



Linnæus University

Sweden

Faculty of Technology

Department of Informatics

Perceived benefits and limitations of chatbots in higher education



Author: Alexander Lidén

Author: Karl Nilros

Supervisor: Nam Aghaee

Examiner: Sadaf Salavati & Patrik Elm

Term: VT20

Subject: Informatics

Level: Bachelor

Course code: 2IK10E



Abstract

Prior to 2012 artificial intelligence, the study of intelligent agents, followed Moore's law which states that compute is doubling every two years. Post 2012 it has been doubling every 3.4 months. However, intelligent agents are focusing on human language, and conversation is rarely developed for education. This study investigates a student's perceived benefits and limitations of chatbots in higher education, by exploring the relative advantage, complexity, and compatibility of a different chatbot functionality. By interviewing students the authors could establish four different themes that perceived to be important when using a chatbot, *Decreasing obstacles*, *Enhanced learning process*, *Hesitance towards complexity*, and *Teacher involvement*. Overall, this study suggests that it is preferable to start with little functionality and then successively improve. Because smaller implementations with basic functionality are more accepted and useful to students compared to complex AI functionality, and for future implementation, this is something that should be accounted for.



Abstrakt

Före 2012 följde Artificiell intelligens, läran om intelligenta agenter Moores lag vilket innebär att data beräkningars kraft fördubblas vartannat år. Efter 2012 har det fördubblats var 3,4 månad. Dock utvecklas ofta de intelligenta agenterna med fokus på det mänskliga språket samt hälsa och sällan för utbildnings syfte. Den här studien undersöker studenters uppfattning om fördelar och nackdelar av chatbotar i högre utbildningssyfte genom att utforska relativa fördelar, svårigheter och kompatibiliteten av olika chatbot funktionaliteter. Genom att intervjua studenter kunde författarna etablera fyra olika teman som uppfattades vara viktiga när en chatbot används, *att minska hinder, förbättra lärningsprocessen, tvivel gentemot svårigheter och lärarens medverkan*. Sammanfattningsvis pekar denna studie på att det är att föredra att börja utveckla chatbotar med lite funktionalitet och att sedan successivt öka. Detta för att mindre implementationer med grundlig funktionalitet är mer accepterad och användbar för studenter jämfört med komplex AI funktionalitet, och detta är något att ta hänsyn till i framtida implementationer.



Key words

Higher education, Conversational agent, Educational chatbot, User perspective, Artificial Intelligence

Acknowledgments

We want to first and foremost thank the five informants who were kind enough to participate even through stressful times. We know for a fact ourselves that the database course is advanced and needs your full attention. Even after the Covid-19 spread, informants agreed to execute the interviews online. Without their cooperation, we would not be able to finish this study. Also, a special thanks to Nam Aghaee who has been our supervisor and friend. She has helped us with all kinds of different aspects regarding thesis structure, data collection, etc. This has been vital for us to perform this study.

Växjö 6/9-2020

Karl Nilros & Alexander Lidén



Table of contents

1	Introduction	1
1.1	<i>Background</i>	2
1.2	<i>Problematization</i>	3
1.3	<i>Purpose</i>	4
1.3.1	Research Question(s)	4
1.4	<i>Limitation</i>	4
1.5	<i>Disposition</i>	4
2	Literature study	6
2.1	<i>Perceived Characteristics of Innovation</i>	6
2.2	<i>Artificial Intelligence and human conversation</i>	7
2.3	<i>Chatbots</i>	8
2.4	<i>Chatbot Mediated Learning</i>	9
2.5	<i>Individual support through Chatbot Mediated Learning</i>	10
2.6	<i>Spaced-out studying and cramming</i>	11
3	Methodology	12
3.1	<i>Scientific approach</i>	12
3.2	<i>Data collection</i>	12
3.2.1	Sample selection & Population	12
3.2.2	Empirical Context	13
3.2.3	Creation of mockup	13
3.2.4	Procedure	14
3.3	<i>Analysis</i>	15
3.4	<i>Validity and reliability</i>	15
3.5	<i>Ethics</i>	16
4	Empirical results	17
4.1	<i>Students individual characteristics</i>	17
4.2	<i>Scenario 1 - Q&A Functionality</i>	18
4.3	<i>Scenario 2 - Quiz Functionality</i>	19
4.4	<i>Scenario 3 - Personalized AI-driven chatbot</i>	20
4.5	<i>Spacing effect</i>	21
4.6	<i>The value of the chatbot</i>	22
5	Analysis	24
5.1	<i>Decreasing obstacles</i>	25
5.2	<i>Enhancing the learning process</i>	26
5.3	<i>Hesitance through complexity and change</i>	27
5.4	<i>Teachers involvement and connection to courses</i>	28
6	Discussion	30
6.1	<i>Perceived benefits</i>	30
6.2	<i>Perceived limitations</i>	31
6.3	<i>Method discussion</i>	33
7	Conclusion/Contribution	34
7.1	<i>Future work</i>	34



References

35

Appendices

Appendix A – Interview Guide

Appendix B – Question & Answer Functionality

Appendix C – Quiz Functionality



1 Introduction

Prior to 2012 Artificial Intelligence (AI), the study of intelligent agents, followed Moore's law which states that compute is doubling every two years. Post 2012 it has been doubling every 3.4 months (Perrault et al. 2019). These intelligent agents perceive an environment through sensors and take actions to achieve a given goal (Russell & Norvig 2009). Perrault et al. (2019) state in their report that the time to train an image classification agent through the cloud has gone from 3 hours in October 2017 to 88 seconds in July 2019, while the cost has fallen similarly. The same report also states that 58% of large companies that were surveyed are adopting AI in at least one function or business unit (Perrault et al. 2019).

Stone et al. (2016) explain in their report on the future of AI that different domains show different challenges in the coming years, with AI in education and healthcare's biggest challenge being smooth conversations and interactions with humans. It is predicted that in 2030 the use of AI in classrooms is going to be increased significantly if they can be integrated with learning in a meaningful way, with chatbots for education at home being a familiar use as well. Mostly since AI shows promise at all levels of education due to the possibilities of personalized AI agents (Stone et al. 2016). Examples of the use of AI for studying are chatbots that simulate patients in medical education or language learning (Zary 2019; Fryer et al. 2019). This use of hardware, software, or other technology to enhance performance in teaching and learning experiences is referred to as Educational Technology (ET). The learning environments that primarily uses this kind of technology to interact with peers, teacher, and material is referred to as Technology mediated learning (TML)(Alavi & Leidner 2001).

Chatbots are described as a conversational agent that primarily uses language and conversation to interact with humans, with Chatbot-mediated Learning (CML) describing a learning environment where the student uses chatbots in their learning process to increase the quality and outcome (Winkler & Söllner 2018). Providing individual support is considered the primary potential of CML since when the classrooms get bigger and individual support becomes more challenging the chatbots can provide support for both teachers and students (Söllner & Winkler 2018). For example, "teacherbots" have been used to mimic teaching assistants in forums at Georgia tech, and not even the students knew that the teacherbot was not a real person until the end of the course (Georgia tech, 2020). There are successful implementations of chatbots, like in language learning or health education. However there is a knowledge gap in understanding how and where a chatbot is preferred by the students (Winkler & Söllner 2018). Although a spaced-out chatbot approach has been suggested in language learning (Fryer et al. 2019).

To our knowledge though there are also few chatbots being used in higher education to assist the student in the study process, apart from applications in the spectrum of online courses (Brinton et al. 2015). By researching students about plausible chatbot functionality this qualitative study intends to understand what the students perceive to be benefits and limitations for a chatbot in higher education.



1.1 Background

A big contributor to this thesis is Winkler and Söllner's (2018) literature analysis reviewing, according to them, the 80 most relevant articles connected to the potential of chatbots in education. They provide a broad introduction to the subject by exploring questions connected to CML. Winkler and Söllner (2018) also show a gap in the use of them in other fields than medical education and language learning. After reviewing articles on chatbots in education, it is clear that there is a lack of qualitative studies exploring how students perceive chatbots and their benefits or limitations for studying.

Artificial intelligence and conversational agents

The field of AI is the study of intelligent agents that perceive an environment through sensors and take actions to achieve a given goal (Russell & Norvig 2009). By utilizing different learning algorithms and techniques intelligent agents try to learn how to achieve that given goal (Russell & Norvig 2009). Education's biggest challenge to utilize AI is to achieve a smooth human interaction. AI was expected to play a bigger part of universities by now, but because of a lack of financial resources and data, AI techniques in education are much less effective (Stone et al. 2016). Moreover, smooth interaction with humans has shown to be more difficult than expected. Chatbots is according to some researchers one of several types of conversational agents (Radziwill & Benton 2017). However other researchers in the field refer to a chatbot as chatterbots, artificial conversational entities, and virtual assistants such as Amazon's Alexa or Google's Home (Söllner et al. 2017). In this thesis, the definition "chatbot" is used.

Chatbots for education

In CML the effectiveness of the chatbot often depends on different factors such as student individual characteristics and how well they can see the value of the chatbot prior to using it (Winkler & Söllner 2018). Different chatbots have been tested in education to some extent with language learning being one of the few fields where chatbots have been adopted and widely used (Winkler & Söllner 2018; Fryer et al. 2017; Fryer, Nakao and Thompson 2019). Language learning chatbots can be used as language partners providing huge help because conversation practice is key to learning a language and not always something that is available. For example, Fryer et al. (2017) conducted a study using chatbots for speaking tasks with their students. Their results showed a drop in the task interest after the first tasks with chatbots, but not for speaking tasks with humans. This suggested a novelty effect, meaning that an initial interest falls off shortly after implementation, thus also performance. Although that suggested a novelty effect, a later study by Fryer, Nakao, and Thompson (2019) that built on the previously mentioned study, showed a rebound in the interest of the chatbots suggesting a spaced-out approach might provide more interest. The study also showed that the interest in chatbot activities arose from the interest in talking to human partners and learning a language. It was suggested that finding and addressing those primary interests should always be first for the technology to have a role in education (Fryer, Nakao & Thompson 2019).

Another field that AI and chatbots are becoming an application in is health and health education, both physical and psychological. Early applications were for example



Alepis and Virvou (2011) who created a model for creating characteristics in e-learning, focusing on showing emotional states that a tutoring agent could use. However that was an early application which is not state of the art today. It showed success in the use of conversational systems with high acceptance from both students and teachers that perceived the system as a very useful tool for health education (Alepis & Virvou 2011).

Benefits and limitations

The benefits of chatbots are often related to always being available to do a task that would otherwise not be possible or take a longer time (Fryer, Nakao & Thompson 2019; Nadarzynski et al. 2019; Palanica et al. 2019). However, there are not only benefits. For example, in a recent study a majority of informants showed interest in using chatbots for minor health concerns and using it instead of health phone lines. However, it also showed hesitance to whether the chatbot was accurate and trustworthy for more serious health concerns (Nadarzynski et al. 2019). Similarly, physicians have shown an interest in using chatbots for simple things in their clinics, like scheduling appointments and provide medical information. Moreover, physicians expressed concerns as well, like problems with patients self-diagnosing too often or a lack of understanding diagnosis through a chatbot (Palanica et al. 2019).

Another example is spaced-out daily reminders through chatbots to improve healthy behavior, which showed the benefits of significantly higher fruit and vegetable intake. Nevertheless, there were also limitations with a majority expressing a problem with repetitiveness while using the chatbot daily (Bickmore, Schulman & Sidner 2013). While chatbots provide benefits, these studies also show limitations through different applications and domains suggesting a variety of possibilities and challenges.

1.2 Problematization

There is limited research on chatbots in education, with most of the research focusing on health and language learning. Thus making inefficient CML experience a challenge to avoid for other fields without the proper research. In CML the students who understand the use and value of the chatbot before using the chatbot usually feel more content about chatbot learning (Winkler & Söllner 2018). Therefore it is critical to understand the students' perceived benefits and limitations of chatbots for higher education, so developers can understand their users better. Winkler and Söllner (2018) also provide an insight to what is expected by a chatbot, with the chatbots adaptability towards the student being one expectation. To achieve adaptability a chatbot would need personalized data from students, which presents a problem since you need data to get the data. Thus initial development would be uncertain together with questions on what students value in the bot. For example, if the value of the chatbot is the information it provides and not how it is provided through the chatbot, then the chatbot provides no value. Thus a regular website or other media might be preferred by the student.

Winkler and Söllner (2018) suggest that one approach for future research should focus on understanding how and where a chatbot would be preferred by the students.



Another suggestion for the future of chatbots in learning is to explore a more spaced out chatbot usage, suggested by Fryer, Nakao, and Thompson (2019). Thus to get a better understanding of the student's perceived benefits and limitations this study will focus on interviewing students about two chatbot functionalities that are currently possible to develop for chatbots internally or through third-party services. Followed by a third scenario of a plausible future functionality with a more personalized approach. It will also use the first two functionalities to demonstrate a spaced-out approach to chatbots, considering it was suggested as a possible approach. These perceived benefits and limitations would provide a bigger understanding of the users and a more stable starting point for developers who aim to develop a chatbot for students in higher education.

1.3 Purpose

The purpose of this thesis is to gain a better understanding of what the students perceive as benefits and limitations when using an educational chatbot in higher education. Based on the purpose of this study the two following research questions shown in the next subsection 1.3.1 were developed.

1.3.1 Research Questions

The questions the study aims to answer are:

What benefits do students perceive to experience from an educational chatbot in higher education?

What limitations do students perceive to experience from an educational chatbot in higher education?

1.4 Limitation

This study has no intention of exploring chatbot's characteristics such as politeness or appearance. This study examines the student's perceived benefits and limitations of an educational chatbot, to describe what the student would value in a potential implementation. This study only focuses on five students currently studying their first year of System science at Linnaeus university. Lastly, this study is only interested in the student's perspective. The teacher or developer's perspective is not accounted for in this study.

1.5 Disposition

In chapter two the theoretical basis of this study is presented. The chapter begins with a description of the theory intended to be used in the study and how the theory relates to the student's perception of chatbots. The chapter ends with an explanation of different aspects regarding chatbots and this study. Chapter 3 explains the method of choice for this study. After that, our procedure and how we performed our method of choice. Then how the result was analyzed and lastly ethical aspects and the validity



and reliability of this study and what actions we took regarding these aspects. The next chapter covers the result from the interviews presented in different sections. The first section focuses on describing the input perspective of CML through the student's prior knowledge, experience, and overall perception of chatbots. In chapter 5 the most significant result and themes are presented in relation to the theory and the perceived characteristics of innovation model. Chapter 6 is where the analyzed result is connected and compared to previous work and literature review. The chapter is divided into two subsections that are based on our two research questions. The final chapter is a short description of the major findings in this study and how they can answer our purpose and our research question. The chapter also concludes what kind of future studies that would be beneficial.



2 Literature study

The literature study that forms the theoretical basis in this study is presented below. The chapter begins with a description of the theory intended to be used in the study and how the theory is connected with the student's perception of chatbots. After that, there is an explanation of different aspects regarding chatbots and this study

2.1 Perceived Characteristics of Innovation

The perceived characteristics of innovation model are designed to measure the different perceptions that an individual may have of adopting an information technology innovation. Moore et al. (1991) explain that to be able to establish the perception of new IT certain aspects should be accounted for.

The theoretical framework measures various aspects and covers different areas that can be used to describe the perceived benefits or limitations of a chatbot. The original model consists of five main characteristics that the authors implemented for this study, relative Advantage, Compatibility, Complexity, Observability, and Trialability (Moore et al. 1991). Due to this study's goals and methodical approach, changes have been made and a more fitting model has been developed. Trialability would in this study refer to which degree the chatbot could be experimented with before any implementation. Thus the trialability aspect was removed given it was not a real implementation. The observability aspect was also removed due to the lack of implementation. This aspect would be about how users see others using the software and how they got affected by it.

Relative advantage

According to Moore et al. (1991), this aspect means to which degree the innovation brings an advantage compared to the original way of doing things. In this study, this aspect refers to which degree the new chatbot brings value compared to the previous method of study.

Complexity

For a real chatbot implementation this aspect would refer to the degree to which students perceive the new chatbot to be hard to use. This aspect includes the ease of use when dealing with the bot, such as getting started, the design, navigating in the bot, understanding the content of the bot, or how to use it.

Compatibility

The compatibility aspect according to Moore et al. (1991) is the degree to which an innovation is perceived as being consistent with the existing values, needs, and past experiences of potential adopters. In this study the compatibility aspect relates to which degree the chatbot delivers the same content as previous studying methods. The content in the chatbot will be questions and information about database development, this is content that the authors have received from the executive professor. With that said the authors have no influence on the compatibility regarding the content of the bot. Although the informants could still express their worries about a fully developed chatbot and how compatible it would be if implemented.



Connection to perceived benefits and limitations of chatbots

There is a variety of models for use to verify software or an implementation success. However, there are not so many models to use when talking about perceived thoughts from informants, what benefits and limitations they anticipate. As previously mentioned Nadarzynski et al. (2019) writes about how a lot of people would use a chatbot instead of phone lines to get medical help. Although the positive mindset, there was a lot of hesitation towards the bot's accuracy and trustworthiness especially for more serious concerns. Thus it was important to investigate the *relative advantage* and *compatibility* when investigating perceived benefits and limitations, especially since they also brought up accuracy as a concern (Nadarzynski et al. 2019). The relative advantage is the extra advantage the student gets from switching to the chatbot compared to their former method. Even if the relative advantage is important when using new software the compatibility affects the outcome the informants in Nadarzynski et al's (2019) study expressed concerns that the chatbot might not be accurate and trustworthy. Here it is clear that compatibility is an important aspect. The chatbot should help and deliver the same content and result as a phone call would, otherwise the users might feel hesitant towards using it.

For a chatbot to be accepted and used it is important that it can be simple to use and that the chatbot is not requiring too much of a hassle to get familiar with. Winkler and Söllner (2018) write in their study that it is crucial that the user beforehand has knowledge about chatbots and what the chatbot's aims and goals are. If not, the user might think that the chatbot appears too complex and would not transition towards using a bot. *Complexity* is therefore taken account for when evaluating perceived thoughts about chatbots. Too complex chatbots implemented with many different functionalities might be hard for the user to fully understand and will therefore rather use something simpler.

The three aspects are together used to reach the answers to our research questions and thereby the foundation of the purpose. The relative advantage, compatibility, and complexity all contains benefits and limitations that can be applied to the perceived experience of an educational chatbot.

2.2 Artificial Intelligence and human conversation

The environments intelligent agents acts in can be described through different dimensions of categories. An example of one dimension is deterministic environments vs stochastic environments (Russell & Norvig 2009). A conversation can be described as a stochastic environment since there are no explicit rules or behavior that could be used to determine with a 100% certainty what a person will say given what has previously been said. Conversational agents use Natural language processing (NLP) to process all the information from a conversation. NLP includes a variety of topics that surrounds the way we process and understand the human language in systems, and is considered a subfield in AI (Otter et al. 2018).

The topic of understanding language is called Natural Language Understanding (NLU) and is more complex than it seems because the human language requires context information to be fully understood (Russell & Norvig 2009). To build a



system that fully speaks and understand the human language just as good as a human is being considered an AI-complete problem (Yampolskiy 2012). AI-complete is a term that is used to describe problems in AI that is being considered extremely complicated that probably not just a single algorithm can solve (Yampolskiy 2012). AI was expected to be more part of Universities and ET by now, but according to Stone et al. (2016) this is mostly because of a lack in financial resources and data that makes the AI much less effective (Stone et al. 2016). However, since teachers communicate through human language with students it might not be that surprising that smooth interaction is hard considering the difficulty of NLU and the stochastic nature of the human conversation.

2.3 Chatbots

There are different ways to categorize chatbots. The intelligence part of the chatbot can be described through adaptation, reasoning, and autonomy. Adaptation refers to how context-aware a chatbot can be and use that context to change the interaction with the user. The reasoning can be seen as the way the chatbot decide, it might be through logical rules set by a human or reasoning learned from data. The autonomy part focuses on whether the chatbot relies on human input if it is totally autonomous or uses a mixed approach (Lebeuf et al. 2017).

Chatbots can also be described by their purpose. Generalist bots are bots like Apple's virtual assistant Siri that is working generally for simple tasks and often refers the user to further resources. Transactional bots focus on executing a task, like purchasing something. Informational bots are referring to bots that focuses on gathering information for users. Productivity bots are bots that are focusing on helping with productivity. For example, simplifying repetitive tasks so the human does not have to do it. Then there are communication chatbots that focus on helping the users communicate, like connecting the user with the right customer service agent for the user's matter (Lebeuf et al. 2017).

While deciding the purpose of the chatbot there has to be a decision if the development of the chatbot is going to be from scratch inhouse or if they are going to leverage third-party services that help with both development and distribution of chatbots. Microsoft, Facebook, and Google all provide services for the development, distribution, and customization (Lebeuf et al. 2017). If an organization does not have the skills to fully develop a chatbot by yourself or do not need a strictly customized chatbot then this might be an attractive choice. In this thesis we will utilize one of these services, QnA Maker, which is one of the services that Microsoft provides that focuses on information bots with Question and Answer (Q&A) functionality for informational bots. However, it will be used both for question and answer, as well as mocking the functionality of a Quiz.

A recent study by Nadarzynski et al. (2019) evaluated the willingness to use informational chatbots and their acceptability in health care. The study was a mixed study with semi-structured interviews with chatbot demos for some informants, and then an online survey. In the interviews, a majority of informants showed interest in using chatbots for minor health concerns and using it instead of health phone lines.



Although there was a motivation for smaller issues, a lot of the informants were hesitant to whether the chatbot was accurate and trustworthy. However, even though the chatbots were unqualified for serious health issues, just being able to receive general advice was seen as acceptable. The survey had results showing that people would use the chatbot for seeking general information and booking appointments, but feel of in some matters like sexual health. All in all the study showed interesting results with mostly accuracy and security being a concern for the informants (Nadarzynski et al. 2019). The use of chatbots in healthcare is closely related to education when it comes to AI, considering the biggest challenge in both fields is to create a smooth interaction with humans (Stone et al. 2016).

Physicians have also expressed that from their perspective there are both benefits and limitations with chatbots. In a recent study by Palanica et al. (2019) they showed that chatbots provide benefits for patient education and treatment. For example, to provide medical information and advice through chatbots. However, the study still showed concerns from the physicians with problems regarding self-diagnosing and not understanding the diagnosis. The reason for their study was similar to this study where there are many positive viewpoints for the technology, but where it still has an uncertainty of the outcomes with a lack of overall real-life implementations (Palanica et al. 2019). Similar chatbots show benefits with the technology in education, but uncertainty with few real-life implementations.

2.4 Chatbot Mediated Learning

Winkler and Söllner (2018) suggest looking at chatbots in education following the theoretical model of TML that Gupta and Bostrom (2009) described. Overall CML can be seen as a term involving describing the entire learning process with a chatbot and how the different parts of the process affect each other. The focus of CML is to allow the student to be more in control of the learning process, which is according to the constructivistic learning theory one of the main success factors for having effective learning. CML considers chatbots in education through three different perspectives. The structure (input), the learning process (process), and the learning outcome (outcome)(Winkler & Söllner 2018).

Input

The input perspective focuses on two parts. Part one of the input perspective is the individual characteristics of the learners. Individual differences in students affect the CML experience and those students who understand the use and value of the chatbot before using the chatbot usually feel more content about the chatbot learning (Söllner et al. 2017). For example individual characteristics like computer experience, self-efficiency, self-regulated, learning styles, and learning motivation are all characteristics that affect the CML process and the outcome (Winkler & Söllner 2018). Likewise did Fryer, Nakao, and Thompson (2019) discuss that the student's primary interest and motivation was something else prior to working with the chatbot. The second part of the first perspective involves the chatbot design and quality which will not be covered in this thesis.



Process

The learning process, which is a phenomenon that through learning methods, individual differences between learners, and other learning scenarios can describe the cognitive process and interactions various students experience with CML. Winkler and Söllner (2018) suggest looking at the learning process in CML, since it is crucial to understand the way technology impacts learning and not only how the individual differences impact it. The process should be synchronous with the student providing them a way to use the chatbot as they fit, with them controlling the learning process and the teacher guiding them. Their findings suggest that the learning process becomes low when the student does not fully understand the value of the chatbot and thus the embeddedness of the technology needs to be addressed (Winkler & Söllner 2018). To actively control the learning process is something that is essential to effective learning, thus the potential of the chatbots is the way it can provide individual support throughout the learning process. Making it so that the student is more in control.

Outcome

It is through the individual control of the process that the learning outcomes become improved (Winkler & Söllner 2018). In TML, which CML is based on, the learning outcome is described as being affected by procedural and structural factors. Structural factors refer mainly to the learning method and the procedural factors to the way that students act in the learning process and to fully understand the technology's impact on the outcome the factors need to be fully understood as well. Although research show that in TML the teacher still has one of the most important roles whether the technology is improving the students (Caporarello et al. 2017). In this study, this perspective will not be the focus since there will not be a real implementation where an evaluation of the outcome would be possible.

2.5 Individual support through Chatbot Mediated Learning

Individual support and human interaction through chatbots in language learning is one field where studies have been done throughout the past years. Although as of now it is still not a substantial technology in language learning (Fryer, Nakao & Thompson 2019). The use of chatbots has also shown positive effects on the willingness to communicate in English as a foreign language with daily conversations. It showed a reduction in anxiety and suggested that daily use of a chatbot increased self-confidence in the students (Ayedoun, Hayashi, & Seta 2015). Frequent spaced out usage for chatbots was also suggested by Fryer, Nakao, and Thompson (2019) with their study on chatbots for language learning. In their study, the students undertook human and human-chatbot pair conversations activities and then was immediately after asked on the interest in the two different partners. Their conclusion and suggestion were not only to provide spaced out usage but also to address the primary interest of the students. In their case, the primary interest and motivation were speaking to a human and learning a language.

Medical education is another field where chatbots have seen some applications throughout the years. Although this chatbot differs compared to language learning



since they are often used to portray patients and has not always been in chat form. An early example was Alepis and Virvou (2011) that created a model for creating characteristics in e-learning, focusing on showing emotional states that a tutoring agent could use. That was appreciated by both teachers and learners. Similar use of chatbots acting as patients for nursing education are currently being tested in Singapore (Zary 2019). The use of chatbots can not only be used to educate students but also as education and support to the patients.

2.6 Spaced-out studying and cramming

Spaced out studying or Spacing effect refers to a learning phenomenon that says that spacing out studying over time is preferred for remembering information in long-term memory. Crammed learning or cramming on the other hand refers to studying that occurs in immediate succession, for example studying the entire day before a test. The spacing effect has been found in hundreds of articles including several meta-analysis and reviews. In the articles, it is typically tested in experiments against crammed learning. It has been tested across different timescales, such as seconds, hours, days, and years. In these experiments, the spaced-out timescales have consistently shown to prove more constant remembering for longer periods than the crammed timescales (Vlach & Sandhofer 2012).

Spaced-out chatbots have been tested with positive effects in another domain than studying. For example, spaced-out conversational agents have been used as a daily intervention to change a user towards a more healthy behavior with more exercise and fruit or vegetable intake (Bickmore, Schulman & Sidner 2013). Especially the fruit and vegetable intake showed a significant change. However, several users also mentioned problems with the repetitiveness of the interventions from the conversational agent (Bickmore, Schulman & Sidner 2013). While this was not studying, it still showed how conversational agents can be used to promote spaced-out usage as Vlach and Sandhofer (2012) explain being preferable over crammed.



3 Methodology

In this chapter, the method of choice for this study will be explained. After that our procedure and how we performed our method of choice. Then how the result was analysed and lastly ethical aspects and the validity and reliability of this study and what actions we took regarding these aspects.

3.1 Scientific approach

Jacobsen (2018) explains that all case studies have in common that they are a profound study of one or a few research objects. In this thesis, we focus on a specific course that system science students have to participate in. The case is thereby all the first-year students on the system science program at Linnæus University. The absolute unit is a single student as an individual. They together form the Collective group, the first-year students at the system science program.

The research questions are about extracting the student's thoughts and requirements, requirements that the authors perhaps never even thought about. A quantitative approach would, therefore, be ineffective. To really comprehend the student's thoughts the authors thought it to be necessary to use a qualitative method with a mockup chatbot as a tool. It might be hard as an interview informant to really understand how AI could help a chatbot to be acceptable if not demonstrated. If there is no explanation with any sort of mockup the whole interpretation is left to the informant's mind and prior experiences. The study will then have a big risk to be pivoted on the informant's behalf and the reliability would be low.

3.2 Data collection

The data collection consists of four parts, the sample selection, followed by the empirical context, the design of the mockup that was later used during the interviews, and lastly how the interviews were executed. These four parts will be further explained in the sections below.

3.2.1 Sample selection & Population

This study will focus on first-year bachelor students on the system science and informatics program. More specifically, students who are taking the "Design of databases" course. A case study was performed with a class of 38 students as the case and population. The reason for the choice of the population depends on two different reasons. The program includes a lot of technological learning for students, both with coding and implementation of software in organizations. Compared to other courses this stretches throughout the whole semester which gave the authors more time to prepare and execute. "Design of databases" is also a course mixed with both coding and conceptual knowledge. This helps with creating different kinds of opportunities and questions for the mockup. Due to the program's IT content, there was hope that the students would be more open to education technology than non-IT programs. It was also more suiting for the authors, who are in the third year of the same program and knew the content of "Design of databases".



Due to the low impact of gender and other demographics regarding technology the authors used a randomized approach with no gender criteria, so that is what this thesis will focus on next. The randomized approach is what it sounds like, the sample is chosen randomly of the targeted population (Jacobsen 2018).

3.2.2 Empirical Context

Jacobsen (2018) explains that it is important that the environment for interviews reflects the purpose of the study. The university was considered a natural environment for the interviews since all the informants were students. A place where educational chatbots someday might be implemented and used by the students. Group or classrooms on school property is the most natural context the authors could think of. There is no reason to consider an artificial context due to the very few negative aspects with the natural context in this case.

Due to Covid-19 circumstances, only 2 of the 5 interviews were held in the previous context described. After the first two interviews changes were made and skype was used to execute the interviews from home. Skype was also considered a natural context, not because of its connection to the school but due to the students free choice of location, a place where they could feel comfortable.

3.2.3 Creation of mockup

The mockup was designed to help examine different chatbot functionality in higher education through previously presented theories, mostly with the modified perceived characteristics of the innovation model. CML's first two perspective, input, and process, was also considered when constructing the interview since these two greatly affected the overall learning outcome. Thus the first part of the interview focused on the input perspective and establishing individual characteristics in the students and their prior perception of chatbots. In the second section, different scenarios were demonstrated to discuss different chatbot functionality in a CML process.

Considering chatbots can have different purposes (see Chapter 2.3) the authors wanted to use separate chatbot scenarios to demonstrate how different chatbots could be used in plausible CML processes. This was seen as crucial since students who understand the use and value of the chatbot before using it feel more content towards it (Winkler & Söllner 2018). Thus for the students to express their perceived benefits and limitations the demonstration was crucial considering a demonstration would supposedly be a part of a real implementation as well. The mockup is using Microsoft Azure and QnA Maker for Scenario 1 and 2. Information from the course "Design of databases" was used, for example information that could appear on the final exam. The chatbot used in the demonstration uses predefined questions and answers to demonstrate Question and Answer (Q&A) functionality (see Appendix B) and to demonstrate a quiz functionality that would require more advanced techniques if in a real scenario (see Appendix C).

Scenario 1 demonstrates a study process using an *informational chatbot* with simple Q&A functionality (see Appendix B). Basically, the student could ask questions about the material in the course or about the course and the chatbot give standard answers.



Scenario 2 focuses on demonstrating quiz functionality with a *transactional bot* with the purpose of testing the student (see Appendix C). It also demonstrates functionality that identifies missing information and entities in answers through NLP techniques to provide hints and help the student find the missing part of the answer. The chatbot functionality was also described to, from previous sessions, identify what subject the user needed to study on and kept it on the desired difficulty level. Scenario 3 demonstrates a more sophisticated behavior of a personalized AI chatbot that is a plausible future according to research (Stone et al. 2016). Although a personalized AI behavior was perceived by the authors to be harder to demonstrate through a mockup for a variety of students with different characteristics and preferences. Thus an explanation and discussion of the behavior were used instead of the demonstration using QnA Maker. It was described to use behavior from an AI chatbot that creates profiles of users from data, combined with NLP capabilities to construct questions and explanations preferred by the users and their learning profiles.

Scenario 4 explained the spacing effect to the student and demonstrates how a *productivity chatbot* could possibly be used to further expand and help the student regulate their studying with similar functionality as scenario 2.

The questions following each scenario were constructed to focus on the relative advantage and complexity of the chatbot functionality. The reason to exclude compatibility in the questions was that the authors felt it to be difficult to ask questions about the compatibility without making the informants focus mainly on the compatibility.

3.2.4 Procedure

Before the interviews, a summarized explanation was posted on the “Design of databases” course room regarding our study, purpose, goals, and what to be expected from the interviews. After that, five informants were randomly selected and contacted by email to book an interview. Around 30 minutes long semi-structured interviews were held with each individual student. The interviews started by reading the pre-written informed consent, encompassing the ethical aspect, the confidentiality to the informant, and eventual risk. After that, the informants were asked if they know what a chatbot is. Regardless of the answer, a prewritten text about chatbots was read to the informant. This to have all the informants on the same level when it comes to the perception of chatbots.

The first questions in the interview before the scenarios focused on the student’s individual characteristics and study process from the input and process perspectives in CML. Following the scenario, the benefits and limitations were examined through different questions about the functionality (see appendix A). After that, the interview progressed into the previously mentioned mockup with the four different scenarios, and questions after each scenario (see appendix A). The questions were used to bring out the informant’s perceptions on the functionality by establishing the *relative advantage, complexity, compatibility* of the chatbot in relation to their current studying process. The mockup used a simple chat window where the informant could move forward in the conversation by using predefined questions or answers given by the authors until the entire scenario was finished. By doing this the informant got a



demonstration of the chatbots functionality and initial perception of how it could be used.

After the scenarios, the last part focused on finishing the interview with their final thoughts on using chatbots as a studying tool. Here two main questions were asked. One to establish what they perceived to benefit them the most in a chatbot and one to describe what they think should be the minimum functionality needed if a chatbot were to be implemented.

3.3 Analysis

Jacobsen (2018) explains that analyzing the collected qualitative data is often done in four steps. Documentation, explore the data, categorize the data, and lastly merging and finding relations. By following these steps from Jacobsen (2018) the authors made sure that the study extracted as much information as possible from the interviews and realizing the relationships between the informants. The authors also used the previously presented theory to help analyze the result and to be able to answer the research questions.

The analysis started by transcribing all the audio files and the comments made during the interviews with something that could be worked with. After that, the authors searched for keywords that appear often in the texts. This gave the authors some idea of what the informants thought to be important and guided the analysis. After that, the authors created codes out of what the informants said and the authors made sure that the codes were formed naturally from the informant's thoughts and not forced upon to strengthen this study. By going through the result of each informant and each scenario, information that was not in line with the study's goal was discarded.

Since the interview guide was designed from our theory (The characteristics of innovation model, see Chapter 2.1) and the analysis of the answers is categorized in line with the model's aspects. Answers from each informant and each scenario was put together to analyze the students' differences and similarities regarding the three aspects from the characteristics of innovation model. The codes in line with the same aspect of the theoretical model were studied with each other and connections were found between them. These groupings of codes were fused together and created the main themes of this study. Codes that were mentioned once or twice were discarded from the study and codes that were mentioned three or more were taken further into the analysis.

3.4 Validity and reliability

Jacobsen (2018) talks about validity and reliability for the quantitative approach and how important the two are. However Jacobsen (2018) explains that for a qualitative study different aspects should be accounted for and other concepts applies. When doing qualitative research it is more accurate to talk about internal validity, external validity, and reliability.

The internal validity is mostly related to the informants, the results are what the informants say. The selection of informants, therefore, has a crucial part when



ensuring a high internal validity. The authors decided early that the study would examine students in the system science program and that the informant's demographics were not very interesting. The authors concluded that the age difference between the students would not be important since there were no seniors in the class, it was also decided that ethnicity and gender would not matter when using a chatbot. The selection was thereby only first-year students studying in Växjö in the system science program.

To reassure high external validity the authors created an interview guide that was followed thoroughly in every interview. The interview guide was created from the presented theory and previous research. This to be able to create questions that were more generalizable and to be able to connect the result findings to our theory and thereby say something about it. The interview guide was always presented and followed to make the interviews similar.

It was tried to assure high reliability by writing and explaining the questions in a way that the result could not be connected to one of the individual informants and thereby not harm them in any way. The informants names were never mentioned, to make sure no informant could be identified through the study. The confidentiality was also explained to each informant before the interviews were held.

3.5 Ethics

Jacobsen (2018) says that there are three basic requirements that are related to the ethical aspects between the researcher and the informant. Informed consent, the right to private life and correct presentation of data.

All three of Jacobsen's requirements have been considered when conducting this research. Informed consent was explained to possible informants at first contact with email and at the beginning of each interview. The authors explained the purpose and goals with the study, free will was very important and it was clearly explained from the start. The authors do feel that competence is high enough in a university class and we made sure that they understood everything about the study before we started. Since our research had few informants it is important to make sure that no informant can be identified from the data. This is something we have thought about during the whole process and we have ensured high privacy. To present the data correctly we have been careful when citing our informants to not take anything out of context. The authors carefully created categories when processing the data, to not claim anything that the informant did not intend. As well as replaced the names of the informants by referring to them as informant 1-5 to increase the confidentiality of the study.



4 Empirical results

Below is the result from the interviews presented in different sections. The first section focuses on describing the input perspective of CML through the student's prior knowledge, experience, and overall perception of chatbots.

The result focus on describing the perceived usefulness of a chatbot described throughout the interview. Section 1 "students individual characteristics" focuses on describing the initial thoughts towards chatbots and the current way they study. The main goal was to establish a description of the informants that further down the study be used in the analysis of the results.

4.1 Students individual characteristics

The interviews began with asking the students about their studying to establish individual differences, similarities, and their attitude towards chatbots. As well as establish if they used a more spaced out or a crammed studying style. Informant 5 did not have any prior knowledge at all of chatbots, and thus informant 5 had a neutral attitude towards chatbots. The other 4 informants had a positive perception of chatbots and their benefits. One important note is that informant 3 was different than the other informants since prior to studying informant 3 had worked at an insurance company and been involved in a successful chatbot project before, thus had a big understanding prior to the interview.

"I've worked in the insurance business and there you're going towards using chatbots in customer service as you said. I was involved during a pilot to test our chatbot, it worked okay it answered a lot of basic questions and such. A lot of the questions we got were basic so a big chunk of the questions could be answered by a chatbot." - Informant 3

After an introduction to different functionality of chatbots, the informants were asked what they perceived would benefit them the most with a chatbot. This is exemplified with the quotes below:

"Something where you could more quickly ask about SQL commandos like 'How do you count an average of something?' would be nice to have." - Informant 2

"[...] sometimes I can get stuck on stuff and it's annoying to ask for the teacher and wait because that wastes time, especially if it's something that is pretty basic that you could just ask a chatbot and get a quicker answer." - Informant 5

To save time and always have answers available was a value that early on was expressed from the 4 informants that had prior knowledge of chatbots. These 4 informants showed to have a much bigger understanding of the value of a bot, compared informant 5 with no prior experiences.

Both crammed studying and spaced-out studying were identified among the informants. They all focused around creating their own notes and material from the books or given by the teacher throughout the course. Three informants also mentioned



in this section that they did not like using the books, except for scamming through the book or to look up specific answers. With the other two not expressing anything about reading the books either, but not explicitly saying they did not do it. One of the informants also pointed out that he did not use any specific time schedule:

“Before I used books but that never worked so now I just use the teacher’s slides from the lectures, my notes, and old exams since I feel that is the most efficient way. [...] I try to split it up as good as I can but it often differs depending on how much study I have to do and how my schedule and life looks in the coming week” - Informant 5

Spaced out studying

Informant 3 began with his studying early in the course to be ready for an exam a week prior to the exam. Informant 4 expressed similar style and liked to study with others to have more of a discussion:

“I study throughout the course not to get a big workload at the end of the course when the exams are. In the end is often me and my buddies that go through powerpoints, concepts and stuff together to get a deeper understanding[...]" - Informant 4

Informant 4’s discussions were more focused on a deeper understanding of concepts and not on learning facts since that was done better alone early in the study process.

Crammed studying

Informant’s 1 and 2 expressed they had a more crammed studying technique that focused on studying intensive the week prior to exams. Informant 2 also mentioned not using the books much, but just focusing on their own material.

“I do my own summary of all material and after that I study practice on the older exams from previous years. [...] I mostly study last week before an exam, intensive.” - Informant 2

Both informants 1 and 2 mention they are starting 1 - 2 weeks prior to the exams with longer sessions.

4.2 Scenario 1 - Q&A Functionality

Four out of five informants explicitly said that the Q&A functionality would provide value for their studying by saving time. The informants state that if such chatbot were implemented they would try the tool. While they agree the chatbot brings value, the motivation from the informants differ to some extent.

“From my perspective, I think it would be good because it is interactive where I have to think and interact with something instead of just reading [...]” - Informant 3

Informant 3 express that the value is the interactive part of a chatbot and when he asks a question he learns better than just reading or taking notes. Informants 1 and 2



answers similar to each other regarding the Q&A chatbot and both express that they would like to use a chatbot to ask the teacher something even if the teacher is not available. They explain that sometimes you have to wait a substantial amount of time for the teacher to respond and that a chatbot could help to erase this time delay.

“It would be nice to answer quickly because the teacher isn’t always available. It happens pretty often that the teacher can’t answer in time.

“- Informant 2

Informant 5 also sees the possibility of saving time but does not feel the same need to ask the professor questions regarding the course.

“I think the advantage is big since it saves me time to find the information for old exams in books and stuff. I’d say 50% of the study time in the beginning for me is just searching for answers in books that [...] Also sometimes you turn to google for this but you can often be unsure if it’s the answer the teacher wants and thus it’s not so easy. “

- Informant 5

The 4 informants that were positive towards the Q&A chatbot had all one thing in common. They all explained that this type of chatbot would not replace their current technique, but as a complement to their current method. They would use the chatbot when they got stuck or wonder about something special. Some would use the chatbot instead of the teacher and some would use the chatbot instead of google and course literature. Informant 2 explained that it depends on how accurate the answer is as well. The ability to ask follow up questions if the answer is still unclear was expressed to be valuable. Informant 4 is the only informant that does not think that the chatbot gives an advantage and that he would not change his current learning strategy if this kind of chatbot were implemented.

“A big advantage is that it’s quick but it would be too narrow with just a simple answer[...].” - Informant 4

Informant 4 agrees that there would be a lot of time saved but that simple answers would not be explainable enough. Informant 4 rather looks it up in his book or google it even if it takes longer time to be sure that he grasps the whole concept. Though informant 4 states that it needs to have more features and that it is a good start for a chatbot that could be developed further.

4.3 Scenario 2 - Quiz Functionality

The quiz functionality was perceived to be a useful functionality for the students. There was not a single negative attitude towards this kind of functionality, except for some speculations towards low engagement teachers in the quiz and how that might affect the student negatively. For example if the quiz is not equivalent enough for the student to pass the exams. The one informant that was hesitant towards using a chatbot only for Q&A was more susceptible towards the ability to quiz yourself with a bot:

“This feels easy and nice, much better than just questions and answers, and good with hints and interaction to find wrongs. I do not think teachers would use the time to make it good enough though. For me it’s



crucial that they are involved and constantly change it as needed.”

- Informant 4

Throughout this scenario there were a lot of comparisons towards other existing quizzes, like older exams. One also mentioned the quiz that people use while studying to take your driving license in Sweden, and how that was great and thus would try this. The teacher’s involvement in these quizzes was a big factor for the perceived advantage by a majority of the informants. Informant 5 for example expressed that it would be a dealbreaker to be certain that the quiz is strictly from the teacher or the course and not something from the internet.

“I think I’d use it if I know it’s material from the course and the teachers.” - Informant 5

Overall the quiz chatbot was not going to completely take over any of the current studying techniques but it was expressed to be a valuable tool to work with together and to test your ability. The interactive functionality to provide hints to what a was missed in a question was also expressed as an appreciated function:

“ This functionality with quiz and hints is basically the same purpose as using older exams, but better, which works well to study with “
- Informant 5

“I think it is smarter and to get hints which feels much better than my current way of studying.” - Informant 1

It was also expressed that you would get more studying and training on this compared to just looking up an answer with the Q&A bot:

“This feels like you could get more training and study compared to the last bot, and it feels good that you can be quizzed in different ways than my current way which is basically just studying on previous exams in a row. “
- Informant 2

4.4 Scenario 3 - Personalized AI-driven chatbot

When the personalized AI chatbot that adapts to users learning profiles was explained and illustrated the informants answered quite differently compared to each other. Three informants thought the AI chatbot to be helpful with personalized answers or questions. They explain that they think they would use a chatbot like this if implemented:

“Yes, absolutely the more adaptability to a person would be good. The more general it is the less this will work I think” - Informant 4

Both informant 3 and 2 express similar thoughts:

“It would feel easier and more quickly to use. I think it would give value if you could get different examples and explanations if you want do not understand the current one. It also feels a little more interesting to use it, a little bit more fun if you could have your own one” - Informant 2



Informants 2 and 3 both express that everybody is different and especially learn differently and therefore they could see it being useful with a chatbot that can tailor answers and explanations depending on the student's profiles. Both informants 1 and 5 were skeptical towards a personalized AI bot.

“I often learn in different ways so creating a profile of how I learn is probably not a good approach. This is maybe good for others, but I want examples sometimes and sometimes explanations.” - Informant 1

Informants 1 and 5 agree with the previous informants, that people are different and learn differently. Informant 1 expressed it not to be beneficial because of a variety in informant 1's learning. Therefore informant 1 expressed hesitance towards this kind of functionality but expressed that it might be more beneficial for others. Informant 1 was not alone in these thoughts:

“I don't know really if it works it would be nice to always get it in a way that is suited for me, but I think it would be hard to make it work and don't really see how. I learn in different ways and often change my mind about stuff.” - Informant 5

Here, informant 5 expresses similar thoughts and expressed hesitance towards the complexity of the chatbot and problems with visualizing how it would work.

4.5 Spacing effect

In the scenario involving spaced-out functionality there was no chatbot demo, only an explanation of the spacing effect and how a spacing effect feature could be achieved in a chatbot. This was received with a lot of different results. Informant 3, that already used a strict spaced out studying process, had no need for this feature because of informant 3's own ability to achieve this. However, informant 3 did express interest in trying the functionality in case he could incorporate it easy and effective in his current process together with the other features expressed in previous scenarios.

Two informants expressed they would like this feature. Informant 4, expressed earlier in the interview that he preferred googling over a Q&A functionality, expressed in this section that this feature would help him significantly. He explained that sometimes it is hard to sit down and study early in the course for the exam. Even if informant 4 often do start early, it is easier said than done. The functionality would benefit informant 4 both through scheduling and quiz:

“I think I would use it and it would be nice to be able to customize schedule and stuff for yourself. And also, it can be really good to help me actually do it since sometimes you say you're going to do it but never do it.” - Informant 4

Similar, informant 5 talked about how it would feel appealing to have something reminding you to pick up your phone and quickly take a quiz for 10 minutes. Informant 5 also expressed that it would not feel like studying with smaller sessions. Compared to how longer studying sessions would. Thus the functionality would help



informant 5 schedule smaller sessions to avoid a negative feeling of having to study. Especially when studying is not something exciting.

“I don’t really know if I’d use it, but I think that I’d give it a shot at least. Although when I think about it this feels more inviting to just go in once a day in your phone your something than logging in to your pc and studying. I think it would feel less like studying while actually being studying so that is nice. Especially if it is something boring to study.”

Informant 5

However, it was perceived to be a good idea in theory by informants 1 and 2. But not any overall interest in changing from their current crammed studying style.

“It’s a good thought but I don’t know if I would use it. It might be just annoying. I think this would be great for some people and not so useful to others, due to our difference in style of learning and personalities, etc.”

Informant 1

“I think it sounds good in theory, but I have tried to use alarms and stuff to space out my studying and It hasn’t worked. I think I would try it out but if I’m being honest, I don’t believe it would help me, but maybe someone else since it sounds like a good idea.”

Informant 2

While not having a huge interest in the functionality, informant 2 did explain previous failed attempts to achieve more spaced out studying with alarms on his phone to start studying earlier.

4.6 The value of the chatbot

In the interviews the authors also wanted to distinguish what the value of this software actually was. If the value was the content in the chatbot and the functionality of the chatbot. Four of the informants said that when interacting with a chatbot by asking something you learn easier. Informant 1 expresses that the information provides its own value apart from the actual bot:

“The chatbot gives value, I think it’s easier to learn while asking and getting an answer instead of just reading. It works better for me. And I can also find the information easily when I ask the bot. I get the information that I really need and filter out unnecessary information.”

Informant 1

“I think a big value is that it’s a bot. Partly because it is easier when it’s interactive. For example, If I read when I’m tired I can’t learn at all. But I have easier studying or doing something more interactive overall when I’m not as alert like after training for example.”

Informant 3

Informant 4 express interest in other solutions, unless the chatbot has the quiz feature. Informant 4 expresses the problem of having to know exactly what to ask the chatbot to be able to use it. Informant 4 rather finds the information on the internet through



other sources. The quiz functionality on the other hand brings him a different value, by being able to test himself.

“I’d rather have the information on a website or something else compared to only Q&A.[...] I think the ability to quiz yourself is really good. I think the information and the chatbot need to work in sort of a harmony to work and be useful to me.” - Informant 4

The biggest benefit for the five informants were not the same, although the answers from four of them were somewhat similar. Informant 1 and 2 said that the most benefits would be the availability to ask questions without having to wait and always get relevant information. Informant 4 expresses the quiz and hints to be the most beneficial and the availability of the bot to be important as well. Informant 5 claims the biggest benefit being a mix between the quiz and Q&A. Informant 3 answer stands out from the others:

“I actually think the most important benefit is the data for the university and the teacher to understand us students and what’s missing. For example, in one of our exams basically the entire class had one question wrong and the teacher can’t change that before the exam if they aren’t aware of the problem prior to the exam. But if they had the data from a chatbot it would probably change and wouldn’t be too late...” - Informant 3

While informant 3’s perceived biggest benefit differs from the others, informant 3 still showed similar perceived benefits of availability and the ability of testing throughout previous sections.



5 Analysis

In this chapter the most significant result and themes are presented in relation to the theory and the perceived characteristics of innovation model.

The first perspective in CML focuses on the students individual characteristics and how that impact the CML experience (Winkler & Söllner 2018). Which in this study was shown mainly through scenario 4, the spaced-out functionality, where it was a clear connection between the answers and prior study techniques. However other than that there was no real theme or connection between different informants and their prior experiences throughout the scenarios. On the other hand there were more similarities in all the informants, especially in how they used the material. For example using the teacher's material for notes, discussing with friends for more in-depth, and not using books for fully reading. But by merely scamming through the books. Given that books are more often than not recommended to be read by teachers they might be more open to trying other things like chatbots if they do not enjoy reading in the books.

Winkler and Söllner (2018) point out that the individual support through chatbots should aim to give more control, and if failed to do this results in a bad CML experience. Thus given the opportunity to not use the books but instead something more interactive could be in line with more self-control to the student. On the other hand if a chatbot is recommended similar to how books are recommended to be read then the students have shown from this that they decide how to use the material and tools given by the teacher. Thus the students could just as well not use a chatbot if they do not see the value in it, similar to how they treat books. This should be considered given that the following analysis only shows the student's perception of chatbots for studying.



Four different themes arose from the analysis that each were connected to the perceived characteristics of innovation model. To get a better overview of this and how the themes were created the analysis chart was shaped (see Table 1)

Table 1 - Analysis chart

Perceived Characteristics of Innovation	Theme	Short Description
Relative advantage	Decreasing obstacles	<ul style="list-style-type: none">● Less waiting time for answers● Quickly find and get simple facts● Always reliable answers
Relative advantage	Enhanced learning process	<ul style="list-style-type: none">● Better knowledge testing● Flexibility in answers● Support for study structure
Complexity	Hesitance through complexity and change in the study process	<ul style="list-style-type: none">● Hesitance towards complex changes in the study process● Too much complexity makes it hard to visualize● Too little complexity can make it hard to value
Compatibility	Teachers involvement and Connection to courses	<ul style="list-style-type: none">● Engaged teachers● Compatible material● Communication on chatbots relation to course

5.1 Decreasing obstacles

The interviews show that simpler functionality like Q&A can provide a good start for a chatbot in education. Mainly because it was perceived to be easiest for the student to implement and use in their study process, without much altering in the way they learn. The possibilities of changing regular actions that currently waste time and become obstacles in the study process were mentioned as to why this could interest the students. These obstacles were related to actions that did not explicitly help the student learn. Obstacles like waiting for an answer from the teacher, looking through an entire book when looking for simple facts, or having to search on google.

This might not be a student's primary interest in studying, however it still shows interest in achieving a smoother studying experience and thereby an advantage compared to their previous technique. Thus the primary interest for the Q&A chatbot appeared not to be the chatbot for actual learning, but more to use as a tool on the side



during studying. All this points to the relative advantage that Moore et al. (1991) talk about in their theoretical model and it shows that a chatbot does not have to be in the center to bring value, it can have different roles in the studying process. Even informant 5, who was hesitant towards using Q&A, could still see the value of the availability and quick answers. Overall the ability to save time seemed to be a relative advantage throughout the study, even if not always the exact words “Saving time” was explicitly used every time. For example one informant mention later in the interview that he did not have to filter through unnecessary information to get all the answers, which would save time.

5.2 Enhancing the learning process

Other advantages was shown later in the study, besides just being a tool on the side for studying. The perceived benefits for a quiz functionality, personalized AI, and spaced out usage was seen to be more influential on the study process. Mainly in the way that it was more involved in the actual learning compared to the Q&A. The quiz for example was expressed as a tool that could help them be more self-regulated and self-efficacy through the way they could test and quickly assess their current knowledge of the material in the course. A student in more control throughout the learning process is something that Winkler and Söllner (2018) expressed to be a big factor that influenced the overall CML quality, thus this would be a good addition to the studying. It was interpreted as this since the quiz functionality was sufficient enough for the informants to express that it was an advantage relative to their current techniques when it comes to testing your ability. Without basically any negative thoughts on the scenario, apart from some concerns about the compatibility of the material.

More specifically was quiz functionality compared to doing old exams which was a regular technique used by a majority of informants. This comparison through older exams was a clear example of a relative advantage with the chatbot since similar comparisons often are used while describing the degree relative advantage (Moore et al. 1991). The interactive hint functionality in a chatbot with a quiz was also seen as a relative advantage to the prior technique of older exams. This might be a unique advantage for software that utilizes NLP over other software trying to achieve similar quiz functionality. However, to what degree the quiz functionality would be an advantage is out of the scope of this study. Yet the result shows an interest in enhancing their learning process with a better method for testing your ability. Given that the compatibility of the chatbot is equivalent to the course.

The way to enhance the learning process with the personalized AI described in scenario 3 was however perceived differently than quiz functionality. A personalized AI chatbot that adapts to the students according to their profile is not a regular feature that could be easily referenced to prior experience in the same way as Q&A and Quiz. Nevertheless informants could still refer it to prior experience where the functionality could fit. For example, one expressed interest in the ability to have different answers if the current explanation is doing a good enough job. It was described as interesting since it might actually help in the learning apart from just given you a standard answer



or asking you the same question all the time like scenarios 1 and 2. However, the answers were not as clear as scenarios 1 and 2 in the sense that those scenarios involved answers with direct comparisons to the current studying technique. But the advantage was still referenced to enhanced functionality that would eventually save them time while at the same time learning.

Another way to enhance the students learning process would be to help them space out their studying. This could be used for a more regulated spaced out study process through a productivity chatbot with functionality that provides reminders to repeatedly have the student test their knowledge through similar quiz functionality as described in scenario 2. Often through smaller amounts of time. This functionality showed immediate interest and relative advantage by only informants 4 and 5 to help self-regulate their study through time schedules. Informant 3 explains he does not need it since he is able to do it alone. Thus Informant 3 had a neutral attitude towards the functionality, but he would still try the functionality if it was compatible with his current technique. It was also interesting because of the fact that it might not feel as actually studying compared to the current technique that would require the student to sit down at a pc, look through their notes or open a book. Still, it was also expressed the opposite effect compared to enhancing the learning process, by informants 1 and 2 that were currently using a crammed study technique.

5.3 Hesitance through complexity and change

The complexity of changing the entire study process from crammed to spaced out was not inviting for the two informants using crammed. These two informants were currently using a more crammed time schedule in their studying, and thus the functionality would supposedly provide them the most assistance. However the complexity of changing and regulating the study process was not perceived as helping them enhance their study process. Informant 2 said that he previously tried similar reminders functionality with alarms to start studying, with no success. Nevertheless informant 2 at least showed that he has had an interest in this spaced out studying before. Thus would supposedly have knowledge of the benefits of spaced out studying and open to trying it. Informant 1 showed a similar problem with this functionality, he expressed that it would be annoying and he would not be interested in using the functionality. The difference in answers was related to the time schedules they used in their studying. The complexity of change for the students using crammed studying techniques was higher, and thus they also expressed hesitance towards implementing this functionality in their studying. This functionality would change, regulate, and push more structure in day to day schedule, which might be a higher boundary for the student to cross.

The results from other scenarios also suggest that when the complexity of implementing the chatbot in the studying becomes harder to visualize, the value of the chatbot also becomes harder to visualize. Likewise there was similar hesitance involving the AI functionality, which suggests hesitance could also arise even as the chatbot becomes more adaptable and more capable of being a personal teacher. However this hesitance was expressed to be more on the complexity of the chatbot



and not the implementation or the way it changed the time schedule. The hesitance was more about not believing in the chatbots presented ability to adapt or the ability to help with more advanced explanations as a teacher can.

When hesitance arose in this scenario it was similar to what was found in Nadarzynski et al's (2019) study on acceptance. In that study, the hesitance was towards using chatbots for more serious health concerns. Similar to hesitance that arises in this study using a more adaptable chatbot functionality, since it might not be able to satisfy the needs for adaptable assistance. Nadarzynski et al's (2019) also showed that a chatbot for more simple health concerns showed much higher acceptance, similar to this study where students described simple functionality as more appealing. This hesitance could arise due to the challenge to compare or reference the functionality to any prior experience. Thus it becomes harder to visualize how this adaptable bot would be used in a real scenario, even if it sounds like an appealing idea. Consequently, this assistance through personalization with more interaction and adaptability might provide another challenge if promised to the student due to a potential comparison of the chatbot to a real teacher. The student's trust can thus perceive to be lower towards the chatbot if promised advanced capabilities similar to how a real teacher.

The study also found smaller hesitance towards the simple task of Q&A. For example the Q&A found hesitance mostly by the one informant that did not have prior experience with chatbots and thus could not visualize the value as the previous could. Here the complexity was more in line with that it was not complex enough, and too simple for the informant. However previous studies in TML points out that the teacher still has the biggest role if the technology would be improving the student or not (Caporarello et al. 2017). Thus to reduce the hesitance the teacher needs to present the chatbot according to the primary interest of the students so they can understand the value of the bot.

5.4 Teachers involvement and connection to courses

The compatibility of the teachers and the information in the chatbot was brought up as a concern. For example, there were concerns that the teachers might not be involved enough to make the quiz equivalent to exams. If so, the chatbot would do more harm than good for the students studying for exams. Which were similar concerns as brought up by informant 4 for the Q&A functionality, where the simple answer or to general answer might not be compatible for exams. This was recurrent throughout the interviews where teachers and compatibility are described as a factor to how the students perceive the chatbot. Thus, the teacher's involvement was confirmed as having a major effect on the advantage of the chatbot. It was even brought up by two informants as a direct requirement that they understand the teachers and the course materials connection to the chatbot.

Moreover, the teacher's involvement and compatibility of the material were not expressed to be only for the students. Informant 3, that had prior knowledge of chatbots from work experience, expressed that the most benefit would be for the teacher. If the teachers were highly involved, then the data from the chatbot would be beneficial for the teachers by constantly updating the teachers on the knowledge gaps in the student. An example brought up was when at one of their exams almost the



entire class failed one specific question, and at the exam it was too late to fix that. However, the lack of knowledge might be caught earlier if the teacher would have had a data from a chatbot. Informant 3 had prior experience from a chatbot project and thus might be more understanding of the value of a chatbot from the teacher's point of view. However, it still suggests how the teacher could use the chatbot as a bridge to the student's knowledge gaps. While the students could see the value in the chatbot through how it affects their studying, this also suggests how the student could see the value through the chatbot's ability to help the teacher.

Another concern brought up was that the students might not always be sure that "Google answers" or other links would satisfy the teacher on exams. Thus, a chatbot would provide more reliability when it comes to getting answers for questions the student would have to google otherwise. Overall, the compatibility of the chatbot is very important considering that the relative advantages described in prior sections are dependent on the compatibility of the material used in the chatbot.



6 Discussion

In this chapter the analysed result will be connected and compared to previous work and literature review. The chapter is divided into two subsections that are based on our two research questions, benefits and limitations.

The contribution shows that Q&A and quiz functionality shows a relative advantage for higher education, by comparing to current studying techniques. It also suggests that complex functionality can provide problems when it is difficult to visualize how and where it would fit in the current study process. However, the result also suggests that for the chatbot to be relevant at all there has to be a plan to address how to make the chatbot compatible the course and clearly presented as such to the students. This chapter focuses on describing these relationships between the relative advantage, complexity, and compatibility. In relation to the results of this study and previous research. It is presented through two sections that discuss benefits and limitations, followed by a conclusion.

6.1 Perceived benefits

The authors did not expect the benefits of the simpler functionality of Q&A and Quiz to be more than the Personalized bot. However if a chatbot with more personalized AI would be developed it would most certainly be capable of doing similar functionality as the Q&A and Quiz. Nevertheless, separating the functionalities shows that simpler functionality is perceived to be more beneficial to the students at the time. The benefits of the Q&A were mostly regarding saving time, but also for more reliable answers than Google. The benefits of Q&A functionality have not explicitly been shown in another study, however, the description of benefits that would result in saving time is a recurring theme (Nadarzynski et al. 2019; Palanica et al. 2019; Winkler & Söllner 2018). For example, Nadarzynski et al.'s (2019) study on chatbots in health care shows similar results where the chatbots show to be more of an interest for simple health problems. In their study, the chatbots showed a relative advantage over phone lines for simple health concerns.

The quiz functionality was also perceived to benefit the students, although addressing an interest from the students other than only saving time. Fryer et al. (2019) stated, from their studies of chatbots in language learning, that having a conversation in the studied language was the primary interest of their students. They suggest that finding this primary interest is important to easier address it with a chatbot. In this study there seemed to be an interest by the students to find a way to test their knowledge better than the current technique, using older exams. Thus results show that the ability to relate the simpler task to prior experiences or techniques seems to be a way to visualize the value, which is important in CML (Winkler & Söllner 2018). Like in this study where the relative advantage for the quiz functionality was directly connected to the use of older exams.

Testing yourself might not be the primary interest for chatbots in higher education as Fryer et al. (2019) points to be important. However, having a conversation is a direct way to test the language ability and thus the primary interest in their studies seems to



be testing knowledge. Moreover with the Q&A functionality the option of googling and searching for the answer still remains as a somewhat reliable source. However, testing the ability of the student is more challenging than just googling. Thus the quiz functionality seems to address interest in students by helping to enhance their study process with better knowledge testing.

In a lot of cases the primary interest of the students might be to actually learn. However, as also pointed out by one of the students, everything you study for is not always fun and interesting. For those cases the primary interest probably switches more towards just passing the course. With that being said the interest of passing a course or exam still resides in both cases and could be seen as something that would persist as interest through all courses, regardless of interest in learning. Thus the benefit of being able to test yourself would make it easier to both pass the course and to learn, thus addressing the interest in both cases. While similar quiz functionality can be done by another software that does not involve chatbots or human conversation, the interactive hints of a chatbot was perceived to be a relative advantage for quiz and chatbot as well. Moreover, the value of the hints might be harder to visualize considering there is no direct comparison to getting hints throughout the current technique of using older exams. At least not if the student is studying alone.

The spaced-out functionality was perceived by the authors like an attractive idea considering the effect of spaced out studying combined with the previous studies that pointed to spaced-out functionality as a good approach to chatbots (Vlach & Sandhofer 2012; Bickmore, Schulman & Sidner 2013; Fryer, Nakao & Thompson, 2019). This study only showed perceived benefits, compared to actual results by Bickmore, Schulman and Sidner (2013) with benefits for more healthy behavior. However in this study, the benefits were only perceived by the students already using spaced-out studying, thus it might only be used by that group. All the while the crammed studying group, that would supposedly get the most use of the functionality can not see how the functionality would benefit them. Nevertheless, this approach still seems promising if the developers can understand how to make the crammed students understand the value. Moreover, the student needs to be in control of the study process and not feel forced or pushed by regulations in the schedule by the chatbot.

6.2 Perceived limitations

Stone et al. (2016) pointed out that the reason AI is not used much in education is because of a lack of data and resources, especially to achieve the personalization that shows much promise. However during the Q&A scenario the students express interest that does not require personalized agents and systems. These were interests that could be satisfied by simpler informational chatbots with Q&A that use little to none AI. By decreasing obstacles, the chatbot could still be able to provide individual support to students, like Winkler and Söllner (2018) points out being a big potential for CML. However, personalized AI functionality was perceived to show benefits to some informants, but the answers describing those benefits were vaguer than the other scenarios. For example, there was no reference to how the personalized AI



functionality would be an advantage compared to the current studying technique, even if the functionality sounded appealing. Moreover, the closest comparison might be a teacher, thus, a successful personalized chatbot would require smooth interactions with humans. Which Stone et al.'s (2016) point out to be a big challenge for the field of AI.

The personalized AI also lacked a clear primary interest in the way Fryer et al.'s (2019) suggests addressing. The primary interest might still be the saving time with the chatbot, thus, it would be achievable by a simpler functionality. In other words, while the personalized AI was described by some informants as being good, there was no real connection to prior studying technique. Given that Winkler and Söllner (2018) point out that CML is highly affected by the student understanding the value there does not seem to be a relative advantage enough for the student to perceive the value. At least not in this study where the answers were more limited and vaguer.

Similar to Nadarzynski et al.'s (2019) study, the hesitance arises when the complexity of the task became higher and the informants' trust lowered. Likewise did Palanica et al. (2019) show, in their study on physicians' perception of chatbots, that there was uncertainty on how a chatbot would be used successfully for more complex tasks. Overall the hesitance and limitations that arose towards the more personalized AI were related to a difficulty in visualizing the value. This scenario was after all a more future functionality than the others, thus, the answers were not expected to be as explanatory compared to the other scenarios.

The spaced-out usage Fryer et al. (2019) suggested was perceived as a benefit by the students who already used a more spaced out approach. However, this functionality was contradicting to the process perspective of CML which states that the students needs to control the study process (Winkler & Söllner 2018). Thus, students currently using a crammed technique would be forced to change their schedules, hence not controlling the process. This also showed to be correct in this study when the spaced-out functionality was neglected by the two students that used a crammed study technique. Thus, the limitations of the spaced-out functionality could be a direct consequence of not being able to control the study process.

Both those informants that used crammed studying showed an interest in the Q&A and quiz functionality from scenario 1 and scenario 2, which was described to be used in the spaced-out functionality. Thus the students can get an experience of self-efficiency with the Q&A and quiz functionality before introducing spaced-out functionality that regulates. Which might be a better approach to achieve a more spaced-out study technique. However, each scenario was dependent on how well the chatbot was compatible with the teacher and the course. The compatibility to the course might be an obvious requirement. However, the teacher must consider how the compatibility is going to be presented to the student for them to understand the chatbots' value.



6.3 Method discussion

The interview was transcribed and the thematic analysis was conducted to identify the key themes in the data. However, there were only five informants, thus the themes described in the thesis should be developed further with a larger sample, considering five people can not be used to generalise to all students' perceptions. Moreover, it might be that the students that would be more hesitant to chatbots would not want to participate in the study, considering the time investment and the lack of interest in the technology. The students who are pro chatbot technology might also be more interested to participate in the study. Therefore, a real-life implementation in a class followed by a bigger thematic analysis of all users might be a better approach.

The COVID-19 situation forced different adaptations to be made in the method. The author's intention was to interview the informants face to face. Also to interview more than five students. After the outbreak of the pandemic, some informants became hesitant to participate, with some informants canceling their interviews and only three agreeing to do online interviews instead. The authors were hesitant towards executing the interviews online, considering there was a live demonstration involved. Although this was inevitable as Sweden's recommendations regarding the pandemic were to reduce physical contact as much as possible. Three out of five interviews were held online instead of face to face and instead of interacting with the chatbot, the authors had to share their screen. The authors had to present the different scenarios with pictures to the informants and at the same time give a verbal explanation. This can affect the informants in a way that they answer, compared to if the interview was conducted face to face.

The understanding of the chatbot might be different if it is explained through pictures online than actually interacting with a real software. Thus affecting the result of this study. Moreover, the context can have an effect as well. The fact that the interview is online might make informants more laid back and perhaps too comfortable when they are at home in front of their computer. The fact that the authors had appointments on school property made the interviews more serious and that can affect the informants and how they choose to reply. Both the context and the understanding of the bot could have been affected by COVID-19, but it is difficult to say how much, or if it had any effect at all. Although it would be a more reliable study if all the interviews were conducted face to face or online instead of mixing the two methods. This was unfortunately not possible due to the low interest of participating after the COVID-19 outbreak.



7 Conclusion/Contribution

The purpose of this study was initially to investigate student's perception of various chatbot functionality and what benefits and limitations they perceive to experience in higher education. By doing that the research questions were formed and then answered.

What benefits do students perceive to experience from an educational chatbot in higher education?

What limitations do students perceive to experience from an educational chatbot in higher education?

The study shows that personalized AI functionality can be beneficial but are not the student's main interest in educational chatbots. On the contrary, simpler functionality such as a Q&A or quiz functionality was perceived to have a strong relative advantage compared to their current study technique. For example, saving time by decreasing obstacles and enhancing the learning process through a quiz chatbot. The hesitation towards chatbots grew for different reasons as the functionality increased. The need for teacher's involvement and compatibility with the courses was vital amongst the students for them to see the value for all the functionality. Moreover, the students showed no clear understanding of where the personalized AI functionality would provide value in the study process. With these findings, the authors recommend for future real educational chatbot implementations to first start small with just Q&A or quiz functionality. Mostly because it is simpler functionality, thus easier for the students to understand how it would benefit them by saving time or enhancing the learning process.

7.1 Future work

A more large scale study would benefit the research society. This to easier generalize the result or find more elements that affect the perceived benefits and limitations. Other types of education could also be investigated. Kindergartens, high schools, or special education would be interesting, to see if the same elements apply or if different aspects matter more. It would also be interesting to see if other studies could be conducted based on the four elements discovered in this thesis. To see other studies using our four elements as a theoretical framework when investigating chatbots and the perceived benefits in higher education.

A quantitative study could be conducted. Instead of interviews looking for original thoughts about chatbots, questions could be asked to back up the findings in this study on a broader scale. Since we just investigated the perceived benefits and limitations, an evaluation study that examines the benefits and limitations after a real implementation would be interesting to see. This would eliminate the "perceived" aspect out of the picture and really focus on how they feel.



References

- Alavi, M. & Leidner, D. (2001). knowledge management and knowledge management systems: conceptual foundations and research issues [1, 2].(Review). *MIS Quarterly*, 25(1), p.107.
- Alepis, E. & Virvou, M. (2011). "Automatic Generation of Emotions in Tutoring Agents for Affective E-learning in Medical Education." *Expert Systems With Applications*,38(8), pp. 9840-847.
- Ayedoun, E., Hayashi, Y. & Seta, K. (2015). A Conversational Agent to Encourage Willingness to Communicate in the Context of English as a Foreign Language. *Procedia Computer Science*, 60(1), pp.1433–1442.
- Bickmore, T., Schulman, D. & Sidner, C. (2013). Automated interventions for multiple health behaviors using conversational agents. *Patient Education and Counseling*, 92(2), pp. 142–148.
- Brinton, C., Rill, R., Sangtae, H., Mung, C., Smith, R. & Ju, W. (2015). Individualization for Education at Scale: MIIC Design and Preliminary Evaluation. *IEEE Transactions on Learning Technologies*, 8(1), pp.136–148.
- Caporarello, L., Magni, M. & Pennarola, F. (2016). Is Technology Mediated Learning Really Improving Performance Of Students? *EAI Endorsed Transactions on e-Learning*, 3(12), pp.1–11.
- Fryer, L.K., Ainley, M., Thompson, A., Gibson, A. & Sherlock, Z. (2017). Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and Human task partners. *Computers in Human Behavior*, 75, pp.461–468.
- Fryer, L.K., Nakao, K. & Thompson, A. (2019). Chatbot learning partners: Connecting learning experiences, interest and competence. *Computers in Human Behavior*, 93, pp.279–289.
- Georgia Tech (2020). The World's First AI Teaching Assistant Turns 4. <https://www.cc.gatech.edu/news/631545/worlds-first-ai-teaching-assistant-turns-4> [2020 - 05 - 25]
- Jacobsen, D.I. (2018). Hur genomför man undersökningar?. 2. ed., Lund studentlitteratur.
- Lebeuf, C., Storey, M. & Zagalsky, A. (2018). Software Bots. *IEEE Software*, 35(1), pp.18–23.
- Moore, G. & Benbasat, I. (1991). Development of an Instrument to Measure the Perceptions of Adopting an Information Technology Innovation. *Information Systems Research*, 2(3), pp.192–222.
- Nadarzynski, T., Miles, O., Cowie, A., & Ridge, D. (2019). "Acceptability of Artificial Intelligence (AI)-led Chatbot Services in Healthcare: A Mixed-methods Study." *Digital Health* 5 : 2055207619871808.



- Palanica, A., Flaschner, P., Thommandram, A., Li, M. & Fossat, Y. (2019). Physicians' Perceptions of Chatbots in Health Care: Cross-Sectional Web-Based Survey. *Journal of Medical Internet Research*, 21(4).
- Perrault, R., Shoham, Y., Brynjolfsson, E., Clark, J., Etchemendy, J., Grosz, B., Lyons, T., Manyika, J., Mishra, S., & Niebles, J. (2019). *The AI Index 2019 Annual Report*, (AI Index Steering Committee, Human-Centered AI Institute), Stanford, CA: Stanford University.
https://hai.stanford.edu/sites/default/files/ai_index_2019_report.pdf
- Radziwill, N. & Benton, M. (2017). Evaluating Quality of Chatbots and Intelligent Conversational Agents. *arXiv.org*, pp.arXiv.org, Apr 15, 2017.
- Russell, S. & Norvig, P. (2009). *Artificial intelligence : a modern approach* 3.ed., Global., Harlow: Pearson Education Limited;
- Stone, P., Brooks, R. Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2016). *Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence* (Report of the 2015-2016 Study Panel), Stanford, CA: Stanford University.
https://ai100.stanford.edu/sites/g/files/sbiybj9861/f/ai100report10032016fnl_singles.pdf
- Söllner, M., Bitzer, P., Janson, A. & Leimeister, J.M. (2017). Process is King: Evaluating the Performance of Technology-mediated Learning in Vocational Software Training. *Journal of Information Technology*, 33(3), pp.233–253.
- Vlach, H.A. & Sandhofer, C.M. (2012). Distributing Learning Over Time: The Spacing Effect in Children's Acquisition and Generalization of Science Concepts. *Child Development*, 83(4), pp.1137–1144.
- Winkler, R. & Söllner, M. (2018). Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. *Academy of Management Annual Meeting (AOM)*. Chicago, USA.
- Yampolskiy, R. (2012). AI-complete, AI-hard, or AI-easy - Classification of problems in AI. *CEUR Workshop Proceedings*. 841. 94-101.
- Zary, N., Sharit, J., Frontoni, E., Shorey, S., Ang, E., Yap, J. & Chui, C. (2019). A Virtual Counseling Application Using Artificial Intelligence for Communication Skills Training in Nursing Education: Development Study. *Journal of Medical Internet Research*, 21(10)



Appendices

Appendix 1: Interview Guide

1. Section 1 - Introduction

Our purpose with the survey is to investigate the acceptance, perceived usability, and requirements of chatbots in education at the university level. The information you provide will be completely confidential. Your name will not be mentioned in our study and no one will be able to link the information you have provided to you.

1.1 Section 1 - Introduction

A chatbot is a software that facilitates a human conversation with a user in chat format, which aims to improve and expand current services. For example, customer service or replace user interface where the task can be performed faster in chat. In education, it often aims to help both the student and teacher as an extra tool in education, but the teacher still has a very important role. Chatbots can be structured in different ways. They can only use question and answer or it can be chatbots that focus on a goal, e.g. to book a movie. It can also be common to use buttons or similar in the conversation to get an easier flow. Objectives for chatbots differ from chatbot to chatbot, but we will focus on functionality that aims to improve and expand learning for students at the university level.

We will go through different scenarios with a chatbot that has different behaviour and functionality, the content is linked to the design of the database course that you are currently a part of. The functionality is fake and it is a predetermined scenario since we did not have time to develop a chatbot. We, therefore, ask you not to take into account other things such as appearance or personality of the chatbot as it is something that has not been the focus or that we had the opportunity to influence.

1.2 Before scenarios

1. Have you used a chatbot before?
2. What are your immediate impressions and thoughts on chatbots?
3. What functionality would you like to see in a chatbot in your education if one were to be implemented?
4. How does your current process look when you are preparing for the exam or doing work?
5. If not addressed in the answer to the previous question: Do you study in long sessions intensively prior to the exam or do you usually divide the work over a longer amount of time?
6. How do you see the current way teachers provide you the information in the courses, ie through lessons, books, and the course management system MyMoodle



2. Section 2 - Scenarios

2.1 Scenario 1 - Q&A Functionality

In this kind of chatbot, there is the only functionality that answers the student's questions with a standard answer. That is, asking a question will then get the same answer every time, basically as a search on google, though your teachers have posted all the answers and questions from the old exams, all the common questions that the teacher usually gets in the course and so on.

2.1.1 Interviewers show demo and example through the chatbot of scenario one.

2.1.2 Follow up

- What advantages or disadvantages do you see compared to your current way of studying?
- Do you think you would use such a chatbot for studying?
- If not addressed in the previous answer: Why or why not?
- Do you think that it is the information and easier access to the information (eg all the exam answers is provided in the bot) that provide the value or is it the chatbot as a service/assistant that provides the value?
In other words, would you still want the information but rather have it via some other channel instead?

2.2 Scenario 2 - Quiz - Functionality

Scenario two is similar to everything from the previous chatbot with questions and answers although, in this version you can be quizzed through the chatbot, it knows which questions you have not answered in a long time and constantly increases the difficulty of the questions the more you answer correctly. If there are any questions or topics that you specifically answered incorrectly, it focuses on learning them. If you answer half the question correctly, it can identify which parts you missed and give hints for what it was you missed. So think of it like the teacher has submitted E, C, and A questions to the chatbot, and the chatbot can quiz you and have a conversation about what you missed, etc.

2.2.1 Interviewers shows a demo example through the chatbot of scenario two.

2.2.2 Follow up

- What advantages or disadvantages do you see compared to your current way of studying?
- Do you think you would use such a chatbot for studying?
- If not addressed in the previous answer: Why or why not?

If the chatbot is not addressed as giving the primary value in scenario 1:



- Do you think that it is the information and easier access to the information (eg all the exam answers is provided in the bot) that provide the value or is it the chatbot as a service/assistant that provides the value?
In other words, would you still want the information but rather have it via some other channel instead?

Scenario 3 - Personalized AI-driven bot

In this scenario, there will not be a demo since it is more about adaptability to the student. Think of it as the functionality as scenario 1 and 2, but with AI that asks relevant questions that the student needs to train on, as well as find the right level of difficulty that suits you, with explanations and functionality that suits you and the way you learn. The more the student uses the chatbot it collects your data, adapts to your learning style, and acts accordingly. So for example, if you like long explanations, short explanations, or specific examples it provides that depending on your data and previous usage.

2.3.1 Follow up

- What advantages or disadvantages do you see compared to your current way of studying?
- Do you think you would use such a chatbot for studying?
- If not addressed in the previous answer: Why do you think that or why do you not think that?

If the chatbot is not addressed as giving the primary value in scenario 1 or scenario 2:

- Do you think that it is the information and easier access to the information that provides the value or is it the chatbot as a service/assistant that provides the value?
In other words, would you still want the information but rather have it via some other channel instead?

2.4 Scenario 4 - Spacing effect

Spacing out your studies over a longer time and start studying in time is something that might be hard to do and that is easier said than done. The spacing effect refers to a learning phenomenon that demonstrates that spacing out study sessions often are more effective for remembering information into long-term memory. Compared to massed studying, like studying the entire day before an exam. This spacing effect can be achieved for example with a flashcards app that sorts the cards depending on how good you know them. So for example, if you answer a question correct two days in a row the app might not ask that question for 3 days.

Think of the same scenarios and functionality as scenarios 1 and 2, but this time the chatbot gives you instruction each evening for studying with a quiz by the chatbot and thus helps you space out the studying. Think of it as something to help you get in 10 minutes studying each day.



2.4.1 Follow up

- What advantages or disadvantages do you see compared to your current way of studying?
- Do you think you would use such a chatbot for studying?
- If not addressed in the previous answer: Why do you think that or why do you not think that?

3. Section 3 - After scenarios

- What do you think are the biggest benefits of these chatbots we have discussed in this interview?
- What functionality do you think should be a minimum added if such chatbot should be implemented?



Appendix 2 – Question & Answer functionality:

what is an sql injection?

[Inspect](#) You

It is one of the most common web hacking techniques. It is when a user tries to write malicious code in SQL statements when the website is asking for something else, like a username. So for example writing in "DROP TABLE Suppliers"

When asked for something else.

[How to prevent SQL injection](#)

Design av databas kursen (Test) at 12:20 PM

How to prevent SQL injection

[Inspect](#) You

You can prevent SQL injection with a parameterized SQL statement

[Show me a parameterized SQL statement!](#)

Design av databas kursen (Test) at 12:20 PM

[Show me a parameterized SQL statement!](#)

[Inspect](#) You

"INSERT INTO MyGuests (firstname, lastname, email) VALUES (?, ?, ?)"



Appendix 2 – Quiz functionality:

quiz me [Inspect](#) You

Sure I think you need to practice on SQL statements, it's a central part of the coming exam.

Sure quiz me on SQL statements

Design av databas kursen (Test) at 12:17 PM

Sure quiz me on SQL statements [Inspect](#) You

Great! Explain the UNION operator and how it's used in SQL statements?

Design av databas kursen (Test) at 12:17 PM

The UNION operator is used to combine the result-set of two or more SELECT statements. It has to be the same datatypes. [Inspect](#) You

Sure, you forgot one key thing though, you want a hint?

Give me a hint

Give me a hint [Inspect](#) You

The thing you forgot has to do with the Select statement and the order of the select statement

Design av databas kursen (Test) at 12:17 PM

they have to be in the same order? [Inspect](#) You

Correct!



Linnæus University

Department of Informatics

351 95 Växjö / 391 82 Kalmar

Tel 0772-28 80 00

Lnu.se