participant mentioned that advancements in the field of generative AI "will improve user-friendliness", which makes it easier to keep up with innovations. The second larger category revealed that writing prompts for ChatGPT are another task which requires effort. Three participants agreed with the importance of crafting appropriate prompts, as getting them right would make other tasks significantly easier. Effective prompts are crucial in guiding ChatGPT to generate relevant and coherent responses.

The smaller categories included one participant who mentioned determining whether ChatGPT is capable of assisting with a certain task. Another participant emphasised the importance of contextualising ChatGPT's answers to ensure relevance with the topic as a task that requires lots of effort. Additionally, keeping up with the rapid innovations in the field of generative AI was among the tasks that requires significant effort. Lastly, two participants claimed that reformulating ChatGPT's answers will be challenging because it might require in-depth knowledge of the subject to accurately convey information.

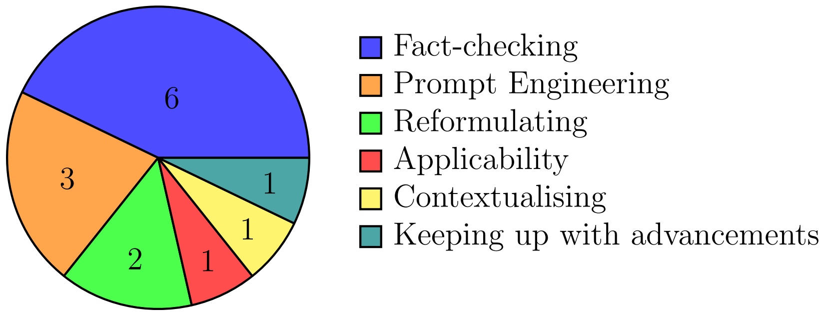

Figure 13: Pie chart of categories for student assistants' views on tasks that require effort

Summary - Effort Expectancy

Overall, the results show a significant number of neutral opinions which may suggest uncertainty among participants. Additionally, several statements shown in Figure 24 were disagreed with relatively often which indicates that interacting with ChatGPT for specific tasks would be perceived as requiring minimal effort. However, the presence of bimodal distributions, which in turn indicates split opinions, and the number of neutral opinions, highlights the existence of diverging views among the participants. With the qualitative data next to the individually analysed questions, the main categories that showed in the qualitative data highlighted that fact-checking is among the tasks that require significant effort, accompanied by writing prompts for ChatGPT.

7.1.5 Social Influence

The next survey section aims to explore the participants' understanding of the Social Influence of using ChatGPT as a supporting tool in their role as student assistants. The objective is to measure whether the potential adoption of ChatGPT is influenced by others and their perceptions of the general acceptance of ChatGPT, as well as to examine if this acceptance is subject to change over time. Table 7 shows the codes that will be used in visualisations and to refer to statements.

Individual Question Data - Social Influence

The results per statement are visualised in Figure 14, with a standalone bar chart in Appendix A.4, in which the statement codes refer back to Table 7. SI1 shows a bimodal distribution, with peaks at "Disagree" and "Agree". This indicates a divergence of opinions regarding the expected Social Influence of Leiden University in accepting ChatGPT as an assisting tool for student assistants. Similarly, SI2 also exhibits a bimodal distribution, but with a stronger peak at "Agree" compared to the peak at "Strongly Disagree". In turn, this shows varying opinions regarding the expected influence of lecturers in accepting ChatGPT for this purpose.

SI3, SI4, and SI5 are showing a left-skewed distribution. For SI3 this indicates that the majority of participants expect ChatGPT to be approved by fellow student assistants, while for SI4 this means that most participants expect ChatGPT to be approved by students. With

Code	Survey Question
SI1	I expect that using ChatGPT is approved by Leiden University.
SI2	I expect that using ChatGPT is approved by lecturers.
SI3	I expect that using ChatGPT is approved by other student assistants.
SI4	I expect that using ChatGPT is approved by students.
SI5	I expect that using ChatGPT is approved by people outside of academia, e.g. parents, potential employees, media, etc.
SI6	I expect that the opinions of others will influence my decision to use ChatGPT.
SI7	I expect that the acceptance of the use of ChatGPT will change over time.
QSI1	How do you expect the opinions about using ChatGPT will change over time?

Table 7: Survey statements and codes for Social Influence

SI5 also showing a left-skewed distribution, the data suggests that participants anticipate general acceptance of ChatGPT as a tool for student assistants by people outside of academia, such as parents, potential employees, media, etc.

In contrast, SI6 reveals a right-skewed distribution, which represents the belief among participants that the opinions of others do not majorly influence their decision to use ChatGPT for student assistant-related tasks. Finally, SI7 shows a strong left-skewed distribution which indicates a widespread agreement among the participants that the acceptance of ChatGPT's use will change over time.

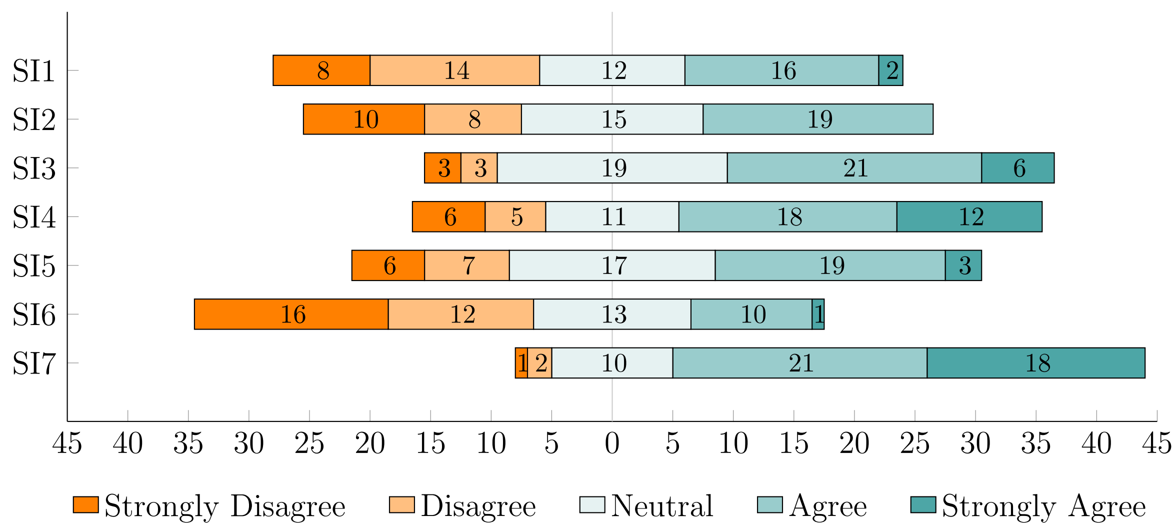

Figure 14: Stacked bar chart of survey results on Social Influence statements ($n = 52$)

Qualitative Data - Social Influence

In total, 21 participants shared their opinion on their expectations on how opinions about using ChatGPT will evolve. The responses are categorised into three groups: decreased opinion, split opinion, and increased opinion. The results have been made visible in Figure 15.

Out of the 21 results, two participants expect a decreased opinion regarding the use of ChatGPT, which is primarily caused by the concerns about it taking over certain jobs and the perception that the "hype" outweighs ChatGPT's capabilities. These participants think the opinion regarding the use of ChatGPT will decrease over time. A single participant said the attitudes toward ChatGPT will vary among different individuals or contexts.

The largest subgroup, which consists of 18 participants, believed that opinions about using ChatGPT will increase over time. Some of these participants mention the potential for improved models, suggesting that newer large language models could enhance acceptance and shift opinions positively. Others stated that the role of academia and enterprise will improve acceptance, for example for similar models developed by institutions or companies themselves. Participants also mentioned that as error rates decrease and more research-based use cases for ChatGPT are developed, the more acceptance will improve. Additionally, once ethical factors are addressed, further acceptance is expected by the participants.

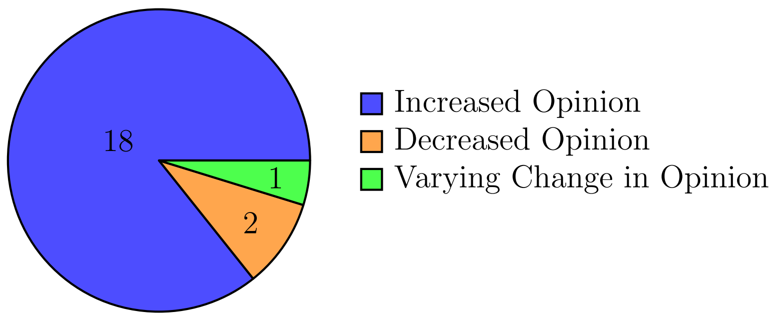

Figure 15: Pie chart of categories for student assistants' expectancies on how the opinions regarding ChatGPT will change

Summary - Social Influence

These results show the diverse perspectives regarding the Social Influence of various subgroups in accepting ChatGPT as an assisting tool for student assistants. While the use of ChatGPT for this goal is expected to be generally approved by fellow student assistants, students, and individuals outside of academia, there are varying opinions regarding the influence of Leiden University and lecturers. Furthermore, the participants also believe that their own decisions to use ChatGPT are not altered by the opinions of others and the participants also generally believe that the acceptance of the use of ChatGPT will change over time.

Together with the qualitative data, a few main categories can be extracted in terms of the expectancy of the opinion of ChatGPT and how it will change over time. The largest subgroup of the participants, which consists of 18 out of 21 participants, believe that the opinion about using ChatGPT will change positively and increase over time, especially with reduced error rates and addressing ethical considerations.

7.1.6 Facilitating Conditions

The following section focuses on Facilitating Conditions, aiming to gain valuable insights into the expectations of student assistants regarding the support and facilitation they anticipate when utilising ChatGPT in practical settings. Participants were asked to respond to five statements, which related to various aspects that Leiden University is expected to provide such as their expectations of training, the level of support, the provision of licenses, and the introduction of policies and regulations regarding the use of ChatGPT. The statements, including the code that is used to refer to the statement in later figures and text, can be seen in Table 8.

Code	**Survey Question**
FC1	I expect Leiden University to provide significant support for using ChatGPT for work-related tasks.
FC2	I expect Leiden University to provide me with a license for ChatGPT.
FC3	I expect Leiden University to provide me with training on the use of ChatGPT.
FC4	I expect Leiden University to ensure ChatGPT is compatible with other university applications and systems.
FC5	I expect Leiden University to introduce policies and regulations regarding the use of ChatGPT.
QFC1	If Leiden University were to provide training on how to use ChatGPT as a student assistant, what would be the most important part(s) of this training?

Table 8: Survey statements and codes for Facilitating Conditions

Individual Question Data - Facilitating Conditions

Figure 16 shows the results regarding Facilitating Conditions and Appendix A.5 shows the standalone bar chart, which visualises the expectations regarding the Facilitating Conditions necessary for using ChatGPT as an assisting tool for student assistants. FC1 displays a left-skewed distribution which indicates that participants are expecting significant support for utilising ChatGPT for work-related tasks.

FC2 shows a bimodal distribution, with a stronger peak for "Agree" and "Neutral", than for "Strongly Disagree". The bimodal distribution does show that the participants hold different opinions regarding the provision of licenses for ChatGPT by Leiden University. This distribution may indicate uncertainty among the participants, for example regarding the necessity or specifics of it. FC3 shows a left-skewed distribution, which shows that participants expect to receive training from Leiden University on the use of ChatGPT. This presses the need for appropriate training and support to make sure the tool is used effectively.

Similar to FC2, FC4 displays a bimodal distribution, reflecting a split opinion regarding the expected compatibility of ChatGPT with university applications and systems. This indicates that participants have varying expectations regarding the potential integration of ChatGPT within university systems and its usefulness. FC5 reveals a strong left-skewed distribution, which emphasises the importance of introducing policies and regulations concerning the use of ChatGPT for student assistants as an assisting tool. This result emphasises the importance of having defined policies and guidelines to ensure the technology is used responsibly.

Qualitative Data - Facilitating Conditions

26 participants shared their opinion regarding the most important aspects of training that Leiden University would have to provide. The distribution is visualised in Figure 17. The

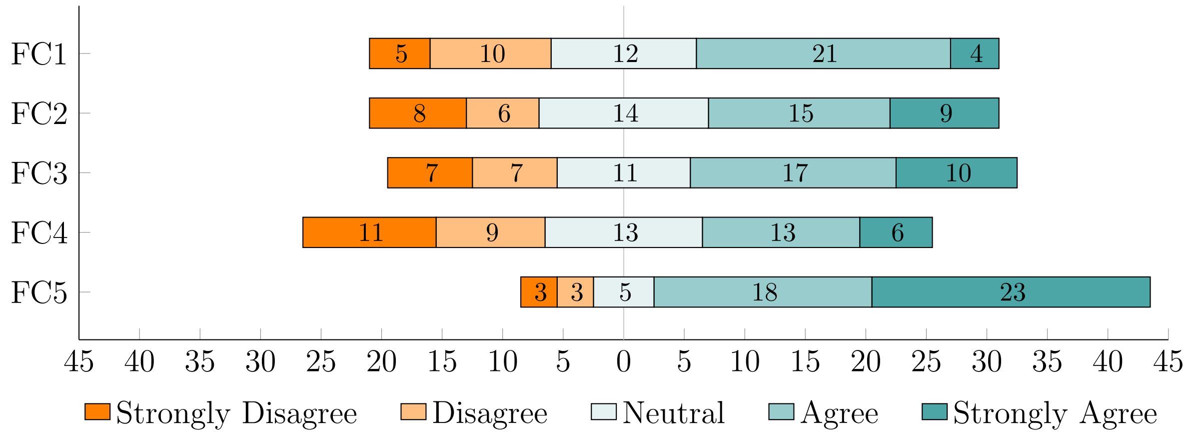

Figure 16: Stacked bar chart of survey results on Facilitating Conditions statements ($n = 52$)

largest group of participants emphasised the importance of training on when and how to use ChatGPT effectively. This includes understanding appropriate situations to use ChatGPT, but also recognising its limitations, such as realising that ChatGPT is a tool and emphasising the need to maintain a critical attitude toward the use of this tool. Ethical considerations were also mentioned as responses to this question, along with understanding the risks associated with relying solely on ChatGPT's answers.

A smaller subgroup emphasised the significance of training on how to contextualise answers provided by ChatGPT. Participants expressed the importance of learning how to formulate answers that are accurate and relevant within the given context. Another subgroup mentioned the importance of training student assistants on fact-checking ChatGPT's answers, which is in line with what is expected to take greater effort and stresses the need to verify and validate the information provided by ChatGPT.

Lastly, another subgroup of participants highlighted the importance of training student assistants on how to write prompts for ChatGPT. This is another aspect that was mentioned in the tasks that might require significant of effort when using ChatGPT, which consequently presses the need for proper training regarding prompt engineering.

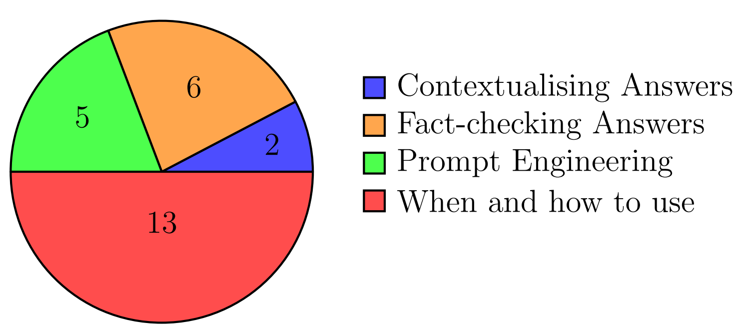

Figure 17: Pie chart of categories for student assistants' expected aspects of training on the use of ChatGPT

Summary - Facilitating Conditions

Overall, the results emphasise the importance of providing adequate support, training, and clear guidelines for the successful implementation of ChatGPT as a tool for student assistants. The participants have different opinions regarding licensing and compatibility with other university systems. Together with the qualitative data, a large group of participants emphasised the need for training on effectively using ChatGPT, including understanding its appropriate use, recognising limitations, maintaining a critical attitude, and considering its ethical implications. Another subgroup did highlight the importance of training on contextualising and fact-checking ChatGPT's answers to ensure accuracy and reliability. A third subgroup emphasised the significance of training in prompt engineering. These topics are in line with the findings from the individual data, but also with the findings for Effort Expectancy in this research, where fact-checking was among the tasks that are expected to require more effort.

7.1.7 Behavioural Intention

In the final survey section, participants were asked to express their opinions on Behavioural Intention based on three statements. These statements aimed to measure their expectations of using ChatGPT in their future work as student assistants, their motivation to become skilled at using the tool, and their incentive to explore its potential uses. Table 9 shows the codes that are used in the following sections and in visualisations of the data.

Code	Survey Question
BI1	I expect to use ChatGPT in the future for my work as a student assistant.
BI2	If I used ChatGPT as a student assistant, I expect I would be motivated to become skilled at using it.
BI3	If I were to use ChatGPT as a student assistant, I expect I would explore different ways to use it.

Table 9: Survey statements and codes for Behavioural Intention

Individual Question Data - Behavioural Intention

Figure 18 shows the visualised results for the statements on Behavioural Intention with the non-stacked bar chart in Appendix A.6. These results provide the expected future behaviour regarding the use of ChatGPT for student assistant-related tasks. In general, the chart shows a left-skewed distribution for all three statements. BI1 shows a slightly weaker left-skew, which indicates that the majority of the participants expect to use ChatGPT in the future as student assistants. However, it is worth noticing that there are still disagreements among the responses to this statement.

BI1 and BI2 both show a slightly weaker left-skewed distribution to BI3, indicating that participants are likely to explore ChatGPT in other ways if they were to use it for student assistant-related tasks. The results for BI1 and BI2 claim that the participants expect to use ChatGPT in the future for their work as student assistants and that they would be motivated to become skilled at using it.

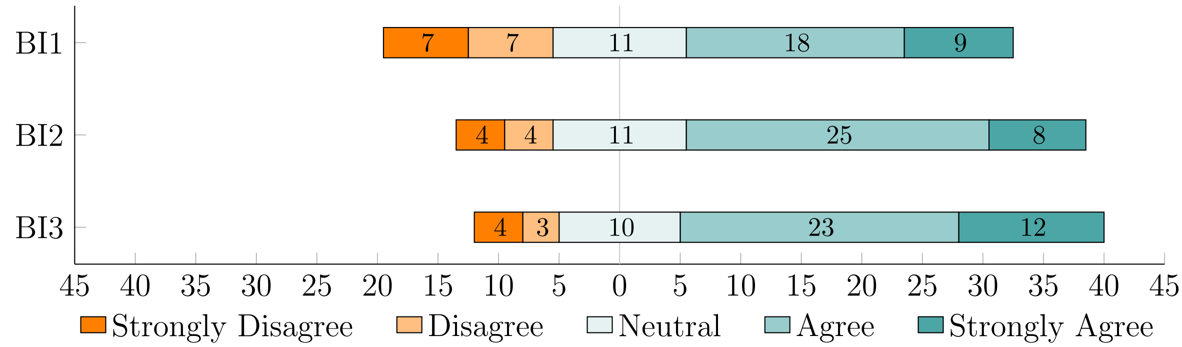

Figure 18: Stacked bar chart of survey results on Behavioural Intention statements

Summary - Behavioural Intention

Overall, the responses to the statements that relate to Behavioural Intention are generally positive which indicates willingness among participants to incorporate ChatGPT into their work as student assistants. Additionally, the results reveal that students show an incentive to explore diverse applications of ChatGPT if they were to incorporate it into their workflow as student assistants. This indicates a sense of curiosity and openness among participants regarding the possibilities and potential benefits of utilising ChatGPT in their work. Although these results are interesting, with the research questions in mind it is also valuable to discover the relationships among the core constructs of the UTAUT model. This will further explore the importance of the various statements and core constructs as a whole with Behavioural Intention as the goal.

7.2 Uncovering Relationships among the Core Constructs

In this phase of the research, the focus shifts to examining the relationships between the core constructs of the UTAUT model. The aim is to gain insights into the connections and associations among these constructs. Initially, the intention was to perform a manual analysis, but it rapidly became clear that this approach would be time-consuming since all statements from all core constructs had to individually be analysed for each of the possible answers.

To overcome this challenge, a methodology used in two recent studies was adopted. One of these studies used the UTAUT model to understand students' acceptance of e-learning systems in developing countries [1], while the other study used the same framework to measure acceptance of electronic payment systems in Serbia [33]. By following these approaches, this research aims to leverage the experience from similar studies using the UTAUT model.

The chosen approach involves utilising Structural Equation Modeling (SEM) to enable a thorough exploration of the relationships between the core constructs and Behavioural Intention with the possible expansion to also researching how the core constructs impact each other. It will assist in answering the research questions as provided in Section 3.1 by providing more insights allowing for a better understanding of the determinants of the acceptance and use of ChatGPT as an assisting tool for student assistants.

Structural Equation Modeling (SEM) is a statistical method that is used to test and estimate complex relationships among variables. It is widely used in various fields [7]. As mentioned previously, it is also used in various studies related to this research, which use the UTAUT model and utilise SEM to analyse the data [1] [33].

The exploration of the relationships through this statistical method will serve as a powerful tool to explore the relationships between the core constructs of the UTAUT model and the Behavioural Intention of the survey participants. The goal is to uncover the effects of the core constructs, Performance Expectancy (PE), Effort Expectancy (EE), Social Influence (SI), and Facilitating Conditions (FC), on the participants' Behavioural Intention (BI) regarding the use of ChatGPT as an assisting tool in their role as student assistants. By examining the relationships, in addition to the results in Section 7.1, the key factors that influence the participants' acceptance can be further explored.

After exploration of the relationships between the core constructs, it will enable further in-depth research to investigate potential relationships between individual statements in the survey. For this, a correlation matrix will be generated to determine the impact of each of the statements on each other. This addition will provide more in-depth insights into the reason behind the responses in the survey, which are important for potential future work.

7.2.1 Structural Equation Modeling

According to a book published by R.B. Kline in 2015 on the principles and practice of SEM [13], the use of a computer programme is essential for analyzing SEM. One popular open-source tool mentioned in the book is Lavaan, which is specifically designed for SEM analysis in the R programming language [27].

Kline notes that no universal rule of thumb applies to all studies in SEM. However, SEM generally benefits from larger sample sizes [13], which is also seen in the sample sizes of the study on the use of e-learning systems in developing countries in which 370 were questioned [1] and the study on the acceptance of electronic payment systems in Serbia which counted 457 participants [33]. According to Kline, a recommended sample size would consist of a sample-size-to-parameters ratio, which would be 20:1 [13]. This study makes use of 25 parameters, meaning a sufficient sample size would be $25 * 20 = 500$. This study only contains 10% of that sample size, which makes the number of participants that were questioned for this study significantly lower.

Kline also points out that larger sample sizes are usually required for more complex models [13]. However, compared to the two aforementioned studies, this research does not take demographic information and Use Behaviour into account, making the model less complex. Furthermore, by investigating the fitment of each statement on the model, there is a chance certain statements can be removed from this analysis, as they do not provide additional value to the analysis. However, it remains important that the number of participants in this study is comparatively lower, meaning careful interpretation of the results is still necessary, considering the potential impact of sample size on the stability of the findings.

Structural Equation Modeling begins with specifying a model that represents the potential relationships between latent and observed variables. In this research, the core constructs are considered latent variables and the observed variables are the individual statements in the survey that were presented to the participants. The measurement model explores the relationships between the latent variables and their observed variables. The structural model then explores these relationships, in this case, PE, EE, SI, FC and BI. In order to measure how well the model fits the observed data, it is important to perform Confirmatory Factor Analysis (CFA). This can potentially weaken the importance of a considerably smaller dataset, by measuring model fitness on the given data, resulting in a possibility that the results contain a degree of usefulness in discovering the relationships between the core constructs of the UTAUT model.

7.2.1.1 Confirmatory Factor Analysis To perform Confirmatory Factor Analysis, it is important to define the model for which this is done. The model is visualised in Figure 4, showing how each latent variable (or core construct) consists of the observed variables (or statements). After executing the program, Lavaan produces a report with statistics from the model which can be used to see how well the model performs. Table 10 contains the data, which shows that the model does not score very well in terms of Root Mean Square Error of Approximation and Standardised Root Mean Square Residual. This is an expected result since the sample size is significantly smaller than the recommended sample size.

Goodness-Fit Indexes	**Range**	**Ideal Result**	**Result**
Comparative Fit Index (CFI)	[0, 1]	1	0,693
Tucker Lewis Index	[0, 1]	1	0,651
Root Mean Square Error of Approximation	[0.05, 0.08]	Within the range	0,134
Standardised Root Mean Squared Residual	[0, 0.08]	Within the range	0,129

Table 10: First Confirmatory Factor Analysis results

To potentially improve these scores, parameters with high modification indices can be inspected. These indicate potential areas of the model where the fit could be improved by removing parameters due to a high modification index. This is possible within the Lavaan library. A small addition to the code resulted in Table 11 showing the modification indices. According to this table, seven statements were best to be removed based on the modification indices, since the value is greater than 10 [1] [12].

Left	**Relation To**	**Right**	**Modification Index**
EE1	→	EE2	35,322
SI2	→	BI1	14,070
EE3	→	EE4	13,798
SI1	→	BI1	11,837
PE3	→	SI2	11,763
PE3	→	BI1	11,138
SI1	→	FC1	11,135
EE4	→	EE6	11,015
PE2	→	BI2	10,754
EE4	→	EE5	10,062
EE1	→	EE4	9,416

Table 11: First result of Modification Indices

Why does EE1 hold such a strong relationship to EE2?

EE1 appears twice in Table 11, particularly the first table entry, showing a modification index of 35,322. This is a logical consequence of the way the statements regarding Effort Expectancy were set up. EE1 states: "I expect that becoming skilled at using ChatGPT will take significant effort.". EE2, EE3, EE4, and EE5 build upon EE1 by focusing on the specific tasks that might require effort to become skilled at, such as fact-checking answers, reformulating answers, contextualising answers, and writing prompts.

After removing the seven statements (PE2, PE3, EE1, EE3, EE4, SI1, SI2) due to their high modification index and altering the model, the Confirmatory Factor Analysis was performed again to see whether the goodness indices had improved. Table 12 shows the results of the Confirmatory Factor Analysis after removing the aforementioned statements. The table shows, when compared to Table 10, that the results have slightly improved, but not yet in the recommended ranges, which invited for re-assessment of the modification indices. Table 13 shows the 5 statements with the highest modification index. As a result, EE2 will be removed as the modification index is greater than 10 for its relation to SI3, but by removing SI3, its relation to SI5 and FC4 is also removed, which have a relatively high modification index.

Goodness-Fit Indexes	Range	Ideal Result	Result
Comparative Fit Index (CFI)	[0, 1]	1	0,846
Tucker Lewis Index	[0, 1]	1	0,811
Root Mean Square Error of Approximation	[0.05, 0.08]	Within the range	0,091
Standardised Root Mean Squared Residual	[0, 0.08]	Within the range	0,102

Table 12: Second Confirmatory Factor Analysis result after removing statements based on Modification Index

Left	Relation To	Right	Modification Index
EE2	→	SI3	12,313
EE2	→	SI5	7,158
EE2	→	FC4	7,058
SI	→	EE6	6,111
BI	→	EE6	6,058

Table 13: Second result of Modification Indices

After removing the additional statement (EE2) due to a high modification index and altering the model again, the Confirmatory Factor Analysis was re-evaluated. The results in Table 14 show that after removing those statements, the scores improved considerably which indicates that the current measurement model is a good fit for Structural Equation Modeling.

7.2.1.2 Results After performing Confirmatory Factor Analysis and defining the new, improved measurement model, Structural Equation Modeling can be carried out. Although 17 statements are remaining, making the recommended sample size equal to $n = 340$, which is still significantly lower than the sample size this study is working with ($n = 52$), the analysis

Goodness-Fit Indexes	**Range**	**Ideal Result**	**Result**
Comparative Fit Index (CFI)	[0, 1]	1	0,935
Tucker Lewis Index	[0, 1]	1	0,918
Root Mean Square Error of Approximation	[0.05, 0.08]	n/a	0,059
SRMR	[0, 0.08]	n/a	0,088

Table 14: Third Confirmatory Factor Analysis result after removing statements based on Modification Index

was carried out. The Lavaan library was able to do this with two simple commands after specifying the model. The results are organised in Table 15 and show that the core constructs hold a positive relationship to BI. Performance Expectancy is a convincing outlier, scoring the highest among all other core constructs, showing that this aspect is very important to the participants for their Behavioural Intention.

Variable	**Score**
PE	1,091
EE	0,289
SI	0,302
FC	0,017

Table 15: Results of Structural Equation Modelling after performing Confirmatory Factor Analysis

When comparing the results to the results that SEM would have obtained without Confirmatory Factor Analysis, the results were slightly different. Table 16 shows these results, which indicate that the relationship of PE to BI would be quite similar. The relationship between EE and BI, as SI and BI would have been weaker compared to the results that were obtained when performing Confirmatory Factor Analysis. The relationship that FC has to BI remains negligible as it is very close to zero.

Variable	**Score**
PE	1,087
EE	0,134
SI	0,162
FC	0,026

Table 16: Results of Structural Equation Modelling without performing Confirmatory Factor Analysis

In summary, Structural Equation Modeling in combination with Confirmatory Factor Analysis and removing certain statements based on their modification indices, shows that Performance Expectancy has the largest impact on Behavioural Intention. Effort Expectancy and Social Influence are indicated to have an impact on Behavioural Intention, although this impact is smaller compared to the impact of Performance Expectancy. Moreover, Facilitating Conditions shows to have a negligible relationship to Behavioural Intention. Confirmatory Factor Analysis and the removal of statements have been shown to impact Effort Expectancy and Social Influence the most, which is a logical consequence as the statements removed

(PE2, PE3, EE1, EE3, EE4, SI1, SI2, and EE2) mostly relate to Effort Expectancy and Social Influence.

7.2.2 Correlation Matrix

In the context of this research, using Structural Equation Modeling presents a challenge due to its need for larger sample sizes. The recommended sample-size-t-parameters ratio of 20:1 indicates that the sample size this study achieved is not sufficient for this research, given the number of statements this study has [13]. However, through Confirmatory Factor Analysis and eliminating certain statements that did not align well with the model, the goodness-fit indices improved and mostly fell in the recommended range.

To compensate for the aforementioned limitations of SEM and to gain further insights into the relationships between statements and core constructs, a correlation matrix was constructed. This correlation matrix, seen in Figure 19 includes all 25 statements, which were presented to the survey participants. The correlation matrix allows for further examination of the interrelation among statements as well as the interrelation between statements and the core constructs.

7.2.2.1 Data Preparation To facilitate the correlation matrix analysis, the survey data required some minor adjustments. Initially, the questions in this section were presented to participants using a 5-point Likert scale. However, SEM and correlation matrices generally work better with numeric data. To address this, the textual responses, such as "Strongly Disagree" or "Very Unsuitable" were converted to numerical values ranging from 1 (indicating strong disagreement or unsuitability) to 5 (representing strong agreement or suitability.

Furthermore, as the survey did not provide a separate score for the core constructs, an average score was calculated for each participant based on the related statements. This average score was then used to measure the participants' opinions regarding each core construct. By including this data in the analysis, it became possible to examine the correlations between individual statements and core constructs, as well as the relationships among different core constructs.

7.2.2.2 Results The complete correlation matrix of all 25 statements is shown in Figure 19. Based on the correlation matrix, the general impression shows that the influence of the core constructs (Performance Expectancy, Effort Expectancy, Social Influence, and Facilitating Conditions) and the statements are clearly visible in the matrix. This is a logical consequence, as within each core construct, there is a certain overlap in the meaning of the statements, e.g. the statements regarding Social Influence are different, but relate to the same construct: Social Influence.

The Performance Expectancy statements show a general correlation with the overall Behavioural Intention of the participants. Specifically, PE1 correlates with Social Influence (except SI6), and EE1 and EE2 show correlations with Facilitating Conditions (Except FC5). EE6 shows a negative correlation with SI1, SI2, SI3, and SI4. Additionally, FC1 and FC3 show a negative correlation with SI4.

Based on the correlations within statements, the core constructs influence each other (rather than just the core constructs influencing Behavioural Intention). Performance Expectancy correlates with Social Influence and Behavioural Intention. Effort Expectancy has minor correlations to every core construct, except for a slight correlation with Facilitating Conditions. The largest correlation shown between Performance Expectancy and Behavioural Intention is in line with the Structural Equation Modeling analysis.

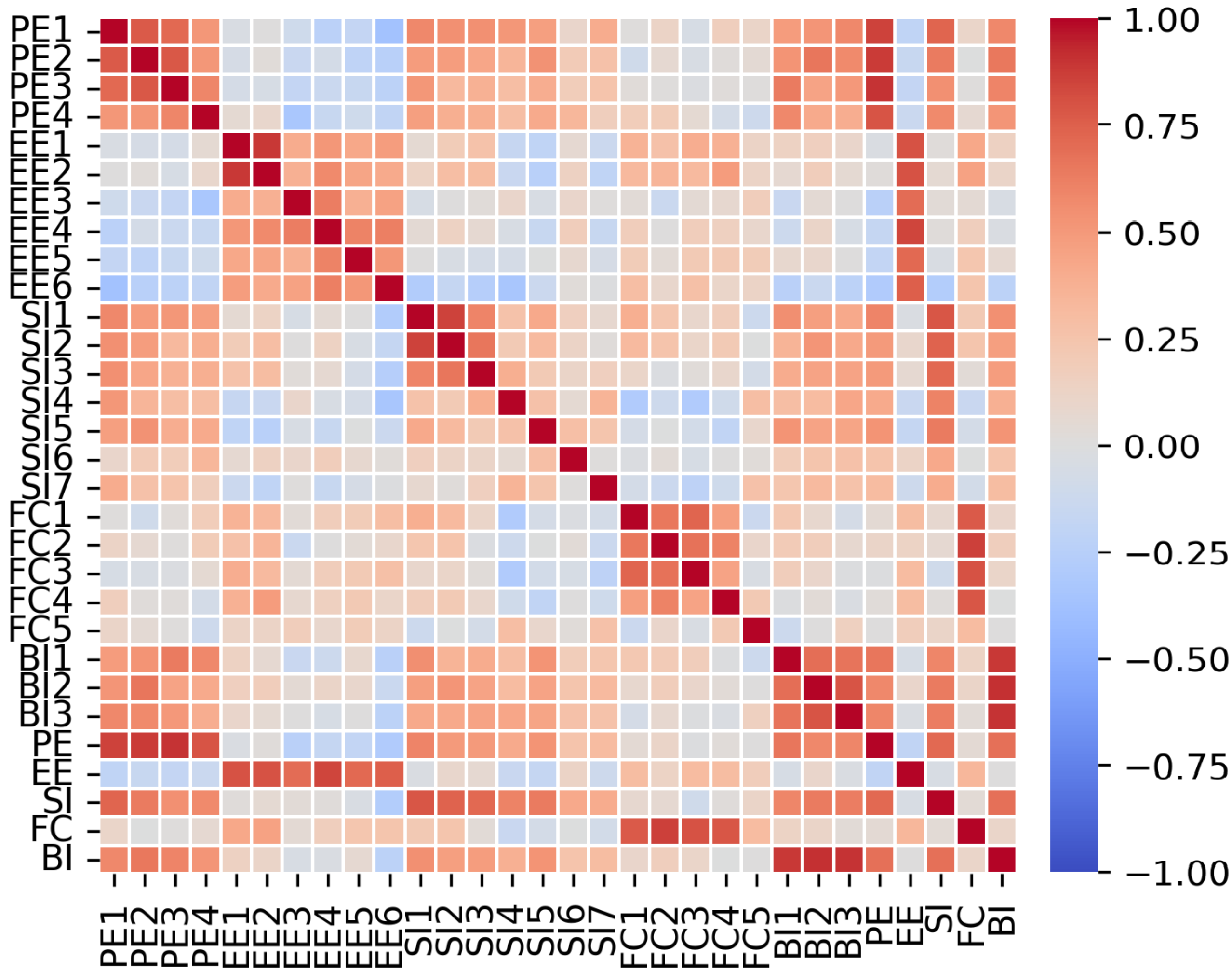

Figure 19: The correlation matrix with all statements from the survey including core constructs.

7.2.2.3 The Link Between Correlation Matrix and Structural Equation Modeling In this research, both the correlation matrix and Structural Equation Modelling (SEM) were employed to examine the relationships between the core constructs of the UTAUT model and the Behavioural Intention of the participants. The results from SEM revealed that Performance Expectancy had a substantial impact on Behavioural Intention, followed by Effort Expectancy and Social Influence, although to a lesser extent. Furthermore, the findings indicated that Facilitating Conditions (FC) had a negligible influence on Behavioural Intention.

The correlation matrix generally aligned with the SEM results, with Performance Expectancy showing the highest influence on Behavioural Intention. Additionally, the correlation matrix highlighted the significant impact of Social Influence on Behavioural Intention, which is consistent with the SEM analysis. Furthermore, the correlation matrix provided further insights into the relationships among specific variables, demonstrating the strong influence of statements related to their respective core constructs. Specifically, Performance Expectancy, Social Influence, and Behavioural Intention were found to be correlated with each other, while Effort Expectancy showed minimal correlation, except for with Facilitating Conditions.

8 Discussion and Limitations

This section relates the survey results back to the theory to discuss the most important findings, relating back to the research questions and hypotheses that were set up in Sections 3.1 and 3.2, respectively. In Section 7, three types of results were analysed: descriptive results, qualitative results, and the relationships among the results. The most important findings of each of these results will be discussed.

8.1 Discussion

As mentioned previously, three types of results were analysed. The individual question results and qualitative results will be discussed together, whereas the relationships among results will be discussed separately. Per type of analysis, the core constructs of the UTAUT model will be used to keep the discussion structured and logical, meaning the results can relate to each other.

8.1.1 Descriptive and Qualitative Results - Discussion

The survey was filled in by 62 people, of which 57 responses were valid. Out of these 57 responses, five participants were not student assistants and have not been in the past, making their contribution to the survey invalid, as this research primarily looks at the potential application for student assistants. In total, the responses of 52 participants ($n = 52$) were taken into account for the survey results.

8.1.1.1 Familiarity Figure 5 and Figure 6 show a brief visualisation of the familiarity levels among the participants. The results in Figure 5 relate to the general familiarity, which is measured through how often the participants use ChatGPT, whereas Figure 6 focuses solely on student assistant-related tasks. It becomes apparent that the results skew toward lower familiarity regarding general familiarity, whereas their experience using ChatGPT for student assistant-related tasks is barely done by the participants.

8.1.1.2 Suitability The participants were asked to give their opinion regarding suitability for tasks that the student assistants at Leiden University are often assigned to. The analysis of this qualitative question showed that there was a general negative consensus toward grading assignments and exams, both theoretical and practical. This applies to providing feedback for theoretical and practical assignments and exams as well, although there were more responses that claimed ChatGPT would be somewhat suitable for this. In the qualitative data, the participants mentioned the lacking ability to apply theoretical knowledge, pressed the need for human verification for grading and providing feedback, and outed their concerns regarding ChatGPT's reliability. As for developing ungraded study material, the general opinion showed to be more suitable for ChatGPT's assistance and the same applies to answering student questions in an asynchronous matter, such as through a forum. Contradictory, answering student questions synchronously, such as in class, was replied to more negatively.

Hypothesis 5: The survey participants have a more negative view of ChatGPT assisting with grading assignments and exams than they have with giving feedback for assignments and exams.

Hypothesis 5 states that it is expected that the survey participants have a more negative view of ChatGPT assisting with grading assignments and exams than they have with giving feedback for assignments and exams. In the results, this hypothesis seems to be true, since more students somewhat agreed or held a neutral consensus toward ChatGPT assisting with providing feedback for assignments and exams.

In the qualitative data, students often pressed the need for human verification to ensure the integrity of the feedback- and grading process and the reliability of ChatGPT's answers was questioned, as ChatGPT can generate confident, but faulty answers. The lack of sources and references was often mentioned as well, which can be crucial in the feedback- and grading process. When tools such as ChatGPT improve the sources and referencing capabilities or potentially introduce a 'confidence meter' (which was suggested in the qualitative data as well), it may improve the suitability regarding these tasks, although human verification is still expected to be required.

Hypothesis 8: The survey participants think that ChatGPT is unsuitable for the use of synchronous question answering (such as in class) whereas it is suitable for the use of asynchronous question answering (such as through a forum)

Hypothesis 8 stated that it was expected that the survey participants would think that synchronous question answering with the assistance of ChatGPT would be unsuitable, whereas asynchronous question answering would be more suitable. Based on the survey results, this hypothesis seems to be true.

In the qualitative data, it was mentioned that if student assistants were to use ChatGPT in class to answer questions, the students might as well interact with ChatGPT directly themselves, which would be the cause of this result. However, the effects of using ChatGPT by student assistants this way should be further investigated, as the student assistants might be able to contextualise the answers for the students to understand the material better.

Research Question 3: What are the perceptions of student assistants regarding the tasks that ChatGPT can assist with, based on their typically assigned tasks?

In the previous section, the figures visualised a split opinion with bimodal distributions regarding the suitability of ChatGPT for providing feedback on assignments and exams. The general opinion leaned towards considering these tasks as unsuitable, as the combined responses of "very unsuitable" and "somewhat unsuitable" outweighed the responses that indicated suitability. The qualitative data further pressed the need for human verification when providing feedback and raised concerns about the reliability of ChatGPT for these tasks.

Regarding grading, a likewise negative trend was observed, but to a greater extent. The support for using ChatGPT in grading exams and assignments was far weaker compared to providing feedback. The qualitative data showed the reasons behind this trend. Grading can be considered an integrity-sensitive task that might require consideration of external factors, such as effort and class participation. Additionally, ChatGPT's reliability for grading tasks was doubted which confirms the results of the survey.

Developing ungraded study material and answering student questions in an asynchronous way, such as through a forum, got more support. The assumption can be made that these tasks are considered less integrity-sensitive since they do not directly impact grades, which is the case for grading. However, synchronous question answering received a more negative response. The qualitative data claimed that if ChatGPT were to be used by student assistants to answer questions in class, it might be easier for the students to directly interact with ChatGPT. However, one participant mentioned that the student assistants could provide better context for ChatGPT's answer which could be beneficial to the understanding of students. The impact of using ChatGPT for synchronous question answering in a classroom setting should be researched in a more practical, field study.

Aside from specific tasks, the participants were also asked to give their opinion on courses that ChatGPT would be suitable to assist with. As described previously, the general findings were that introductory, practical courses such as "Algoritmiek", "Introduction to Programming", "Programming Methods", and "Databases" were mentioned more often than theoretical-based courses, such as "Introduction to Logic", "Research Methods in CS", "Studying and Presenting", "Orientation Informatics", and the mathematical courses "Calculus" and "Linear Algebra".

Hypothesis 6: The student assistants that participated in the survey believe that ChatGPT's assistance is more suitable for simple, practical courses such as introductory programming courses.

Hypothesis 6 stated that it was expected that the survey participants think that introductory, practical courses (such as first year programming courses) suit ChatGPT's assistance for student assistants better. The results of the open-ended question on suitability for courses in the Computer Science Bachelor programme are in line with the hypothesis, as the results often mention first-year practical courses related to introductory concepts of programming, such as in Python or C++. This verifies the sixth hypothesis that was made for this thesis.

Research Question 4: Are there specific courses in the Computer Science Bachelor Programme that student assistants believe could benefit from the integration of ChatGPT

The fourth research question aimed to determine if there was a preference for certain courses in the Computer Science Bachelor Programme that student assistants believe could benefit from the assistance of ChatGPT. The analysis of the qualitative data showed that there is greater support for its assistance in introductory programming courses compared to theoretical courses. Interestingly, most qualitative responses mentioned first-year courses, except for one second-year course, "Datastructures".

The more specific insights of the qualitative data showed that "Algoritmiek" was supported the most, followed by "Introduction to Programming," "Programming Methods," and "Databases.", which are all first-year courses in the Computer Science Bachelor Programme at Leiden University. On the other hand, the mentioned theoretical courses, which were given less support, included "Introduction to Logic," "Research Methods in CS," "Studying and Presenting," and "Orientation Informatics." A qualitative response did mention that ChatGPT often performs worse when asked to apply theoretical data, which could be a cause of these results, favouring practical courses over theoretical courses.

Notably, the implementation and use of ChatGPT as assisting tool in these courses require further research, especially looking at its performance in said courses. Nonetheless, the data suggest that there is generally greater support for the use of ChatGPT in introductory programming-related courses compared to theoretical courses.

8.1.1.3 Performance Expectancy The more in-depth results discussed in Section 7.1 regarding Performance Expectancy show that the participants in the survey have relatively high expectations in terms of performance about ChatGPT being an assisting tool for student assistants. Most participants expect that the adoption of ChatGPT will be useful in their work as student assistants. Similarly, they also expect it will boost their efficiency in terms of completing tasks, and additionally, a majority also expect it will boost their productivity. A relatively large part of the participants agrees with ChatGPT's potential in terms of assisting with difficult or impossible tasks, but more people are negative regarding its expectations for this task in comparison to the other statements.

The open-ended question on the potential integration of ChatGPT in the Performance Expectancy section showed that the majority of participants believe that ChatGPT can be integrated effectively. A large part of this subgroup specifically mentioned its potential for assisting with writing feedback and grading, but this subgroup did press the need for human verification which is still necessary. The lack of sources and references within ChatGPT's answers is also mentioned as a pitfall occasionally, together with a potential "confidence meter" that would improve ChatGPT's confidence in its generated answers.

Hypothesis 7: The survey participants worry about the lack of sources that ChatGPT provides with its generated answers.

Hypothesis 7 hypothesises that the survey participants will worry about the lack of sources that ChatGPT provides with its generated answers. This hypothesis was confirmed through the qualitative data where participants emphasised ChatGPT's lack of citing and referencing sources where the information was retrieved. Additionally, with ChatGPT occasionally generating faulty replies while coming over confidently, the need for a "confidence meter" might be a valuable addition before it can be utilised for integrity-sensitive tasks, such as grading assignments or exams.

8.1.1.4 Effort Expectancy The survey measured Effort Expectancy through six statements. The more in-depth analysis of these results was discussed in Section 7.1. This analysis shows that people hold the opinion that it does not take significant effort to become skilled at using ChatGPT for tasks related to being a student assistant. That does not take away the fact that there are still people who hold a neutral position and participants who think it takes significant effort.

According to the results, reformulating answers is among the tasks that are not expected to require significant effort, just like writing prompts. However, contextualising and fact-checking answers are tasks that are expected to take significant effort to become skilled in. Additionally, most people expect it does not require large amounts of effort to keep up with the innovations made by ChatGPT.

As for the open-ended question on the tasks that are expected to require the most effort, over a quarter of the total participants shared an answer to an open question, the results show that fact-checking is among the tasks that require the most effort, whereas keeping up with innovations will be made easier as the user-friendliness will increase as the tools improve. These results are in line with what was seen for the individual statements.

8.1.1.5 Social Influence Within the survey, Social Influence was measured through seven statements. Based on the in-depth findings that were discussed in Section 7.1, it shows that the participants have mixed feelings about Leiden University and lecturers accepting it as a tool to assist student assistants. The results also show positive opinions about other student assistants approving of the use of ChatGPT, which is similar for students and people outside of academia. In general, people also claim that others do not influence the participants' opinions on their decision to use ChatGPT or not in their work as a student assistant. Additionally, a large majority of people expect that the acceptance and use of ChatGPT will change over time.

Hypothesis 3: Opinions from influential individuals (such as lecturers, colleagues, or supervisors) and influential authorities (such as academic institutions) influence the likelihood of student assistants adopting ChatGPT.

The third hypothesis stated that it is expected that the opinions of influential individuals and authorities will influence the likelihood of student assistants adopting ChatGPT. The results showed that the participants are expecting less support from lecturers and Leiden University. However, they are expecting support from colleague student assistants and students. Furthermore, it is also expected that the opinion of others will not influence the participants' decision to use ChatGPT for student assistant-related tasks. Whether this is true in practice or not would have to be confirmed through future studies, e.g. through the "Use Behaviour" core construct from the UTAUT model.

The main takeaways regarding the open-ended question on how the participants think opinions on ChatGPT will change over time resulted in a vast majority claiming it will improve in the near future. The newer large language models could enhance acceptance and shift opinions positively, while some participants also mentioned that the role of companies and educational institutions might influence the opinion on ChatGPT and future similar technologies.

8.1.1.6 Facilitating Conditions Regarding expectations for Facilitating Conditions, people generally expect Leiden University to provide training, support, licenses, and policies and regulations regarding using ChatGPT to assist student assistants in their work. One statement asked whether the participants expect Leiden University to ensure compatibility between systems, which resulted in a mixed response.

The open-ended question regarding the expectancies of training on the use of ChatGPT showed that the largest group of participants expects training on when and how to use ChatGPT effectively. Among these responses, ethical considerations were also mentioned. Additional participants mentioned that training on how to contextualise and fact-check answers should be addressed, which is in line with the expectancy of those tasks requiring significant effort to become skilled at. Some participants also mentioned training to assist in better prompt writing, which was also mentioned as a task that required some effort to become skilled at.

8.1.1.7 Behavioural Intention As part of the UTAUT model, Behavioural Intention is used to measure whether people will use technology. In this study, it is based on the previous core constructs. It was measured through three statements, which indicated that the student assistants that participated in the survey are likely to use the tool for their work. These participants also expect to become motivated at using ChatGPT for their work and expect to explore it in different ways if they were to use it for their work as student assistants.

8.1.2 The relationship among core constructs

The UTAUT framework served as a theoretical basis in this study for examining the relationships between the core constructs and Behavioural Intention regarding the use of ChatGPT as an assisting tool for student assistants. The core constructs were individually measured and their relationships with Behavioural Intention were measured.

8.1.2.1 Structural Equation Modeling This was done through Structural Equation Modeling (SEM), which has been used in various similar studies [1] [33]. It is worth noting that, according to a handbook on SEM [13], the minimum recommendation of 20 responses per statement was not met. To address this limitation and improve the various model-fit measures for SEM, Confirmatory Factor Analysis (CFA) was performed to improve the data to fit the model and ensure the suitability to assess the relationships among core constructs through SEM.

Hypothesis 1: Student assistants who expect ChatGPT to be capable of improving performance in their tasks are more likely to consider its adoption.

The first hypothesis stated that student assistants who are expecting ChatGPT to be capable of improving performance in their tasks are more likely to consider its adoption. This hypothesis addresses the importance of Performance Expectancy on Behavioural Intention. Based on the findings through Structural Equation Modeling, Performance Expectancy was found to have the highest influence on Behavioural Intention among all core constructs, both with and without performing Confirmatory Factor Analysis.

In addition, based on the correlation matrix, it became apparent that PE2 influenced Behavioural Intention the most, followed by PE1, PE3, and lastly PE4. This shows that improving the increased speed at which tasks can be accomplished is of the greatest importance to the Behavioural Intention, followed by ChatGPT's usefulness and the increase in productivity while using ChatGPT. Lastly, ChatGPT's ability to efficiently complete tasks that are considered difficult or impossible also influences Behavioural Intention, which showed mixed results in its individual analysis.

Hypothesis 2: Effort expectancy has a minimal impact on the likelihood of adopting ChatGPT as an assisting tool for student assistants, given its already relatively simple usage.

Hypothesis 2 states its expectancy that required effort will have minimal impact on the likelihood of adopting ChatGPT as an assisting tool for student assistants, given its already relatively simple usage. After analysing the more specific tasks that ChatGPT will be used for and the required effort it takes to do prompt engineering, fact-checking, reformulating, and contextualising answers, it became clear that some of these tasks are expected to require some effort to become skilled at, especially fact-checking ChatGPT's answers. Based on the Structural Equation Modeling analysis, it showed that Effort Expectancy has some influence on Behavioural Intention, but not as significant as the influence of Performance Expectancy. The analysis of the correlation matrix showed a somewhat similar result, although the strength of the relationship was somewhat weaker. An outlier was EE6, which mentioned that keeping up with the innovations from ChatGPT will take significant effort.

After excluding certain statements through Confirmatory Factor Analysis, SEM indicated that Performance Expectancy plays a crucial role in shaping participants' Behavioural Intention. Effort Expectancy and Social Influence also contribute to the Behavioural Intention of the participants, although their impact is not as significant as that of Performance Expectancy. Additionally, Facilitating Conditions exhibit a positive relationship with Behavioural Intention, which can be neglected, as the score is close to zero.

Hypothesis 3: Opinions from influential individuals (such as lecturers, colleagues, or supervisors) and influential authorities (such as academic institutions) influence the likelihood of student assistants adopting ChatGPT.

Although the third hypothesis was already briefly discussed previously, the opinions from influential individuals and authorities were shown to be similarly correlated to Effort Expectancy, the correlation matrix showed a slightly different result, which indicated a weaker correlation than Performance Expectancy had to Behavioural Intention. This highlighted SI1 and SI5 to be most influential to Behavioural Intention, indicating that the approval of Leiden University influences Behavioural Intention and that the acceptance by people outside of academia also influences Behavioural Intention to some degree. Additionally, SI2, SI3, and SI4 were also important, although to a lesser extent than SI1 and SI5, highlighting that the approval by lecturers is also somewhat important to the Behavioural Intention of the participants, together with the approval of other student assistants and students.

Hypothesis 4: Facilitation of the use of ChatGPT as an assisting tool for student assistants have minimal influence on the adoption of the technology, considering its relative ease of use.

The impact of facilitation of the use of ChatGPT as an assisting tool for student assistants from Leiden University was expected to be minimal. Although the individual questions showed general support for introducing policies and regulations, but also showed expectancies in terms of support, licensing and training provided by Leiden University. However, according to the Structural Equation Modeling analysis and the correlation matrix, it became clear that the Facilitating Conditions hold a negligible relationship with Behavioural Intention. This implies that regardless of the expectancies in terms of facilitation by Leiden University, Behavioural Intention does not change. This does not change the fact that the survey participants agreed that facilitation in several areas is necessary according to the results from the individually analysed data.

These findings provide insights into the determining factors of Behavioural Intention within the UTAUT model regarding the integration of ChatGPT as an assisting tool for student assistants. As shown, Performance Expectancy emerges as a primary factor that influences Behavioural Intention, followed by Effort Expectancy and Social Influence impacting Behavioural Intention as well, but to a lesser extent. Facilitating Conditions contributes to a minimal extent, making its contribution negligible.

8.1.2.2 Correlation Matrix With the limitations of Structural Equation Modeling for interpretation because of sample size, a correlation matrix was also used to provide insights into the relationships among the statements within each core construct. As expected, statements within the same core constructs show correlation. For instance, the statements PE1, PE2, PE3, and PE4 demonstrate mutual influence, which is logical given these statements belong to the same core construct. Similar patterns can be seen for the statements within Effort Expectancy, Social Influence, Facilitating Conditions, and Behavioural Intention.

The statements that belong to Performance Expectancy show correlations with the overall construct of Behavioural Intention, which shows the influence of performance expectations on Behavioural Intention. This finding is consistent with the results from Structural Equation Modeling. Moreover, Performance Expectancy also influences Social Influence to some degree, meaning that expectations in terms of performance influence how big of a role Social Influence plays in participants using ChatGPT for student assistant-related tasks.

Research Question 1: What factors are important for student assistants when considering the adoption of ChatGPT as an assisting tool? How do factors such as performance and required effort influence the intention to adopt a technology?

The first research question of this thesis aimed to determine the important factors for student assistants when considering the adoption of ChatGPT as an assisting tool, particularly focusing on whether performance and required effort influence the intention to adopt this technology. This aligns with the constructs of the UTAUT model Performance Expectancy and Effort Expectancy.

The Structural Equation Modelling (SEM) analysis and correlation matrix have shown that Performance Expectancy is the most influential factor on Behavioural Intention by quite a margin. This result was similar both before and after performing the confirmatory factor analysis and removing statements that would not fit the model well. The results also align with the correlation matrix.

Examining the correlation matrix in more depth, it can be observed that PE2, which measured the expectancy of accomplishing tasks more quickly with the assistance of ChatGPT and has shown that it has a strong influence on general Behavioural Intention. The results generally show that the student assistants who took part in the survey expect to value the overall usefulness, increased productivity, and efficient problem-solving capabilities of ChatGPT.

Effort Expectancy has demonstrated a less influential role in Behavioural Intention. While the relationship before the Confirmatory Factor Analysis was somewhat weaker compared to the relationship after this analysis, the overall influence of Effort Expectancy on Behavioural Intention remains notably weaker than that of Performance Expectancy. The SEM analysis showed that Effort Expectancy has a relatively small influence on Behavioural Intention, and the correlation matrix suggested that its influence on Behavioural Intention is negligible.

However, based on the findings related to participants' expectations regarding training on specific tasks, which were also explored in the Effort Expectancy section (SI2 through SI5), it is possible to assume that student assistants anticipate receiving training in the future. This expectation may explain a reduced concern regarding the required effort, as the participants expect Leiden University to provide them with training to mitigate potential challenges regarding the required effort. It should be said that this assumption requires further investigation in future studies, as the current study serves more of a preliminary exploration of the influential factors.

Social Influence seems to have a general influence on Behavioural Intention, which also aligns with SEM findings. This implies that the participants' perception of Social Influence plays a role in shaping their Behavioural Intention. Additionally, Effort Expectancy and Facilitating Conditions show a relatively weak correlation with the statements that measured Behavioural Intention. This is partially in line with the findings from SEM, which indicates that FC had very little impact on Behavioural Intention, while Effort Expectancy had far less influence on Behavioural Intention, which does not seem to be similar in the correlation matrix.

Overall, the findings from the correlation matrix support the results obtained from the Structural Equation Modeling technique mostly. This helps to use the foundation of this research, the UTAUT model, to figure out what is important for the adoption of this technology. These findings also found a lesser correlation between Effort Expectancy and Behavioural Intention, which is contradictory to the SEM findings and may be a result of the incorrect sample size. These results also show that SEM was carried out most successfully, through Confirmatory Factor Analysis which resulted in removing certain statements, making the model-fit measures mostly within appropriate ranges.

Research Question 2: To what extent do external influences, such as the opinions of others and the facilitation provided by the willingness to adopt ChatGPT as an assisting tool?

The second research of this thesis aimed to explore the influence of external factors, such as the opinions of others and the facilitations provided by Leiden University, on the readiness of student assistants to adopt ChatGPT as an assisting tool. This question is in line with the Social Influence construct and the Facilitating Conditions construct of the UTAUT model.

The results that were achieved from the SEM analysis and correlation matrix shows that Social Influence holds a similar level of importance in influencing Behavioural Intention as Effort Expectancy. This similarity is observed in the scores both before and after performing the Confirmatory Factor Analysis, with Social Influence slightly outscoring Effort Expectancy. However, this difference is negligible.

The correlation matrix has revealed a stronger correlation between Social Influence and Behavioural Intention compared to the correlation between Effort Expectancy and Behavioural Intention. As mentioned in the discussion regarding hypothesis 3, outliers such as SI1 and SI5 mention the influence of approval by Leiden University and acceptance by individuals outside of academia, whereas SI2, SI3, and SI4 mention the approval of lecturers, colleague student assistants, and students which hold importance in the stance on Behavioural Intention.

With the individual results regarding the influence of others' opinions on the use of ChatGPT, it became clear that among the survey respondents, the majority claimed it would not let the opinion of others influence their decision to use ChatGPT as a student assistant. However, whether this is the case in real-world situations would have to be further investigated, as that was not part of the study.

The results that were obtained from the SEM analysis and correlation matrix regarding Facilitating Conditions show a negligible relationship with Behavioural Intention. This result was obtained both before and after the Confirmatory Factor Analysis. The individual data which was discussed beforehand showed certain expectancies concerning training, support, licensing, and policies and regulations for the use of ChatGPT for student assistants. However, the experiences related to its compatibility with university systems were relatively lower, which could be caused by the stand-alone nature of ChatGPT. Despite these findings in the individual data, the Facilitating Conditions do not influence Behavioural Intention.

This observation may be related to the expectancy in terms of performance. While Facilitating Conditions may not be important for the Behavioural Intention of ChatGPT as an assisting tool, they might be seen as additional support to enhance the effectiveness of using ChatGPT in this situation. The individual data of Performance Expectancy indicating a left-skewed distribution in general responses further supports the assumption that Facilitating Conditions may be seen as optional 'nice-to-have' additions, rather than essential requirements for working with ChatGPT.

8.2 Limitations

Within this study, it is important to acknowledge the limitations as they can affect the reliability of the findings. Throughout the study, several limitations have been identified and should be taken into consideration when interpreting the results.

Firstly, the survey was distributed to students at Leiden University that are studying Computer Science, which introduces the potential for sampling bias. Since all participants are from the same university, the findings may not be suitable or representative for interpretation of the broader population, such as other universities and/or other studies. This also relates to the tasks that the participants think ChatGPT is suitable or unsuitable for, as these tasks may differ per university, faculty, or study. The characteristics and experiences may differ from those of students at other universities.

Secondly, the survey did not collect any demographic data such as age, gender, level of education, experience as a student assistant, courses in which they are assisting or have assisted, or educational background. These variables could influence participants' perceptions and behaviours related to the integration of ChatGPT. Although the demographic data did not add value to the research question answers and the effect of leaving out the demographic data was minimised due to the homogeneity of the target population, this limitation should be noted.

Another limitation is that the study did not measure Use Behaviour, which is an additional core construct used in other studies using the UTAUT model [1] [33]. Measuring the Use Behaviour of ChatGPT and assessing its effect through the UTAUT constructs could provide a better understanding of their Behavioural Intention. This could be achieved through a control group, which would complete a follow-up survey to determine Use Behaviour and assess the technology's impact on the various core constructs of the UTAUT model. However, this aspect was beyond the scope of this thesis.

Furthermore, the results obtained by Structural Equation Modeling require careful interpretation, as SEM typically requires a relatively large sample size to ensure reliable results [13]. As observed in previous studies, [1] [33] larger sample sizes are used to meet this criterion. Although this study made efforts to assess model fitness and improve the model fitness through Confirmatory Factor Analysis, and scoring within the proper ranges on four different model-fitness indices, the results should be carefully interpreted.

Since various other frameworks are available for measuring technology acceptance, the survey design is quite limited to the framework chosen, which is the UTAUT model in this study. The use of different frameworks might find different results, while adding additional core constructs, might yield more surprising results and may change the strength of relationships between the core constructs and Behavioural Intention.

The survey's findings may be influenced by the specific period in which the data was collected. With the advancements made in the general field of Generative AI, the study may achieve completely different results in a year, perhaps even months from the point of writing. As people become more experienced with similar tools and as these tools become more skilled at their tasks, the results may vary, meaning this research is time-bound.

Lastly, the survey primarily collects quantitative data based on the core constructs. This is typical of the UTAUT model [35]. Although some efforts were made to allow participants to share their thoughts on certain parts of the core constructs, not every participant shared their opinion through the open-ended questions. Different qualitative methods, such as interviews or focus groups could provide additional, valuable data to understand the data in a better, more in-depth way.

9 Conclusion and Future Work

This section discusses the conclusion that can be drawn from this research and outlines future work regarding the exploration of integrating ChatGPT as an assisting tool for student assistants. The conclusion section summarises the key findings and insights obtained from the analysis, including the significant role of the core constructs of the UTAUT model. Additionally, the future work subsection identifies potential valuable areas for further investigation, such as studies that also measure Use Behaviour.

9.1 Conclusion

The objective of this thesis was to explore the potential integration of ChatGPT as an assisting tool for student assistants in Computer Science Education using the Unified Theory of Acceptance and Use of Technology (UTAUT). The findings that were revealed during this research between the core constructs of the UTAUT model and Behavioural Intention provide answers to the research questions of this study.

Firstly, the Structural Equation Modeling analysis and correlation matrix have shown that Performance Expectancy emerged as the most influential factor in Behavioural Intention. The individual data indicated that participants held relatively high expectations in terms of performance, including overall usefulness, improved efficiency, and increased productivity. Effort Expectancy, although influential to a lesser extent, still played a role in determining Behavioural Intention. The individual data displayed more mixed responses regarding the required effort for specific tasks, with a higher perceived effort to become skilled at fact-checking ChatGPT's answers.

Secondly, the Structural Equation Modeling analysis explored the external influences on Behavioural Intention, such as Social Influence and Facilitating Conditions. Social Influence was found to hold similar significance in determining Behavioural Intention as Effort Expectancy. It became apparent that the approval of Leiden University, lecturers, individuals outside of academia, as well as students and other student assistants, came up as influential factors. The Facilitating Conditions highlighted a clear need for support from the university in terms of required facilitation which includes licensing, general support, training, and setting up policies and regulations. However, the relationship analysis showed that these expectancies did not directly impact the willingness to use ChatGPT.

Thirdly, regarding the tasks that are typically assigned to student assistants at Leiden University in Computer Science, the study demonstrated limited support for ChatGPT's assistance in grading and providing feedback, with slightly more support for providing feedback. The qualitative data indicated that this is currently caused by concerns related to ChatGPT's false confidence and reliability. However, the inclusion of human verification for integrity-sensitive tasks, such as grading and providing feedback for assessments and exams could potentially improve the reliability. Moreover, answering student questions synchronously did not receive support, whereas answering questions asynchronously was supported by the survey participants. Additionally, the survey participants showed support for the assistance of ChatGPT with the development of ungraded study material.

Lastly, the study revealed that ChatGPT's assistance received more expected support in introductory, practical courses that have a focus on programming concepts, such as introductory courses to programming languages, for instance, Python and C++, but also an introductory course to databases. The support for theoretical and mathematical courses, which are also part of the Computer Science Bachelor curriculum at Leiden University, received considerably less support.

By specifically focusing on student assistants, this research revealed the expectations regarding the use of ChatGPT in their work. As potential future lecturers, their perspectives are significant in shaping the future of education with these continually improving tools. In an era where such tools are widely accessible, it is important to explore how they can be effectively used for educational purposes. This research contributes to embracing AI in education and encourages exploration of the field in more detail for the possibilities and challenges associated with AI integration in education.

9.2 Future Work

This research opens up several paths for future research that can build from the findings of this study to further advance the understanding of integrating ChatGPT as an assisting tool for student assistants. The following areas can be further explored in this topic.

Firstly, exploring Use Behaviour is crucial to assess whether the Behavioural Intention translates into real-world usage. Investigating the impact of Behavioural Intention on Use Behaviour, as well as examining the influence of other core constructs on Use Behaviour will provide additional insights into the practical implementation of ChatGPT as an assisting tool and its effectiveness.

Furthermore, while this study focused on student assistants at Leiden University, specifically within the field of Computer Science, future research can enlarge the scope to include student assistants from different universities. This broader sampling technique will improve the generalisability of the findings by offering different perceptions and expectations of a more diverse population. Future research can also be expanded to different fields, as the field of Computer Science might yield different results than the field of Psychology, Physics or others.

Lastly, given the rapidly evolving nature of generative AI and the continuous advancement in language models, it will be valuable to monitor and measure the changes in results over time. Tracking the acceptance, use, and perceptions of ChatGPT among student assistants and other groups in the upcoming months and years can provide insights into the potential of integrating AI assistance in educational, and academic settings.

References

[1] Muneer M. M. Abbad. "Using the UTAUT model to understand students' usage of e-learning systems in developing countries". In: *Education and Information Technologies* 26.6 (Nov. 2021), pp. 7205–7224. DOI: 10.1007/s10639-021-10573-5. URL: https://doi.org/10.1007/s10639-021-10573-5.

[2] Patrick Ajibade. "Technology Acceptance Model Limitations and Criticisms: Exploring the Practical Applications and Use in Technology-related Studies, Mixed-method, and Qualitative Researches". In: (2018).

[3] Jill Anderson. *Harvard EdCast: Educating in a World of Artificial Intelligence.* https://www.gse.harvard.edu/news/23/02/harvard-edcast-educating-world-artificial-intelligence. Accessed: 2023-04-26. 2023.

[4] Som S. Biswas. "Potential Use of Chat GPT in Global Warming". In: *Annals of Biomedical Engineering* 51.6 (June 2023), pp. 1126–1127. ISSN: 1573-9686. DOI: 10.1007/s10439-023-03171-8. URL: https://doi.org/10.1007/s10439-023-03171-8.

[5] "Computer Self-Efficacy: Development of a Measure and Initial Test". In: *MIS Quarterly* 19.2 (1995), pp. 189–211. ISSN: 02767783. URL: http://www.jstor.org/stable/249688 (visited on 05/23/2023).

[6] Mohammad Shams Ud Duha. "ChatGPT in Education: An Opportunity or a Challenge for the Future?" In: *TechTrends* (Apr. 2023). ISSN: 1559-7075. DOI: 10.1007/s11528-023-00844-y. URL: https://doi.org/10.1007/s11528-023-00844-y.

[7] Yi Fan et al. "Applications of structural equation modeling (SEM) in ecological studies: an updated review". In: *Ecological Processes* 5.1 (Nov. 2016), p. 19. ISSN: 2192-1709. DOI: 10.1186/s13717-016-0063-3. URL: https://doi.org/10.1186/s13717-016-0063-3.

[8] Mohammadreza Farrokhnia et al. "A SWOT analysis of ChatGPT: Implications for educational practice and research". In: *Innovations in Education and Teaching International* (2023), pp. 1–15. DOI: 10.1080/14703297.2023.2195846. eprint: https://doi.org/10.1080/14703297.2023.2195846. URL: https://doi.org/10.1080/14703297.2023.2195846.

[9] Suzanne Fergus, Michelle Botha, and Mehrnoosh Ostovar. "Evaluating Academic Answers Generated Using ChatGPT". In: *Journal of Chemical Education* 100.4 (2023), pp. 1672–1675. DOI: 10.1021/acs.jchemed.3c00087. URL: https://doi.org/10.1021/acs.jchemed.3c00087.

[10] Paul R. Warshaw Fred D. Davis Richard P. Bagozzi. "User Acceptance of Computer Technology: A Comparison of Two Theoretical Models". In: *Management Science* 35 (1989), pp. 982–1003.

[11] Chong Guan, Jian Mou, and Zhiying Jiang. "Artificial intelligence innovation in education: A twenty-year data-driven historical analysis". In: *International Journal of Innovation Studies* 4.4 (2020), pp. 134–147. ISSN: 2096-2487. DOI: https://doi.org/10.1016/j.ijis.2020.09.001. URL: https://www.sciencedirect.com/science/article/pii/S2096248720300369.

[12] Joseph Hair et al. *Multivariate Data Analysis: A Global Perspective*. Mar. 2010. ISBN: 0135153093.

[13] Rick Hoyle and Cahyono St. *Handbook of Structural Equation Modeling*. July 2022. ISBN: ISBN 978-1-60623-077-0 (hardback : acid-free paper).

[14] Yijia Lin Jianyang Deng. "The Benefits and Challenges of ChatGPT: An Overview". In: *Frontiers in Computing and Intelligent Systems* 2.2 (2022). ISSN: 2832-6024. DOI: 10.54097/fcis.v2i2.4465. URL: https://www.researchgate.net/publication/367106604_The_Benefits_and_Challenges_of_ChatGPT_An_Overview.

[15] PC Lai. "The Literature Review of Technology Adoption Models and Theories for The Novelty Technology". In: *Journal of Information Systems and Technology Management* 14.1 (2017), pp. 21–38.

[16] Yi-Hsuan Lee, Yi-Chuan Hsieh, and Chia-Ning Hsu. "Adding Innovation Diffusion Theory to the Technology Acceptance Model: Supporting Employees' Intentions to use E-Learning Systems". In: *Journal of Educational Technology & Society* 14.4 (2011), pp. 124–137. ISSN: 11763647, 14364522. URL: http://www.jstor.org/stable/jeductechsoci.14.4.124 (visited on 05/23/2023).

[17] Chung Kwan Lo. "What Is the Impact of ChatGPT on Education? A Rapid Review of the Literature". In: *Education Sciences* 13.4 (2023). ISSN: 2227-7102. DOI: 10.3390/educsci13040410. URL: https://www.mdpi.com/2227-7102/13/4/410.

[18] Aleksandra Luszczynska and Ralf Schwarzer. "Changing Behavior Using Social Cognitive Theory". In: *The Handbook of Behavior Change*. Cambridge Handbooks in Psychology. Cambridge University Press, 2020, pp. 32–45. DOI: 10.1017/9781108677318.003.

[19] Lyndsey Middleton, Hazel Hall, and Robert Raeside. "Applications and applicability of Social Cognitive Theory in information science research". In: *Journal of Librarianship and Information Science* 51.4 (2019), pp. 927–937. DOI: 10.1177/0961000618769985. eprint: https://doi.org/10.1177/0961000618769985. URL: https://doi.org/10.1177/0961000618769985.

[20] Gary C. Moore and Izak Benbasat. "Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation". In: *Inf. Syst. Res.* 2 (1991), pp. 192–222.

[21] Tolorunleke Caroline Aduke Opara Emmanuel Chinonso Adalikwu Mfon-Ette Theresa. "ChatGPT for Teaching, Learning and Research: Prospects and Challenges". In: *Global Academic Journal of Humanities and Social Sciences* 5.2 (2023), pp. 33–40. DOI: 10.36348/gajhss.2023.v05i02.001. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4375470.

[22] Juliana E. Raffaghelli et al. "Applying the UTAUT model to explain the students' acceptance of an early warning system in Higher Education". In: *Computers & Education* 182 (2022), p. 104468. ISSN: 0360-1315. DOI: https://doi.org/10.1016/j.compedu.2022.104468. URL: https://www.sciencedirect.com/science/article/pii/S0360131522000392.

[23] B. et al. Rahimi. "A Systematic Review of the Technology Acceptance Model in Health Informatics". In: *Applied clinical informatics* 9.3 (2018), pp. 604–634. DOI: 10.1055/s-0038-1668091. URL: https://pubmed.ncbi.nlm.nih.gov/30112741/.

[24] Md. Mostafizer Rahman and Yutaka Watanobe. "ChatGPT for Education and Research: Opportunities, Threats, and Strategies". In: *Applied Sciences* 13.9 (2023). ISSN: 2076-3417. DOI: 10.3390/app13095783. URL: https://www.mdpi.com/2076-3417/13/9/5783.

[25] Waleed Mugahed Al-Rahmi et al. "Integrating innovation diffusion theory with technology acceptance model: supporting students' attitude towards using a massive open online courses (MOOCs) systems". In: *Interactive Learning Environments* 29.8 (2021), pp. 1380–1392. DOI: 10.1080/10494820.2019.1629599. eprint: https://doi.org/10.1080/10494820.2019.1629599. URL: https://doi.org/10.1080/10494820.2019.1629599.

[26] Vanessa Ratten. "Social Cognitive Theory and the Technology Acceptance Model in the Cloud Computing Context: The Role of Social Networks, Privacy Concerns and Behavioural Advertising". In: *Adoption of Innovation: Balancing Internal and External Stakeholders in the Marketing of Innovation*. Ed. by Alexander Brem and Éric Viardot. Cham: Springer International Publishing, 2015, pp. 43–56. ISBN: 978-3-319-14523-5. DOI: 10.1007/978-3-319-14523-5_4. URL: https://doi.org/10.1007/978-3-319-14523-5_4.

[27] Yves Rosseel. "lavaan: An R Package for Structural Equation Modeling". In: *Journal of Statistical Software* 48.2 (2012), pp. 1–36. DOI: 10.18637/jss.v048.i02. URL: https://www.jstatsoft.org/index.php/jss/article/view/v048i02.

[28] Ronny Scherer, Fazilat Siddiq, and Jo Tondeur. "The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education". In: *Computers & Education* 128 (2019), pp. 13–35. ISSN: 0360-1315. DOI: https://doi.org/10.1016/j.compedu.2018.09.009. URL: https://www.sciencedirect.com/science/article/pii/S0360131518302458.

[29] Artur Strzelecki. "To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology". In: *Interactive Learning Environments* 0.0 (2023), pp. 1–14. DOI: 10.1080/10494820.2023.2209881. eprint: https://doi.org/10.1080/10494820.2023.2209881. URL: https://doi.org/10.1080/10494820.2023.2209881.

[30] Hamed Taherdoost. "A review of technology acceptance and adoption models and theories". In: *Procedia Manufacturing* 22 (2018). 11th International Conference Interdisciplinarity in Engineering, INTER-ENG 2017, 5-6 October 2017, Tirgu Mures, Romania, pp. 960–967. ISSN: 2351-9789. DOI: https://doi.org/10.1016/j.promfg.2018.03.137. URL: https://www.sciencedirect.com/science/article/pii/S2351978918304335.

[31] Christian Terwiesch. "Would Chat GPT Get a Wharton MBA? A Prediction Based on Its Performance in the Operations Management Course". In: (Jan. 2023). URL: https://mackinstitute.wharton.upenn.edu/wp-content/uploads/2023/01/Christian-Terwiesch-Chat-GTP-1.24.pdf.

[32] James O. Abiola Thomas Olushola. "The Efficacy of Technology Acceptance Model: A Review of Applicable Theoretical Models in Information Technology Researches". In: *Journal of Research in Business and Management* 4.11 (2017), pp. 70–83.

[33] Nenad Tomić, Zoran Kalinić, and Violeta Todorović. "Using the UTAUT model to analyze user intention to accept electronic payment systems in Serbia". In: *Portuguese Economic Journal* 22.2 (May 2023), pp. 251–270. ISSN: 1617-9838. DOI: 10.1007/s10258-022-00210-5. URL: https://doi.org/10.1007/s10258-022-00210-5.

[34] UNESCO. *Artificial intelligence in education*. https://www.unesco.org/en/digital-education/artificial-intelligence. Accessed: 2023-04-26.

[35] Gordon B. Davis Viswanath Venkatesh Michael G. Morris. "User Acceptance of Information Technology: Toward a Unified View". In: (2003).

[36] Ricardo Herrera Martinez Yvette Mucharraz y Cano Francesco Venuti. "ChatGPT and AI Text Generators: Should Academia Adapt or Resist?" In: (Feb. 2023). URL: https://www.hbsp.harvard.edu/inspiring-minds/chatgpt-and-ai-text-generators-should-academia-adapt-or-resist.

[37] Ke Zhang and Ayse Begum Aslan. "AI technologies for education: Recent research & future directions". In: *Computers and Education: Artificial Intelligence* 2 (2021), p. 100025. ISSN: 2666-920X. DOI: https://doi.org/10.1016/j.caeai.2021.100025. URL: https://www.sciencedirect.com/science/article/pii/S2666920X21000199.

A Visualised Survey Data

A.1 Suitability

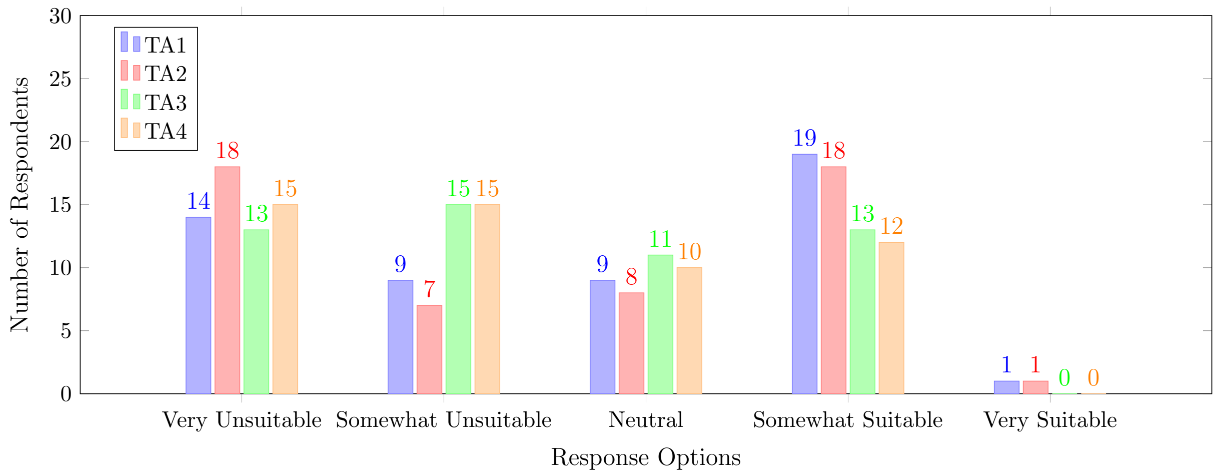

Figure 20: Results for suitability providing feedback for theoretical and practical assignments and exams

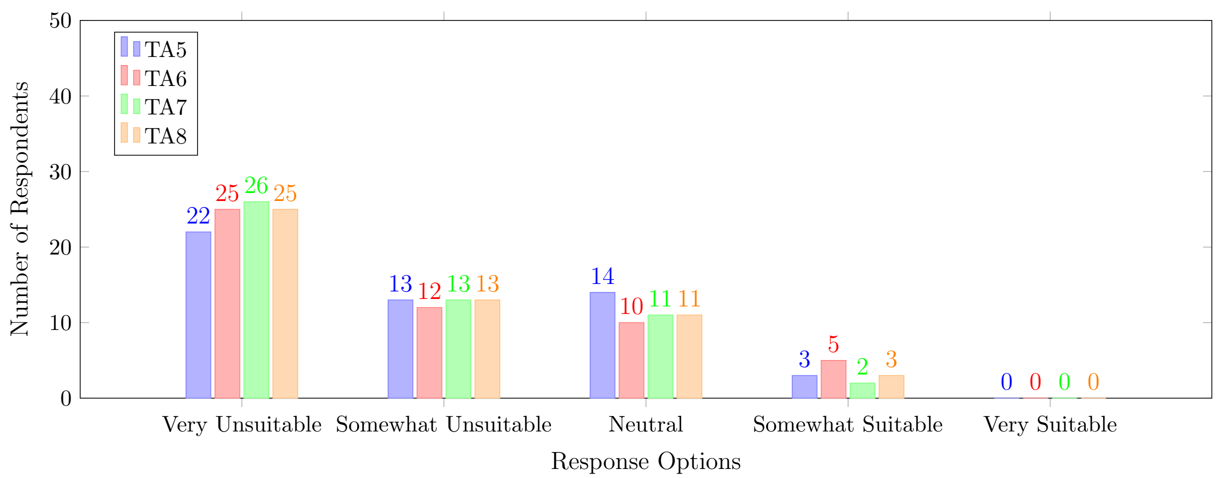

Figure 21: Results for suitability grading theoretical and practical assignments and exams

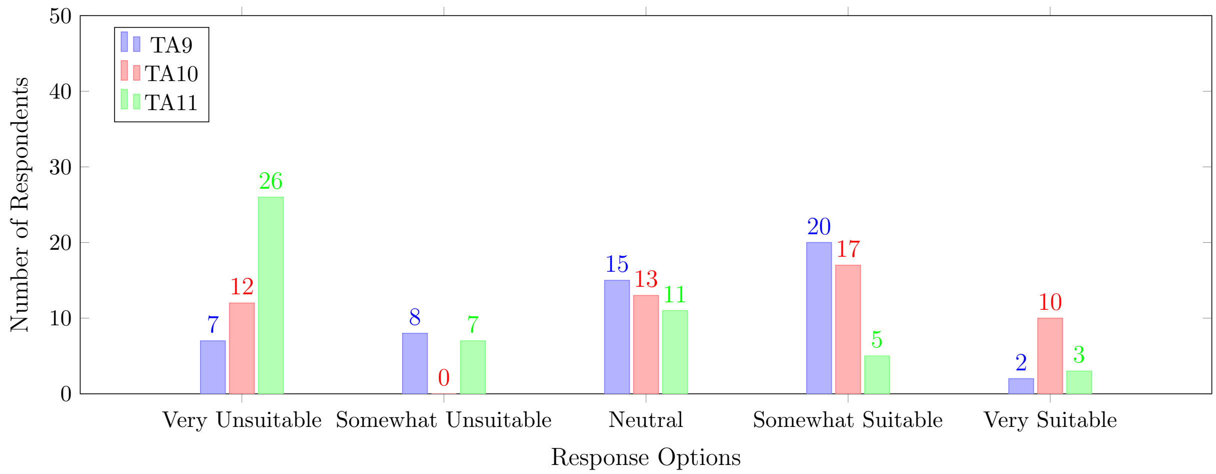

Figure 22: Results for suitability developing ungraded material and answering student questions

A.2 Performance Expectancy

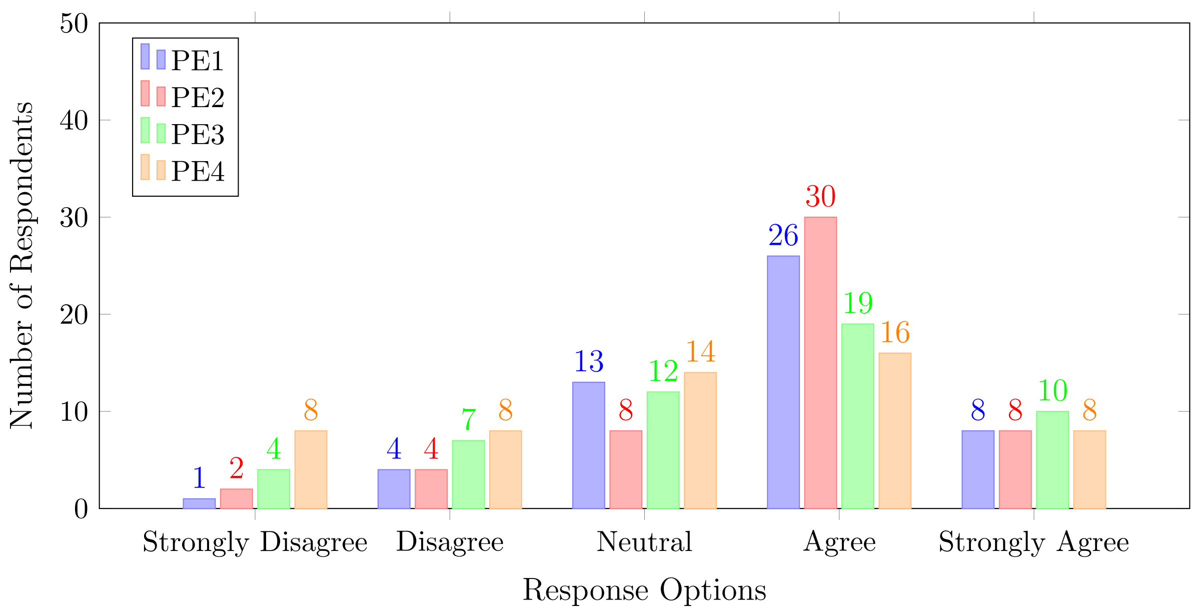

Figure 23: Performance Expectancy results

A.3 Effort Expectancy

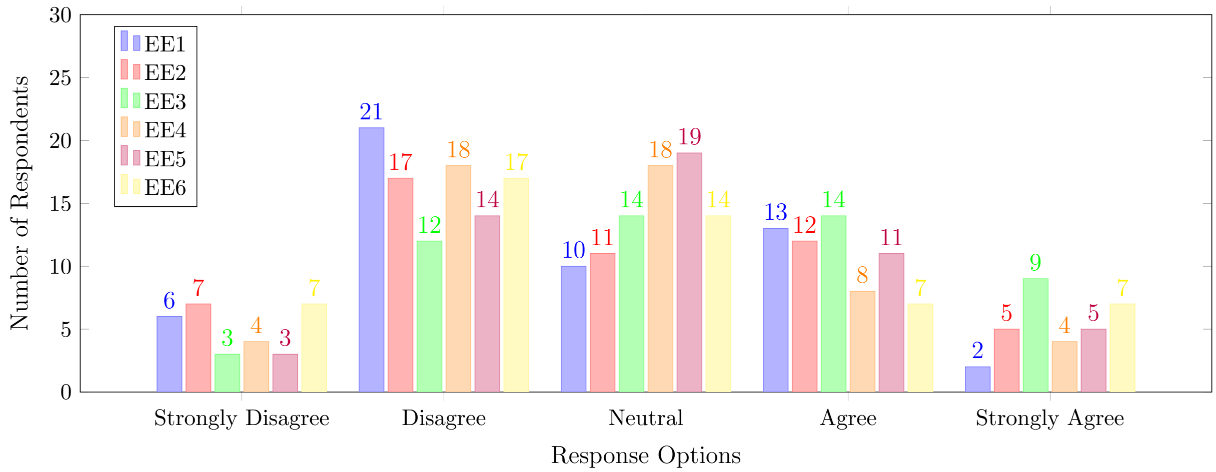

Figure 24: Effort Expectancy results

A.4 Social Influence

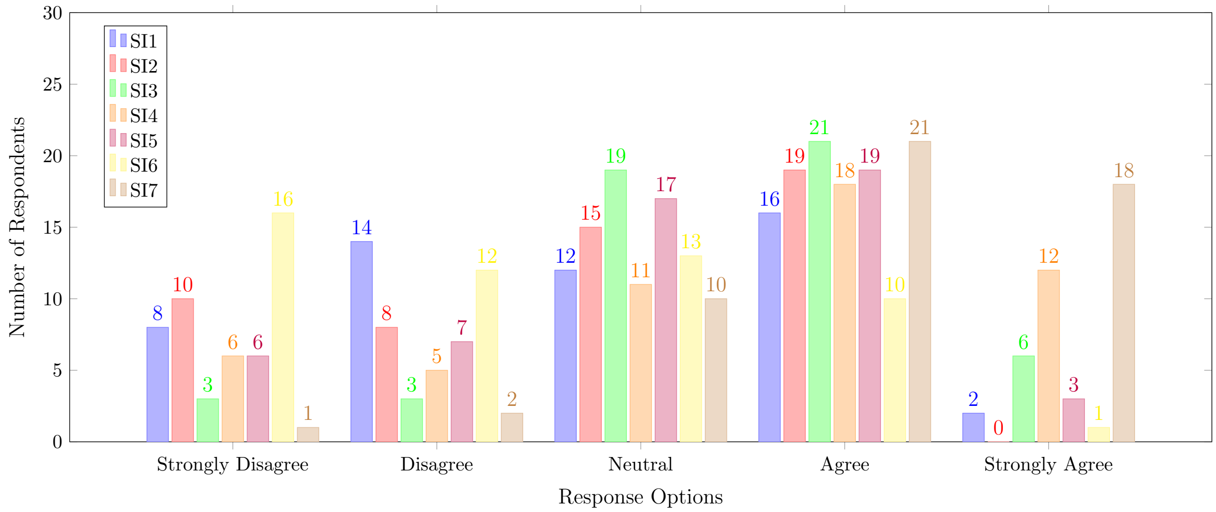

Figure 25: Social Influence results

A.5 Facilitating Conditions

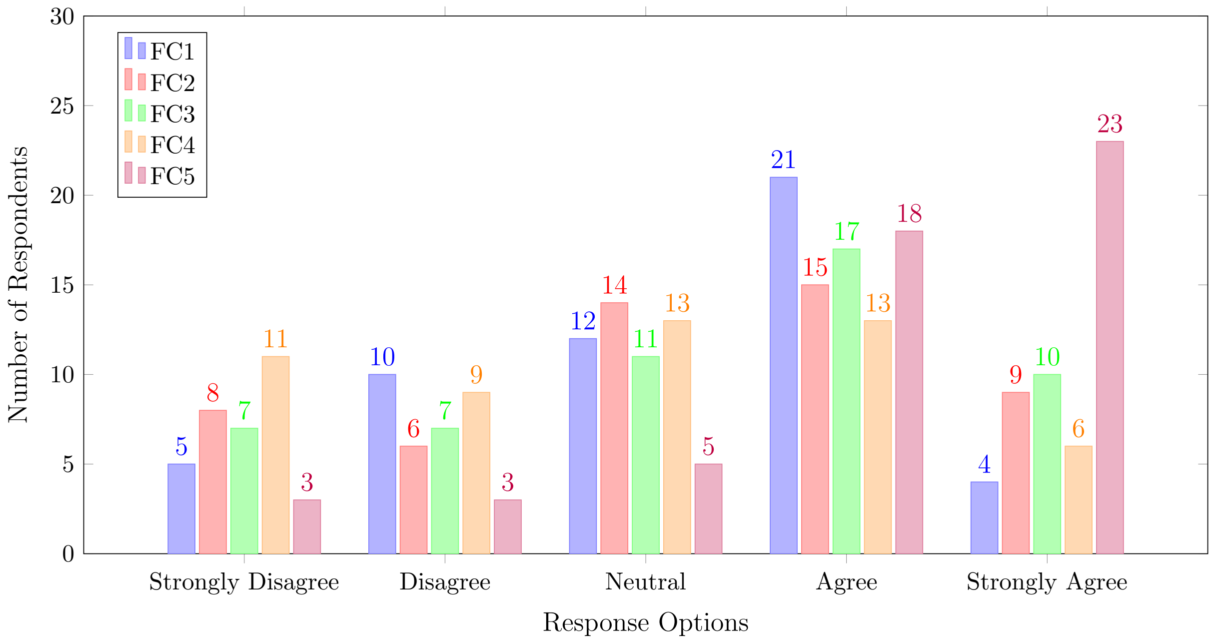

Figure 26: Facilitating Conditions results

A.6 Behavioural Intention

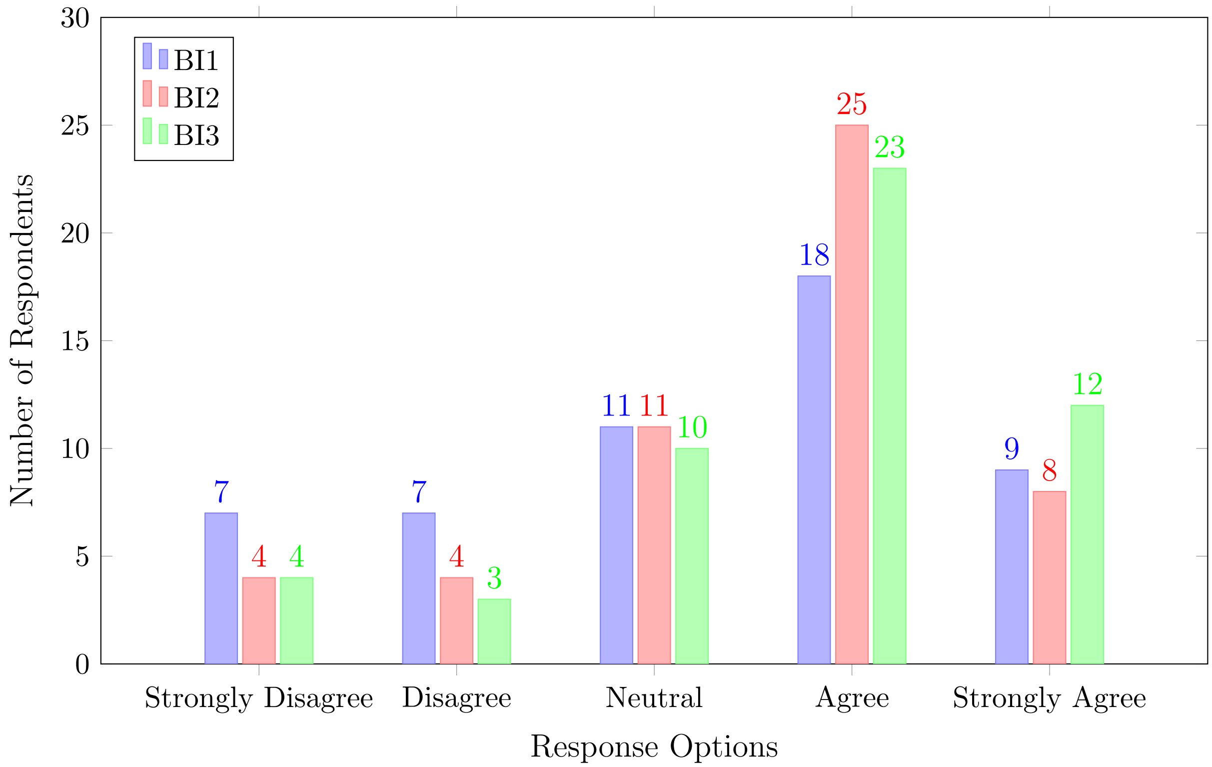

Figure 27: Behavioural Intention results