Vaasan yliopisto
UNIVERSITY OF VAASA

Joona Mäkinen

Consumer Perceptions of AI-driven Chatbots in Digital Platforms

Factors Affecting Service Quality and Satisfaction

School of Management
Master's thesis in Strategic
Business Development

Vaasa 2024

UNIVERSITY OF VAASA
School of Management
Author: Joona Mäkinen
Title of the Thesis: Consumer Perceptions of AI-driven Chatbots in Digital Platforms : Factors Affecting Service Quality and Satisfaction
Degree: Master of Science in Economics and Business Administration
Programme: Strategic Business Development
Supervisor: Anni Rajala
Year: 2024 Pages: 70

ABSTRACT:
In the digital age, artificial intelligence (AI) has become a significant part of how customers interact with online store assistants and customer service. Despite AI chatbots' impact on service delivery speed, cost efficiency, and around-the-clock customer support, the impact on customer satisfaction and service quality remains poorly comprehended. This thesis uses hybrid quantitative and qualitative methods to answer the research question on drivers and factors affecting perceived service quality and satisfaction. Additionally, providing insights for managerial considerations for the future development of AI chatbots. This thesis found that the results are in line with previous research on models TAM, UTAUT and SERVIQUAL. Concluding that performance, ease of use and reliability are correlated and in causal relation with the perceived satisfaction and service quality. Enhancing capabilities in these areas would positively affect the experienced satisfaction and service quality. Current overall impressions and attitudes towards AI chatbots are critical, and most customers would rather speak with a human customer service agent. Yet, there is an indication of better applications in the future. The issues, especially with the performance, should be addressed to provide accurate, fast and valuable information for the customer. Currently, the performance is perceived as lacking when it comes to more complex problem-solving and understanding. The AI chatbots should be able to actually solve customers' problems. Otherwise, they will have to seek human assistance. Future development, especially in generative artificial intelligence, could be seen as an opportunity for businesses to solve more complex customer problems with precision and speed, leading the better customer experience in digital platforms. In a complicated digital world, consumers appear to appreciate the ease of use in AI chatbots, advising businesses to focus on customer-centric design and UX/UI of the chatbots. Reliability and trust in a company's compliance with ethical operating models and data privacy also affect the experience with the service provider. This research proposes further future research into factors that affect the perceived service quality. The suggestion is to focus on anthropomorphism and personalisation of AI chatbots since there can be found factors to better identify since this research did not manage to find consistent data on the issue.

KEYWORDS: Service Quality, AI, Chatbot, Satisfaction, TAM, E-SERVIQUAL, Development

Contents

Figures

Tables

1 INTRODUCTION

In the rapidly technologically advancing world, artificial intelligence (AI) has gained popularity on digital platforms, affecting how customers interact with businesses online. AI is implemented into online store assistants and customer service chatbots. Businesses use AI chatbots to provide automated customer support fast, cost-efficiently and with around-the-clock availability. Yet, the consideration of how these AI Chatbots affect service quality and customer satisfaction remains speculative, presenting a significant gap in both academic research and practical business application.

Despite the promising capabilities of evolving AI, there is a gap in understanding how this adaptation is related to perceived service quality and satisfaction. The relevance of this thesis is to apply service quality and technology adaptation frameworks in the context of AI chatbots in digital platforms. The focus is on customers' perceptions to gain better insights into the customer-centric future development of AI chatbots, particularly how the solution is received by users since AI chatbots are often compared to interactions with human service agents. This thesis aims to fill the academic gap in the literature and provide insights into the future development of AI chatbots in business applications.

This research aims to answer, in the context of AI chatbots, the main research problem: "*What factors are perceived to influence the adoption of new AI technology into use, create satisfaction, and affect the service quality in digital platforms?".* Answering the research question provides valuable knowledge on drivers for satisfactory service, perception of quality service and increase the overall use of the solution.

This thesis combines quantitative and qualitative methods to establish insights from the research survey conducted. Data analysis is used to measure different factors established in the theoretical framework to validate the research model. Qualitative survey questions are added to support findings.

The structure of this thesis consists of first introducing the previous literature from the research area and establishing a theoretical framework. Then, the methodology is described in detail before analysing the results. In the end, the thesis discusses the topic and concludes the research with additional managerial considerations.

2 THEORETICAL FRAMEWORK

In this chapter, the paper introduces the foundation for the research and an established framework to evaluate the use of AI in digital services. To answer the research question, it is fundamentally important to comprehend human psychology and reasoning when it comes to accepting new technologies into usage—let alone delivering valuable customer experience in the platform.

2.1 Theory of planned behaviour

The theory of planned behaviour (TPB) by Icek Ajzen (1991) is a cornerstone of modern psychology and a widely accepted model of aspects that affect individuals' attitudes towards something. Theory suggests that attitudes, subjective norms, and perceived behavioural control determine the perception (Ajzen, 1991).

Previous experiences re-frame the attitudes toward the situation at hand. Subjective norms from peer pressure, social pressure, and empirical perceptions of the environment from social circles form the behaviour. Perceived behavioural control refers to the ease and difficulty of performing a certain way (Ajzen, 1991). The theory is applicable in the context of automated customer service. Previous experiences and attitudes with AI and chatbots will form behaviour for future use of these digital services. The simplicity and usability of the solution will affect customer behaviour. Also, a person's environment and social circles will be a part of dictating the eagerness to use these tools.

2.2 Technology acceptance model (TAM)

The first frameworks were formed during the rise of computing at the beginning of the 1990s. Theories are parallel with the TPB model by focusing on technology perception. Fred Davis et al. (1989) introduced the Technology Acceptance Model (TAM), which is currently relevant in the field. The model aims to provide insights into how users accept and adapt to new technology.

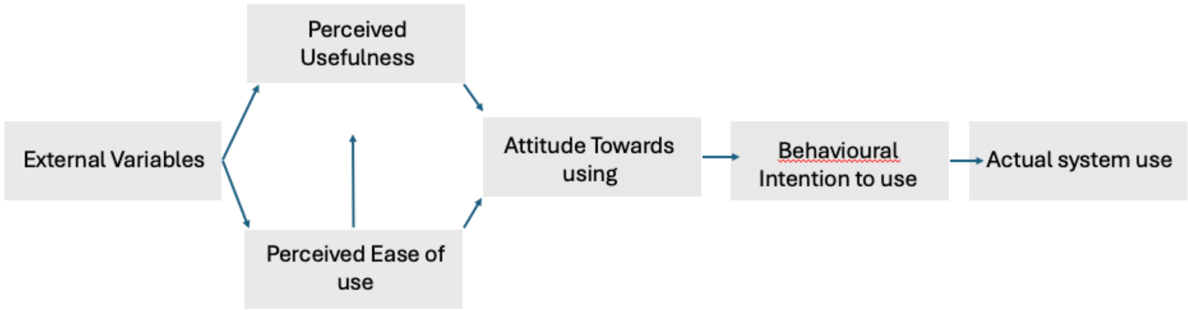


Figure 1 Technology acceptance model (Davis et al., 1989).

The TAM framework (Figure 1) gives a basic understanding of issues affecting the use of a technical system in general. The key elements differing in TAM compared to the TPB model are Perceived Usefulness (PU) and Perceived Ease of Use (PEOU) (Ajzen, 1991; Davis et al., 1989)

PU refers to the perceived value the technological solution potentially offers its users. The paper argues that if the solution is perceived as useful for its purpose and offers value to its user, it is more likely to be accepted (Davis et al., 1989). In the context of chatbots and delivery channels, it comes to the end of the question: How much value is the customer getting by using the service provided? PEOU, on the other hand, refers to the user's belief in the technology's capabilities. If the technology is easy and convenient, it would be more likely to be accepted (Davis et al., 1989). These attributes, based on David F. et al., form the Attitude Towards using (ATU) aligned with the Behavioral Intention of Use (BI), which creates the actual use of the technological system (Davis et al., 1989).

As the TAU model has received consensus from the scientific community, the research has tried to deepen the knowledge of the TAU model into the Unified Theory of

Acceptance and Use of Technology model (UTAUT), focusing more on the effect of behaviour intention (Venkatesh et al., 2003). The updated model consists of the same stages as the TAU -model in Figure 1. UTAUT adds more potential factors to the model, Figure 2., to better explain behavioural intentions when using technology.

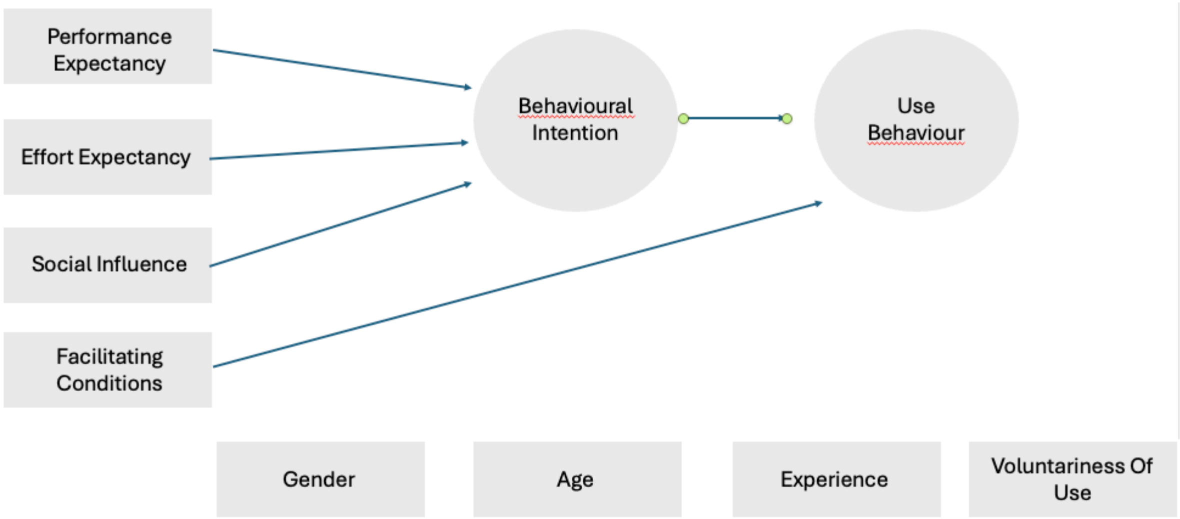


Figure 2 Behavioural Intentions in the UTAUT model (Venkatesh et al., 2003).

The study concludes that UTAUT attributes presented in Figure 2. explain statistically as much as 70% of the variation in behavioural intentions (Venkatesh et al., 2003). It found that performance expectancy is stronger in men and younger people. Effort expectancy is stronger for women, older people, and people with limited experience. Social influence is the strongest among women, older people, and less experienced people. Mandatory use also affects the conditions. Other statistically significant findings from the study have not been made (Venkatesh et al., 2003). Even though the results are statistically significant, they must be examined in cultural contexts cautiously. The study suggests future research on the topic but also wonders how comprehensive studies could be made examining these selected variables as explanatory factors.

2.3 Expanding the TAM -model into Information Technology

The TAM model was later developed and formed to explain various technology acceptance situations. Considering the complexity of human behaviour, it could be hard to find one unified theory to explain the use of technology. The variables used to measure the likelihood of using the solution provided can also differ. Lee et al. (2003) examined the TAM model more in the context of information technology aimed at the end-users. The research found significant factors as variables (Picture 1.) to enhance the TAM model's applicability in evaluating acceptance.

Perceived enjoyment	End-user support	Experience
Accessibility	Anxiety	Attitude
Compatibility	Complexity	Result demonstrability
Facilitating conditions	Image	Relevance
Managerial support	Playfulness	Personal innovativeness
Relative advantage	Self-efficacy	Social influence and pressure
Social presence	Trialability	Subjective norms
Visibility	Voluntariness	Usability

Figure 3 TAM variables in the Information Technology (Lee et al., 2003)

The variables presented in Picture 1. are not comprehensive but possible variables that have received consensus in the scientific community. One factor in the acceptance of technological solutions (Lee et al., 2003). The main problem with the TAM model itself is the challenge of measuring, in practice, how much variables affect the result. Lee et al. (2003) measure the indicated use of technology, not the actual use. The study does not provide the level of effect on variables or correlation between variables. It merely raises awareness of the need to focus on these variables when developing information technology systems.

Dwivedi et al. (2019) pondered the future of the TAM model and its relations to the TPB model. Even though the model will not fully explain technology acceptance, it will give perspective on issues affecting the solutions' actual use. The TAM stays relevant when combined with TPB since, in business, companies' monetary investments in information technology are significant (Dwivedi et al., 2019). Investment in fancy digital projects and tools could waste resources if the products are never used.

2.4 Expectation-Confirmation Theory in Information Technology

Expectation-confirmation theory (ECT) takes a similar view as previously introduced theories on accepting new technology into use. The model is extensively used, especially in consumer behaviour and customer satisfaction literature (Bhattacherjee, 2001). The focus in ECT is customer satisfaction with the solution provided in the context of information technology; in Figure 3. the model is presented.

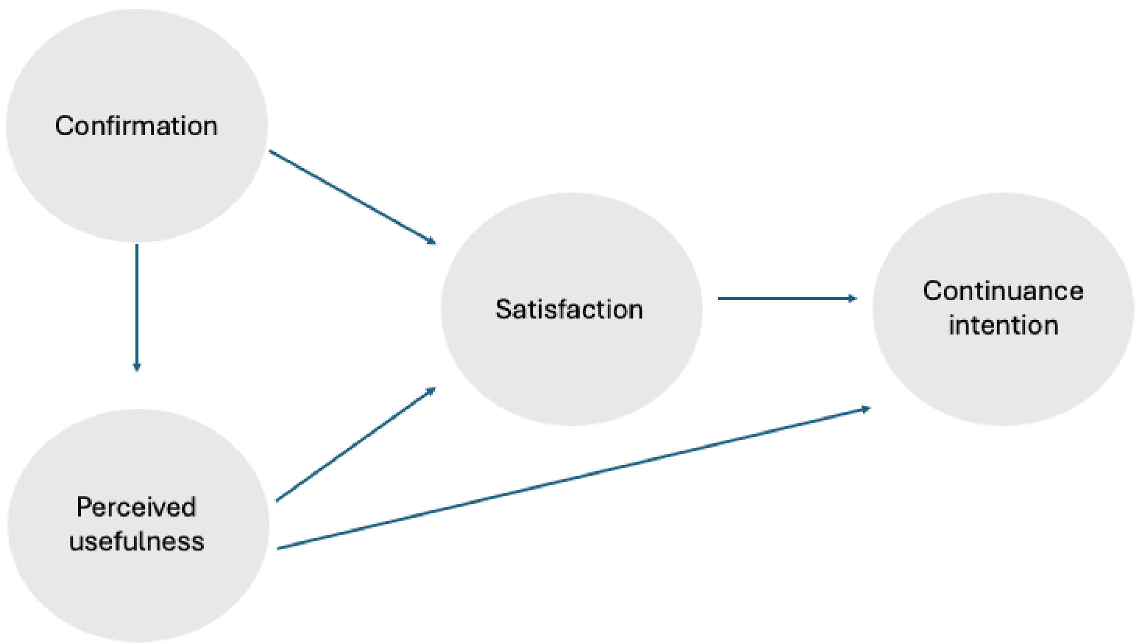


Figure 4 Expectation-confirmation theory (ECT) (Bhattacherjee, 2001).

In the ECT, consumers have some pre-expectations about the product or service offered. The extent to which these expectations are met will form the perception of usefulness. Perceived usefulness and satisfaction correlate; usefulness will positively correlate with satisfaction and vice versa. Satisfaction with the solution will determine the actual usage of the solution. (Bhattacherjee, 2001)

Compared to TAM, the ECT focuses on post-acceptance and variables to analyse how well the solution met customers' expectations. This allows the model to predict the future use of the solution. The model provides a framework to better understand customers' expectations, satisfaction, and confirmation/disconfirmation (Bhattacherjee, 2001).

2.5 Service quality and SERVIQUAL dimensions

The concept of quality service is an extensively researched topic in business literature. What does the quality of service mean, and what creates exceptional quality service? Parasuraman A. et al. (1985) established a foundation for the service quality literature by introducing the core of the SERVIQUAL -model with its dimensions. The model is presented in Figure 4.

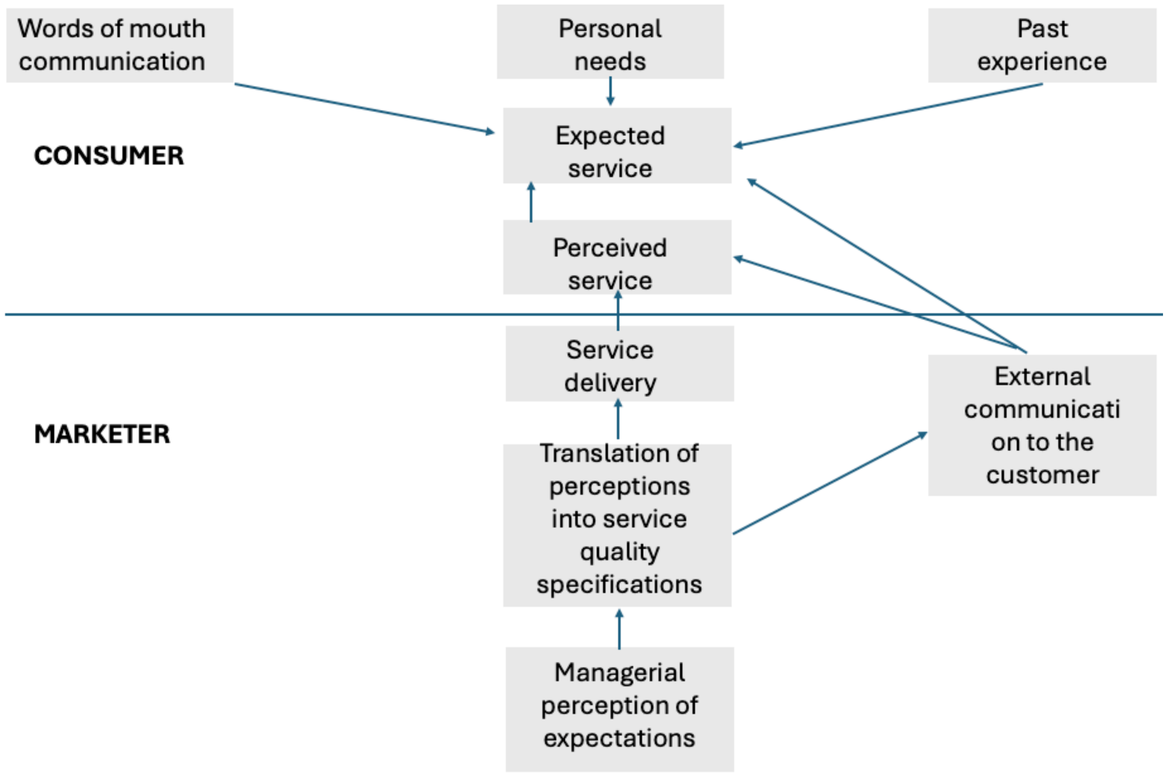


Figure 5 Foundation of SERVIQUAL model (Parasuraman et al., 1985).

The model introduced the concept of measuring the quality of service through the gap between consumers' perceptions and expectations of service quality (Parasuraman et al., 1985). Considering the model presented in Figure 4. with the TAM model and ECT, the role of consumer expectations and perceived value of service define the quality and satisfaction of the service. The research points out the subjectivity of the perception of quality of service. Personal needs, past expectations, and word of mouth from the service will create expectations about the quality of the service (Parasuraman et al., 1985). On the other hand, the quality of service that is experienced depends on how well the provided service meets personal expectations (Parasuraman et al., 1985).

The study aims to explain service quality through dimensions that affect the experience of service (Parasuraman et al., 1985). In Figure 5, the dimensions affecting the perception of quality are listed.

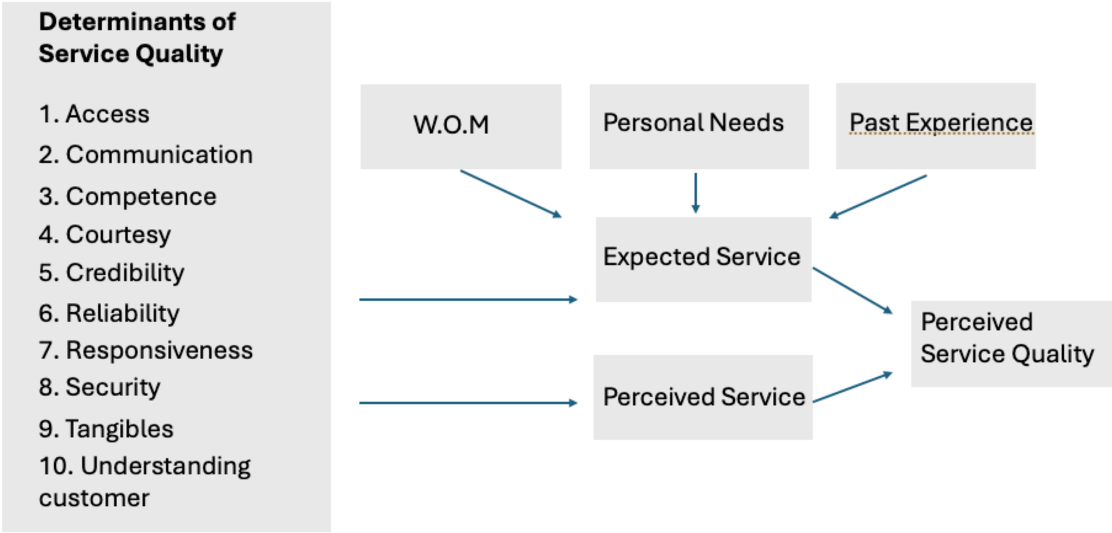


Figure 6 Dimensions of service quality (Parasuraman et al., 1985).

Parasuraman et al. (1985) focused on the main dimensions of reliability, assurance, empathy, responsiveness, and tangibles. Reliability is in parallel with the consistency of service performance and dependability. Assurance depends on service agents' knowledge, courtesy, and ability to inspire trust with confidence. Empathy relates to individualising service and personal attention and care for the customer. Responsiveness is the willingness and ability to provide service to the customer without delays. In classic theory, Tangibles are the physical facilities where the service is delivered and the overall appearance of the environment. (Parasuraman et al., 1985)

2.6 Expanding SERVIQUAL into the digital platforms

Considering the Parasurama et al. (1985) foundation of the concept of service quality, it cannot be directly applied to the modern digital economy. Many services are delivered on digital platforms, and it is important to distinguish key differences.

In business literature, the term e-Service Quality (e-SQ) has emerged in recent years (Zavareh et al., 2012). The rapid growth of services on the Internet and new channels of communication with consumers have raised the importance of e-SQ theories in identifying key variables affecting the perceived quality of service on digital platforms (Zavareh et al., 2012). Zavareh et al. (2012) suggest updating SERVIQUAL dimensions around information technology into e-SERVIQUAL. Updated dimensions evolve around the efficiency and reliability of services, fulfilment, security and trust, site aesthetics, responsiveness, and ease of use (Parasuraman et al., 2005; Zavareh et al., 2012).

Effectiveness and reliability refer to the speed of internet service, technical, functional reliability, and fast navigation in the service provided. Fulfillment provides easy-to-follow pages, accurate value promises made by the service, and accuracy of information presented. Security is a rising concern in digital platforms, and as a variable, it refers to the ability to create trust in service safety, the right use of personal information, and confidence in data privacy. Site aesthetics can also affect the fulfilment of using the service. Proper user interface (UI) design will make e-SQ better due to the pleasing design that is visually pleasing. UI design also affects the effectiveness of the service since easy navigation through the service makes the use easier and faster. Responsiveness refers to access to the service, support, and availability of the platform. Ease of use can be seen in combining various attributes to create a seamless and pleasant user experience in the service, enhancing usability (Parasuraman et al., 2005; Zavareh et al., 2012).

Santos Jessica (2003) has also presented her model on virtual service quality dimensions based on the SERQUAL model. The approach is more technical and takes on the actual measures to enhance e-SQ. The model is presented in Figure 6.

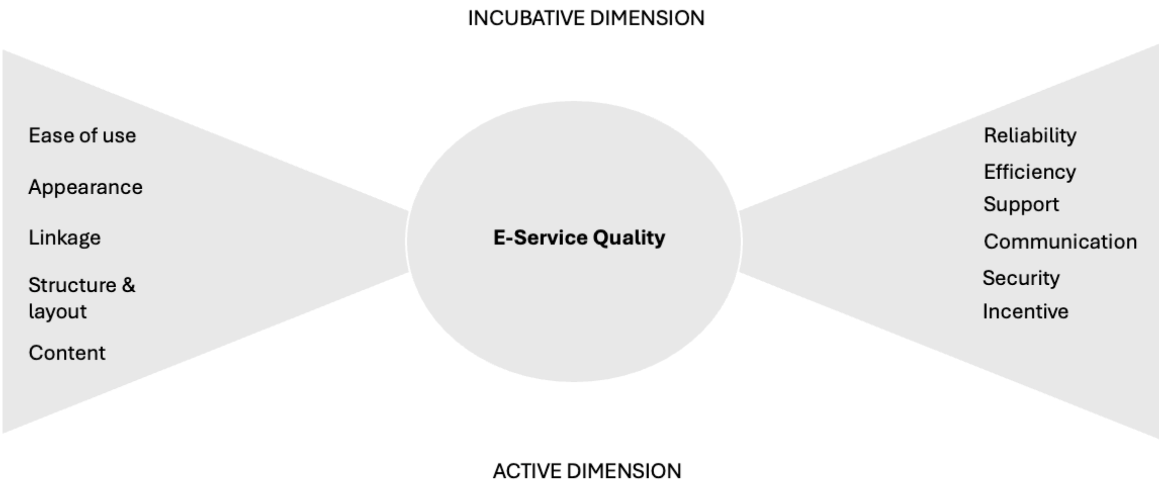


Figure 7 E-Service Quality model (Santos, 2003).

The environment of digital services is different from the traditional view of what the service is and should be considered to have different dimensions (Santos, 2003). Also, Parasuraman (2000) later made a proposition that major positive themes in digital platforms are flexibility, efficiency, and pure enjoyment of using pleasing digital services (Parasuraman, 2000). Negative variables raise security and privacy concerns, as well as a lack of personalisation and obsolescence in the technology in use. Parasuman A. (2005) sees it as an opportunity to create flexible and personalised customised digital experiences that could positively affect the e-SQ. Exceptional e-SQ could be possible by focusing on the suggested variables in digital platforms, paving the way in this digitalised world. (Parasuraman et al., 2005)

Parasuman A. (2005) measured that efficiency, system availability, and fulfilment are the most important qualities to provide positive experiences in digital platforms. Interestingly, the study ranked privacy as the least significant factor. The study also considered customers' loyalty and found significant results on the effect of word of mouth and recommendations from others to be positively correlated with satisfaction with the service (Parasuraman et al., 2005). Parasuman A. (2005) named this extension of SERVQUAL as an E-S-QUAL model, focusing on these mentioned variables.

Considering the different models presented, these are incomplete, but later research is suggested. The main focus is to give insights on what issues to address when planning service delivery in the digital platforms (Parasuraman et al., 2005).

2.7 Artificial intelligence in service

The concept of intelligence can be philosophically disputed, as can the relationship between artificial intelligence (AI) and human intelligence (HI). Artificial intelligence refers to machines exhibiting human-like behaviour and aspects of HI (Huang & Rust, 2018). Currently, AI can be divided into different stages based on technological advancement and functionality. Generally, AI is divided into categories: robotic process automation (RPA), machine learning (ML), and deep learning (DL) (Hollebeek et al., 2021).

Furthermore, in general, AI can be divided into four categories based on its human-like abilities that can be distinguished. Huang M. et al. (2018) categorised stages as mechanical-, analytical-, intuitive-, and empathic intelligence. Stages are presented in Figure 7.

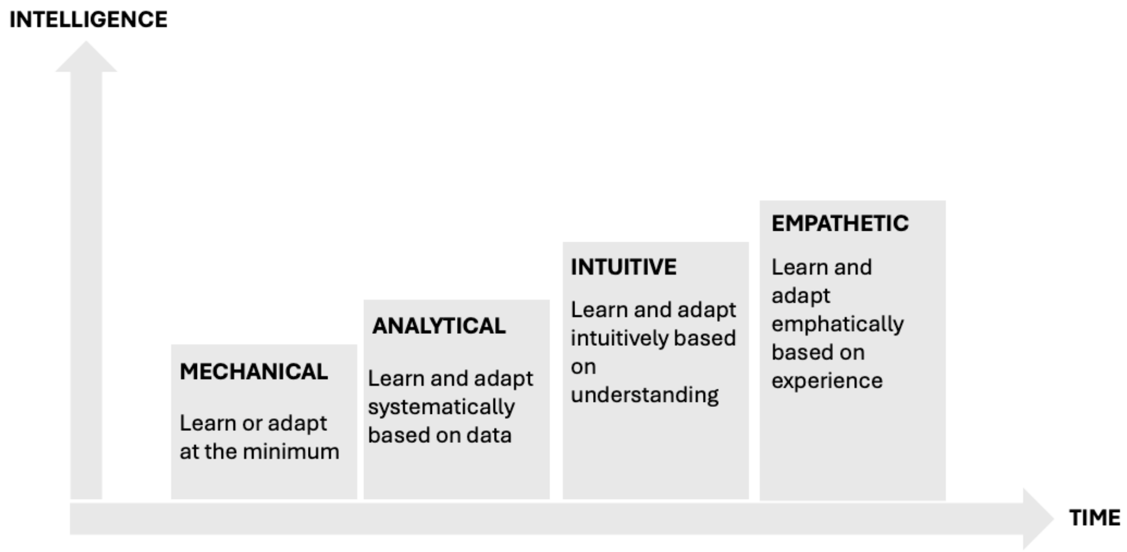


Figure 8 The four intelligences (Huang & Rust, 2018).

The lowest level of artificial intelligence is the mechanical level, where Hollebeek L. et al. (2021) set the RPA, which is purely based on rule-based automation of processes. RPA automates routines by performing repetitive tasks (Hollebeek et al., 2021). When programmed, the RPA is given a set of rules that it follows. It can be purchased as a software solution, where the RPA is set up in a graphic user interface (GUI) to follow the rules the user sets for the robot. To illustrate the point, Figure 8., contains the simplified script of an RPA chatbot that is programmed to look for keywords from customers' messages, and by rules, if-else statements determine the answer to users' messages.

```
def check_input(input_string):
    if 'pillow' in input_string or 'duvet' in input_string or 
'blanket' in input_string:
        print("Here *link* you can find our bedroom products.")
    else:
        print("I´m sorry. I didn´t quite understand. Please ask me an-
other question.")
```

Figure 9 Example of possible RPA Chatbot application in Python

In this example, if the message refers to products often found in the bedroom, based on the rules, it gives the customer the link to the correct webpage in the online store.

Otherwise, it will apologise for not understanding the message. A complete RPA program would contain many more scenarios and rules to determine the answer to the customer.

Analytical intelligence advances from the described RPA robotics level, adding machine learning into action (Huang & Rust, 2018). At this level, responses are no longer determined by simple step-by-step rules but rather by using linear regression and statistical processes to adopt and adjust the action based on its datasets and data points. Decision-making is based on the most probable response (Hollebeek et al., 2021).

Huayn M. et al. (2018) described the last levels of intelligence, Figure 7, as being close to human intelligence when a machine adopts qualities of human emotions and empathy. Intuitive considers some form of creativity in problem-solving. When empathetic intelligence advances to understanding and recognising people's emotions and emotional responses while reacting to those feelings (Huang & Rust, 2018). Deep learning could be seen as close to human intelligence as possible, in the sense of customer experience, since the machines could adopt emotional responses to the customers (Hollebeek et al., 2021).

Future considerations in the service context are whether artificial intelligence can replace human interaction and be considered human enough to provide emotional service.

2.8 Consumer trust in AI assistants

Consumer trust in AI assistants is a crucial component when considering the business-wise applicability of the desired solution. Perception of trust towards AI solutions can create attitudes toward use in the TAM -model (Davis et al., 1989). Also, the solutions that AI assistants provide must be accurate and correct; it could affect the perception of performance according to the TAM (Davis et al., 1989).

From the SEVIQUAL standpoint, trust and security are listed as affective variables in Figure 6 and Figure 7. It can be argued that this affects the perceived quality of service provided (Parasuraman et al., 1985; Santos, 2003).

Khan A. (2021) considers the dimensions that affect trust in AI. The study divides trust into cognitive trust and emotional trust. Cognitive trust considers factors such as transparency, reliability, tangibility, and fast communication responses. Emotional trust is mostly formed through the AI's anthropomorphism (Khan A., 2021). These findings further correspond to earlier established theories. The main finding of Khan A.'s (2021) study is that the expectation of AI operating on commonly accepted moral and ethical grounds creates trust.(Khan, 2021)

Studies have found that trust in AI services is positively correlated with trust in the company that adopts the AI into use. Trusted companies adopting AI services are likelier to be trusted (Frank et al., 2023).

2.9 Human-Computer Interactions and Anthropomorphism in AI

Nass C. et al. (1994) formed the paradigm called “Computers are Social Actors” (CASA), sometimes referred to as Social response theory, that has established fundamental acknowledgements in scientific society. The main idea is that humans tend to interact with computers like real people. This leads to a phenomenon where socially accepted norms, biases, and beliefs apply to the communication between humans and computers. Based on the study, these attributes are shown unconsciously, even though humans understand that computers cannot have feelings (Nass et al., 1994). This can be considered to affect communication between humans and the AI.

The term anthropomorphism in the context of human-computer interactions refers to the human-like behaviour of the computer. The AI can adopt human-like behaviour, mimicking emotions, personalities, or feelings. The anthropomorphism of AI comes into action in the later stages of the four bits of intelligence (Huang & Rust, 2018).

Adam M. et al.'s (2021) study found that adding clues of anthropomorphism into AI chatbots in customer service would increase the likelihood of customers complying with instructions given by the AI. The study notes that many customers still hesitate to comply with AI due to scepticism and resistance against the kind of technology. (Adam et al., 2021)

Human-centred AI (HCAI) adds anthropomorphism to the design of the AI chatbot. The anthropomorphism in AI and customer satisfaction seem to be correlated (Klein & Martinez, 2023). The study examined adding anthropomorphic design clues into customer service chatbots that appeared to add perceived enjoyment and trust and enhance attitudes toward AI (Klein & Martinez, 2023).

The father of modern computing, Alan Turing, presented his famous considerations on the question "Can machines think?" in 1950. The theory known as The Imitation Game, later the Turing test, exhibits the intelligent behaviour of the machine. Machines can be seen as intelligent if humans cannot tell if communication is between another human or a machine (Turing A., 1950). If adding anthropomorphism into AI could enhance customer service chatbots in a way that solution would pass the Turing test, it could raise some ethical considerations. Because humans perceive computer interactions as relational to human communication and add anthropomorphism to increase the trust, satisfaction, and overall customer experience of the service (Adam et al., 2021; Klein & Martinez, 2023; Nass et al., 1994), one cannot dismiss the ethical considerations.

3 METHODOLOGY

Considering the phenomenon's nature, it can be described as a personal perception and attitude toward disruptive technology that creates new ways of communication and transactions with the company. Thus, examining personal perceptions in a vast population would be an appropriate approach to the issue. Due to this, empirical research is conducted using a quantitative questionnaire aiming to evaluate attributes that factor into the research question. This study mixes methods with qualitative research by adding open-ended questions to gain more insights beyond the statistical analysis.

3.1 Sampling Method, Data Collection & Analysis Plan

Since the thesis does not specify a specific industry or customer segment, the sampling of data can be collected from every and any demographic. The plan to collect data is to publish the questionnaire on social media platforms, university students and employees, and different employers. The aim is to get as many replies as possible by diversifying the questionnaire to different platforms to reach different demographics. The goal is to get as broad data as possible from different demographics to establish a holistic view of customers' perceptions.

The research will classify the answers based on the survey responder. Demographics are classified based on age, gender, educational background, and income. Geographically, the survey is targeted at people living in Finland. Classification allows the study to identify if statistically significant patterns can be found based on different demographics.

The survey questions will be formatted as a Likert scale ranging from Strongly Disagree to Strongly Agree. This methodology will provide the survey to gain a more dimensional view based on the answers. Also, the evaluation of what factors are perceived as the most important can be identified. The neutral answers can also be considered, and the research can focus on the key issues raised by the survey.

Multiple-choice questions measure the demographics, and the limits of values will be set based on previously established research practices. Also, the questionnaire will include Boolean parameters. The survey also allows written answers to the questions, where responders can clarify or share thoughts on the issue. Scientifically, the Likert scale will provide the best data to be analysed considering the research question, hypothesis, and previous research. The function of the voluntary open-answering option is to gain insights into issues during the survey and gain considerations for future research. The written answers will be evaluated separately from the Likert and Boolean scale answers.

From the ethical perspective of this research, the answers are collected from the individuals with their consent. The survey does not ask for or use sensitive personal information, and answers are anonymous. Anonymity is used to ensure the integrity of the answers and to gain unbiased opinions of respondents.

The research will use a variety of analysis methods. Descriptive statistics will measure median, mean, and distribution to gain a unified view of the data in total. Frequency distribution will provide a visual approach to the frequency of each category of the Likert scale. Factor analysis will measure similar constructs to identify underlying factors explaining patterns of correlation among variables. This will help find potentially statistically significant results from the survey.

Cronbach´s Alpha can be used to measure internal reliability and consistency by measuring how well a set of items measures a single unidimensional latent structure. Chi-square, Mann-Whitney U, and Kruskal-Wallis H tests evaluate the significance of the results between groups. T-tests can be used for interval data, but their use in Likert scales is limited.

This paper combines these methodologies to broadly analyse the data, aiming to explain consumer perceptions. The data analytics are compared to the presented research model and previous literature on the topic.

3.2 Measures

In this part, the questions and measures are introduced. The questions are categorised into groups: Past Experiences, Expected Performance, Ease of Use, Reliability & trust, Personalization & Anthropomorphism, and Satisfaction.

Demographics give background information about survey respondents and are analysed with the intention of seeing if there is differences on answers based on the demographics.

Age;
Gender;
Education;
Income;

The Past Experience measures the level of exposure to AI Chatbots and previous experiences. Based on the TAM and UTAUT models, past experiences have an effect on the use of technology (Davis et al., 1989; Venkatesh et al., 2003). Also, the SERQUAL and E-SERQUAL have attributes for expectations and how the previous experiences from expectations factor in service quality and satisfaction. (Parasuraman et al., 1985, 2005; Santos, 2003)

PA1 I regularly use AI Chatbots (Camilleri, 2024; Gerlich, 2023)
PA2 I have a basic understanding of how AI and customer service chatbots work (Gerlich, 2023)

PA3 I am comfortable using different AI chatbot services. (Gerlich, 2023)
PA4 I have encountered issues and problems with AI chatbots in the past (Gerlich, 2023)

The Expected Performance factor measures the extent to which the consumer perceives the AI chatbot's performance in achieving their intentions and how well it performs in the task. For the solution to be used, the performance has to be good enough for the user (Davis et al., 1989; Venkatesh et al., 2003). From the experienced service quality perspective, the solution has to provide value with performance in resolving users' issues. The service needs to meet the customer's needs and be useful to be relevant. (Parasuraman et al., 2005; Santos, 2003)

EP1 AI Chatbots provide useful assistance to my problems with the service. (Camilleri, 2024)
EP2 AI Chatbots provide quick assistance to my problems (Camilleri, 2024)
EP3 AI Chatbots make customer service more efficient (Camilleri, 2024)

The ease of use factor measures the effort required to use the AI chatbot. To be adopted into use, the solution needs to be reasonably effortless to use (Davis et al., 1989; Venkatesh et al., 2003). User-friendly design and easiness of use, in theory, have been seen to affect satisfaction. Service is experienced with better quality if these requirements are met (Parasuraman et al., 2005; Santos, 2003).

EU1 It is easy to use AI Chatbots (Camilleri, 2024)
EU1 AI Chatbots answer me clearly and understandably (Mark Anthony Camilleri)
EU3 I feel that AI Chatbots are user-friendly (Mark Anthony Camilleri)

Reliability and trust measure the ethical ways of companies using AI chatbots. From perceived reliability of using the data privately and complying with ethical standards. This can be seen to affect the perceived service quality and overall satisfaction with the solution. (Parasuraman et al., 2005; Santos, 2003)

RT1 I feel AI chatbots are reliable (Camilleri, 2024)
RT2 Assistance that AI Chatbots provide is correct and answers my question. (Camilleri, 2024)
RT3 I feel that AI Chatbots use my personal data securely (Camilleri, 2024)
RT4 I feel companies use AI Chatbots ethically (Gerlich, 2023)

Personalization & Anthropomorphism measure the perception of human-like behaviour and how emotionally tailored the service is. Research suggest that the more emotional and human-like service would increase the satisfaction of use (Klein & Martinez, 2023).

PA1 AI Chatbots are polite
PA2 I feel the AI Chatbot should acknowledge and be able to respond to me with emotion
PA3 I feel more confident if the AI Chatbot has a human name and avatar

Satisfaction measures the individual's willingness to use and keep using the provided solution (Davis et al., 1989; Venkatesh et al., 2003). These questions are determined to answer how satisfied the customer is with the AI chatbot solution and how willing they are to continue using it.

S1 Most likely, I will continue using AI Chatbots in the future (Camilleri, 2024)
S2 The AI is the future of digital customer service services (Gerlich, 2023)
S3 I am satisfied with the service I have received from AI chatbots
S4 I find AI chatbots useful in customer service
S5 AI chatbot service exceeds my expectations
S6 I prefer speaking with a human in customer service chat

For the qualitative part, the following qustions are used to gain as relevant insights as possible.

Q1 Why have you previously been satisfied or dissatisfied with customer service AI chatbots ?
Q2 In some situations, is AI chatbot a better way to contact for issue resolution than conversing with a human ? In what situation ?
Q3 I would be more satisfied with customer service Ai chatbots if…

3.3 Sample

In total, the research survey received 148 answers from people of various demographics, which was sufficient to conduct this research. The survey can be found in appendix 1 in English

The Webropol survey was posted in different Facebook groups, shared on LinkedIn, and shared on various social media platforms. The target audience was Finnish people from different demographics. The questionnaire was mostly answered in Finnish.

3.4 Valuating survey category results and reliability of the results

The questions were ranked into categories, presented in Chapter 3.2: Past Experiences, Performance, Ease of Use, Reliability, Personalization, and Perceived Service Quality. Each question in the category provides a Likert chart value from 1-7, combined towards the total value of the category. Two questions, "*I have encountered issues and problems with AI chatbots in the past*" and "*I prefer speaking with a human in customer service chat*", needed to be reversed to gain the actual score that reflects the participants' intended perception towards the question for the analysis.

The concept of combining and valuating categorized answers, the combined total results need to be validated though reliability analysis. Considering the ordinary results, the method will be Cronbach's Alpha. This value evaluated the inner consistency and reliability of combined results. The values under 0,7 can be seen as unreliable, and hence these could be considered to be discarded from the research.

The Past Experience gained the value of 0.234, Table 1.

Reliability Statistics	
Cronbach's Alpha	N of Items
0.234	4

Table 1 The Past Experiences Cronbach´s Alpha reliability

This statistic is unreliable, and hence the further research based on this would not yield any relevant statistics. The Performance category received Cronbach´s Alpha of 0.919, Table 2.

Reliability Statistics	
Cronbach's Alpha	N of Items
0.919	3

Table 2 Performance Cronbach´s Alpha reliability

This is a reliable and valid result >0,7. Also, the Ease of Use received a value of 0.802, Table 3.

Reliability Statistics	
Cronbach's Alpha	N of Items
0.802	3

Table 3 Ease of Use Cronbach´s Alpha reliability

The reliability statistic of the Reliability category is also reliable with the value of 0,792, Table 4.

Reliability Statistics	
Cronbach's Alpha	N of Items
0.792	4

Table 4 Reliability Cronbach´s Alpha reliability

Yet, the value of Personalization did not have inner consistency and is discarded with the value of 0.333, Table 5.

Reliability Statistics	
Cronbach's Alpha	N of Items
0.333	3

Table 5 Personalization Cronbach´s Alpha reliability

To conclude, the results on reliability, personalisation, and past experiences did not have reliable consistency together. This study suggests discarding and not examining the categories that were below 0,7 Cronbach´s Alpha at this phase. Yet, it could be meaningful for future research. The problem of unreliable results might be the sample size n or the questions applied in the context of AI Chatbots.

Even though the results in grouped categories were inconsistent, these questions can be evaluated separately but not together. A comparison of individual questions will be done later in this research.

3.5 Limitations

Due to the selected research method, some issues need to be considered. The anonymity of the survey always poses a risk to the integrity of the answers. The factor analysis aims to explain one factor at a time to gain the most reliable results and measure the exact effect of the variable. The goal is to eliminate excess factors from the statistics that could falsely affect the result.

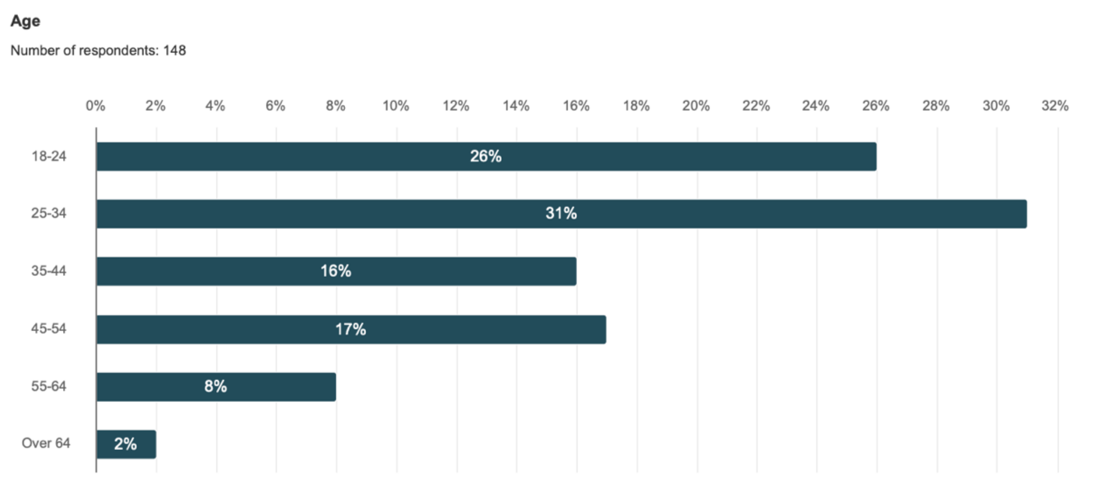
The study is geographically targeted towards the perception of Finnish people. This might limit the scalability of applying results in different cultures and social constructs.

The survey had 148 participants, which can be seen as somewhat small for broader analysis. Partly due to this, some parts of the analysis did not yield significant results, and some parts of the desired analysis had to be discarded. It could be possible that with a larger sample size, the research might have yielded more extensive material for more in-depth quantitative analysis.

Considering the research model itself, the survey questions and the model were based on the currently recognised theories such as UTAUT and SERQUAL models (Davis et al., 1989; Parasuraman et al., 2005; Venkatesh et al., 2003), and applied to the context of AI Chatbots. The transition to the topic and applicability can cause some errors in causality. There could be other factors that explain the perceptions towards AI chatbots. This research established the view that there is a correlation between factors and satisfaction.

4 RESULTS

Considering the number of answers, it is necessary to consider the demographics of the answers to understand better the factors affecting the analysis. The use of data can somewhat limit the possibility of the research making certain assumptions if the data is biased or heavily focused on a certain demographic. The optimal situation would be if the results were normally distributed. Below, in Figure 9, the demographic distribution of all the answers is presented.

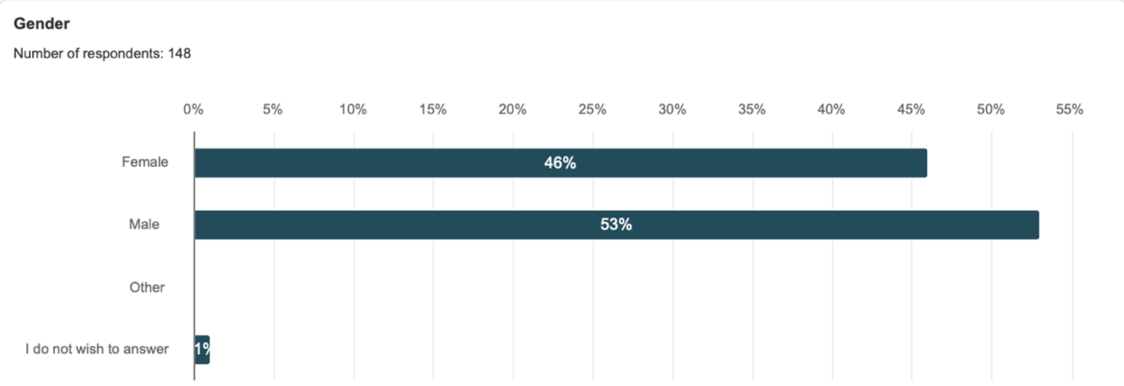


	n	Percent
18-24	39	26,3%
25-34	46	31,1%
35-44	24	16,2%
45-54	25	16,9%
55-64	12	8,1%
Over 64	2	1,4%

Figure 10 Age distribution

As Figure 9 shows, the age is heavily focused on younger persons. Over 50% of the answers were from people under 35 years old. The reason for this height in statistics in participants' age varies a lot, and there would be so many variables to consider accurate reasons for this. The key takeaway is to consider that results are heightened by younger people’s perceptions of AI chatbots. Later, this paper will examine if differences between younger and older people can be found in relation to the specific questions. Yet, since the n value is relatively low, conclusions cannot be withdrawn just based on age.

The gender of the participants was more even, Figure 10.

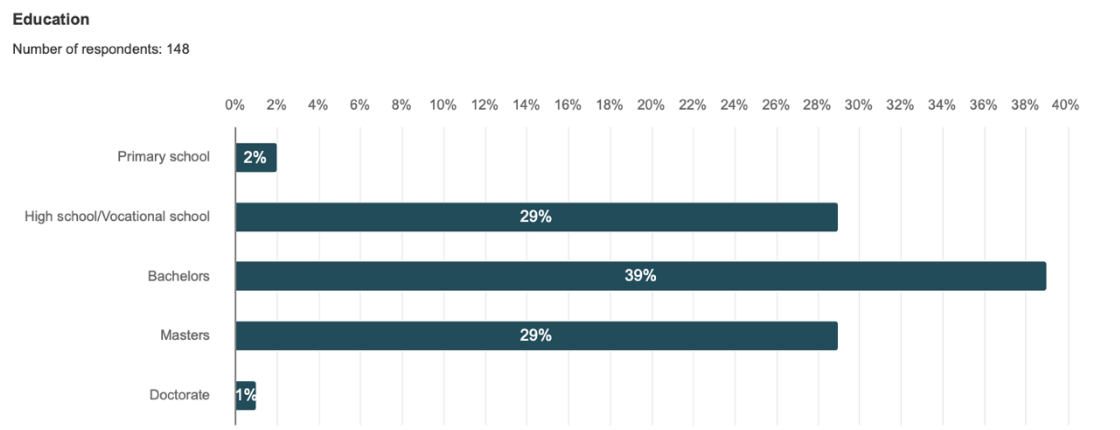


	n	Percent
Female	68	45,9%
Male	78	52,7%
Other	0	0,0%
I do not wish to answer	2	1,4%

Figure 11 Gender of the participants

Gender as an attribute is distributed quite evenly, and from that, some conclusions can be drawn about whether there is a difference between females and males. Other genders were not given any answers to consider in this research.

Educational distribution, Figure 11, has some variance in the answers.

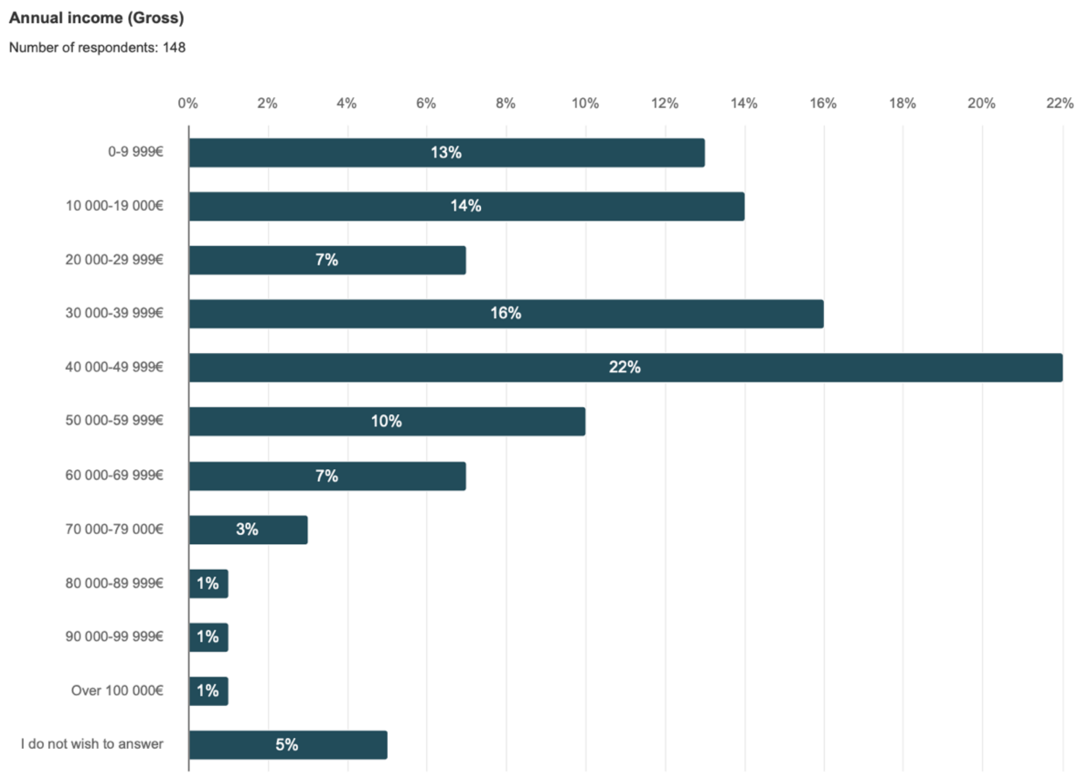


	n	Percent
Primary school	3	2,0%
High school/Vocational school	43	29,0%
Bachelors	58	39,2%
Masters	43	29,1%
Doctorate	1	0,7%

Figure 12 Educational backgrounds of participants.

In Figure 11, It is notable that there were relatively more highly educated participants when considering the total demographics of Finland’s population. Primary school and doctorate education backgrounds had low n values, so the conclusions from there are statistically irrelevant. The bachelor’s and master’s educated participants represent over 50% of the total statistics. This can somewhat weigh the overall results. Based on the answers, meaningful statistics can be found when comparing different educated people in high school & vocational school compared to higher education.

Considering the participants' incomes, the results were leaning towards the left, the lower end of the scale. Figure 12 presents the distribution of the participants' different incomes.



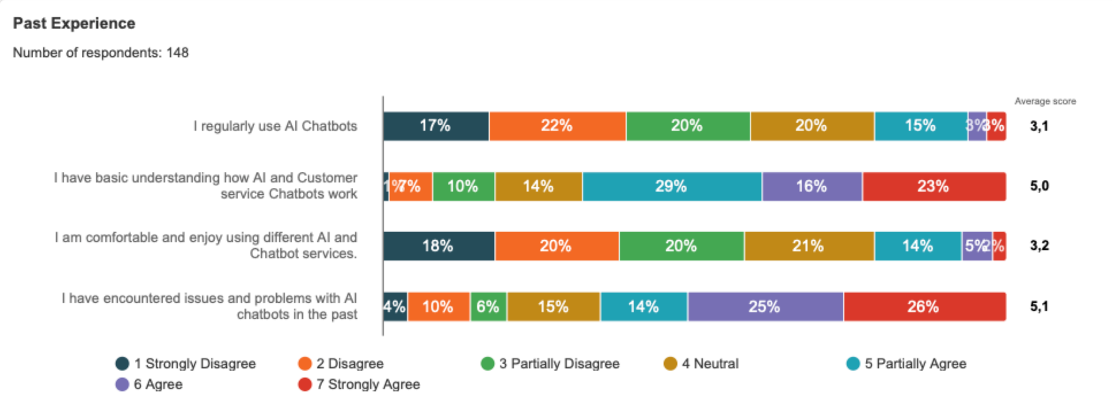
	n	Percent
0-9 999€	19	12,8%
10 000-19 000€	21	14,2%
20 000-29 999€	10	6,8%
30 000-39 999€	24	16,2%
40 000-49 999€	33	22,3%
50 000-59 999€	15	10,1%
60 000-69 999€	10	6,8%
70 000-79 000€	4	2,7%
80 000-89 999€	1	0,7%
90 000-99 999€	2	1,3%
Over 100 000€	2	1,4%
I do not wish to answer	7	4,7%

Figure 13 Income of participants

Considering the distribution in Figure 12, it is possible to conclude that this statistic could represent the actual distribution of income in Finland. The difference between low- and high-income persons' answers in this study would be irrelevant due to the low n-values of high-earners. Grouping the data into groups such as low-income, medium-income, and high-income could give more insight. The notable problem would be that the differences between income stages would be much further away. This paper will later take this issue into account.

4.1 Past experiences

One reason for past experience-related questions is to gain insight into a back group of participants and how well they are familiar with and accustomed to the idea of AI Chatbots. Based on the theories presented earlier, it can be seen that the attitudes towards using the solution and the customer experience related to the topic are affected (Ajzen, 1991; Davis et al., 1989; Venkatesh et al., 2003). First, this paper introduces the answers received in total, Figure 13, and then it goes through them more accurately to distinguish some insights.



	1 Strongly Disagree	2 Disagree	3 Partially Disagree	4 Neutral	5 Partially Agree	6 Agree	7 Strongly Agree	Average	Median
I regularly use AI Chatbots	17,6%	21,6%	20,3%	19,6%	14,8%	3,4%	2,7%	3,1	3,0
I have basic understanding how AI and Customer service Chatbots work	0,7%	6,8%	10,1%	14,2%	29,0%	16,2%	23,0%	5,0	5,0
I am comfortable and enjoy using different AI and Chatbot services.	17,6%	20,3%	20,3%	20,9%	13,5%	5,4%	2,0%	3,2	3,0
I have encountered issues and problems with AI chatbots in the past	3,4%	10,1%	6,1%	14,9%	14,2%	25,0%	26,3%	5,1	6,0
Total	9,8%	14,7%	14,2%	17,4%	17,9%	12,5%	13,5%	4,1	4,0

Figure 14 Past Experiences

First, considering how fairly common the use of AI chatbots in customer service is nowadays, a fairly significant portion of participants appears to use them. The average 3,1 and median 3,0 implicate translate to “Partially Disagree”. Only 20,9% of the answer were above the neutral answer. This finding indicates that AI Chatbots are not regularly used by participants – but might be used occasionally or depending on the situation. Strongly disagree and disagree answers were gained in 39,2% of the total answers, indicating the unlikeness of this group using the AI Chatbots that often.

Considering the knowledge of participants, the majority of the participants claimed to partially agree to have a basic understanding of how AI and AI Chatbots function. The average and the mean to this question was 5,0 (Partially Agree). Only 7,5% strongly disagreed or disagreed, and most of the answers fell more into agreeing over the neutral option. 23% of the people strongly agreed to have a basic understanding. To conclude, in this sample size, most of the participants had a fairly good knowledge of the basics of AI.

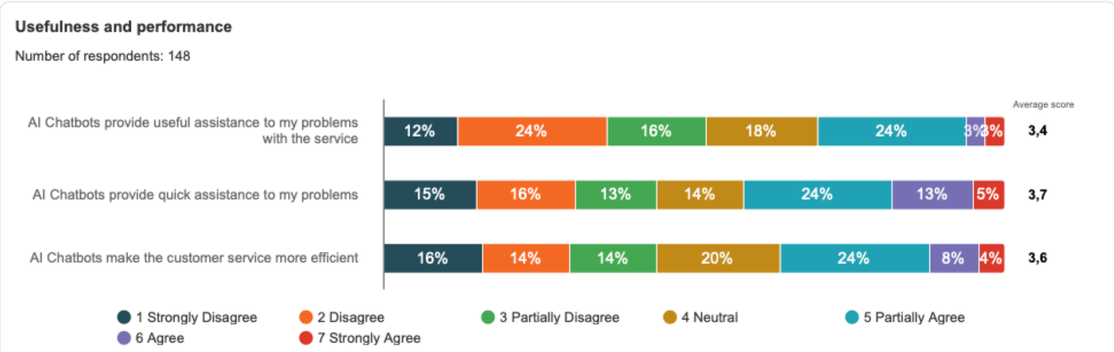
The last two questions in this category aimed to reflect the past comfortability and enjoyment of using the AI Chatbots. The last question, on the other hand, focused on disruption, problems and issues with the AI Chatbots. The question on how comfortable and enjoyable the user experience has been was fairly divided, averaging 3,2 with the mean of 3.0 scoring partially disagree. 18% Strongly disagreed, 20% disagreed, and 20% partially disagreed. This distribution is fairly similar to the first question in this category on how regularly you use AI Chatbots.

The last question clearly illustrates that the majority of the participants had experienced problems and issues with AI Chatbots in the past. Averaging 5,1 with a median of 6, the scores agree. Only 20% of the answers disagreed, claiming that there have not been issues or problems. Yet, the most over 51% of the answers strongly agreed or agreed to have problems. This can be seen as a clear result, as most people have experienced problems in the past.

To conclude these results, this paper can summarise that AI chatbots are used somewhat regularly, but the majority of the participants use the solution more or less occasionally. The knowledge of the basics of AI and Chatbots is fairly good in this sample, and participants partially agreed. Not that many participants enjoyed or were comfortable using these solutions, which could reflect the regularity of the usage. Most of the participants have encountered issues and problems with AI Chatbots.

4.2 Perceived usefulness and performance

This section will introduce the results related to the category, aiming to evaluate the perceived usefulness and performance presented in Figure 14.



	1 Strongly Disagree	2 Disagree	3 Partially Disagree	4 Neutral	5 Partially Agree	6 Agree	7 Strongly Agree	Average	Median
AI Chatbots provide useful assistance to my problems with the service	11,5%	24,3%	16,2%	17,6%	24,3%	3,4%	2,7%	3,4	3,0
AI Chatbots provide quick assistance to my problems	15,6%	16,2%	12,8%	14,2%	23,7%	12,8%	4,7%	3,7	4,0
AI Chatbots make the customer service more efficient	16,2%	14,2%	14,2%	20,3%	23,6%	8,1%	3,4%	3,6	4,0
Total	**14,4%**	**18,2%**	**14,4%**	**17,4%**	**23,9%**	**8,1%**	**3,6%**	**3,6**	**4,0**

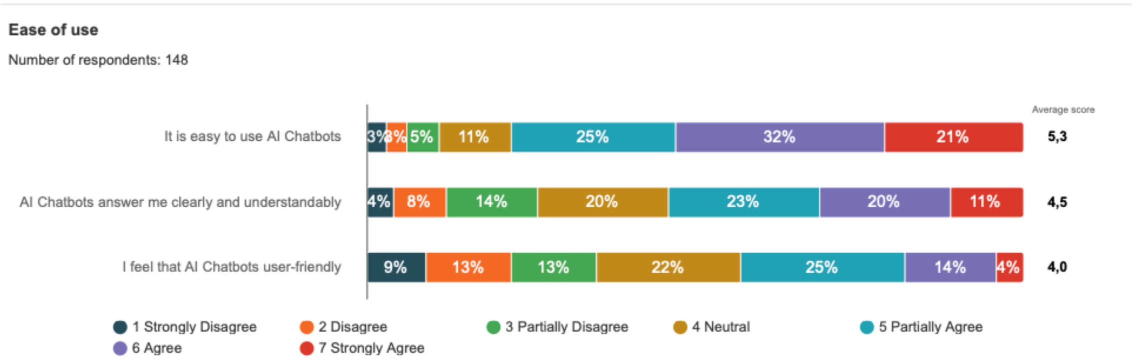
Figure 15 Perceived usefulness and performance

In total, this paper can note that most of the answers were quite neutral. The perception of whether AI Chatbots provide useful assistance with the service shared opinions scoring 3. Partially disagree with an average of 3,4 and a median of 3,0. Only 6,1% strongly agreed or agreed, and most were more sceptical about the usefulness of the AI Chatbots. More positive responses were received regarding the speed and quickness of the assistance received from the AI Chatbots, averaging 3,7 median of 4,0. This can be seen as a neutral response. The same kind of distribution and metrics received the question about AI Chatbots' ability to make customer service more efficient.

To conclude, the topic shared opinions between participants, and the results could mostly be described as neutral.

4.3 Ease of use

As part of the relevant metrics in modelling the customer experience and technology acceptance model, there is a component of ease of use for the offered solution. The questions aimed to gain insights into the participants' perceptions of ease. The results are presented in Figure 15.



	1 Strongly Disagree	2 Disagree	3 Partially Disagree	4 Neutral	5 Partially Agree	6 Agree	7 Strongly Agree	Average	Median
It is easy to use AI Chatbots	3,4%	2,7%	4,7%	10,8%	25,7%	31,8%	20,9%	5,3	6,0
AI Chatbots answer me clearly and understandably	4,0%	8,1%	14,2%	19,6%	23,0%	19,6%	11,5%	4,5	5,0
I feel that AI Chatbots user-friendly	8,8%	13,5%	13,5%	21,6%	25,0%	13,5%	4,1%	4,0	4,0
Total	**5,4%**	**8,1%**	**10,8%**	**17,3%**	**24,6%**	**21,6%**	**12,2%**	**4,6**	**5,0**

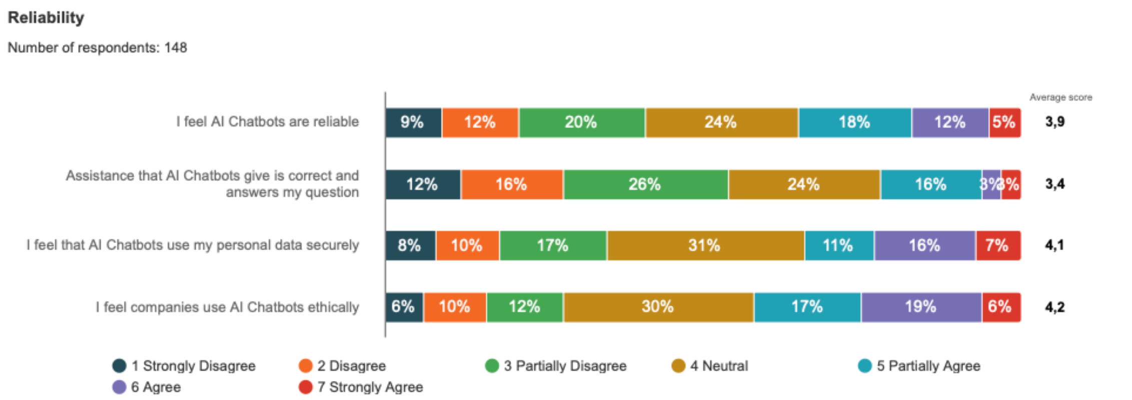
Figure 16 Ease of use

The most of the participants regarded the use of AI Chatbots to be easy. The 78,4% of surveys answers considered it to be at least partially easy – above the neutral answer option. The 52,7% strongly agreed or agreed. Only 6,1% strongly disagreed or disagreed. The question averaged 5,3 with a median of 6,0 (Agreed). This paper concludes that, in general, AI Chatbots are considered to be easy to use.

Considering the clarity and understandability of the answers received from AI Chatbots, the general opinion falls under the category of partially agreeing, with an average of 4,5 and a median of 5,0. This indicates pretty neutral answers regarding the communication of AI chatbots. Close to this, the sum-up question on the feeling of user-friendliness of AI Chatbots received an average and mean of 4,0 neutral.

4.4 Reliability

Considering the reliability category, the aim was to map the responses' perceptions of the accuracy of information, issues with privacy, and the reliability of using personal data. The answers are presented in Figure 16.



	1 Strongly Disagree	2 Disagree	3 Partially Disagree	4 Neutral	5 Partially Agree	6 Agree	7 Strongly Agree	Average	Median
I feel AI Chatbots are reliable	8,8%	12,2%	20,3%	23,6%	18,2%	11,5%	5,4%	3,9	4,0
Assistance that AI Chatbots give is correct and answers my question	11,5%	15,5%	26,4%	24,3%	16,2%	2,7%	3,4%	3,4	3,0
I feel that AI Chatbots use my personal data securely	8,1%	9,5%	16,9%	31,1%	10,8%	16,2%	7,4%	4,1	4,0
I feel companies use AI Chatbots ethically	6,1%	10,1%	12,2%	29,7%	16,9%	18,9%	6,1%	4,2	4,0
Total	**8,6%**	**11,8%**	**19,0%**	**27,2%**	**15,5%**	**12,3%**	**5,6%**	**3,9**	**4,0**

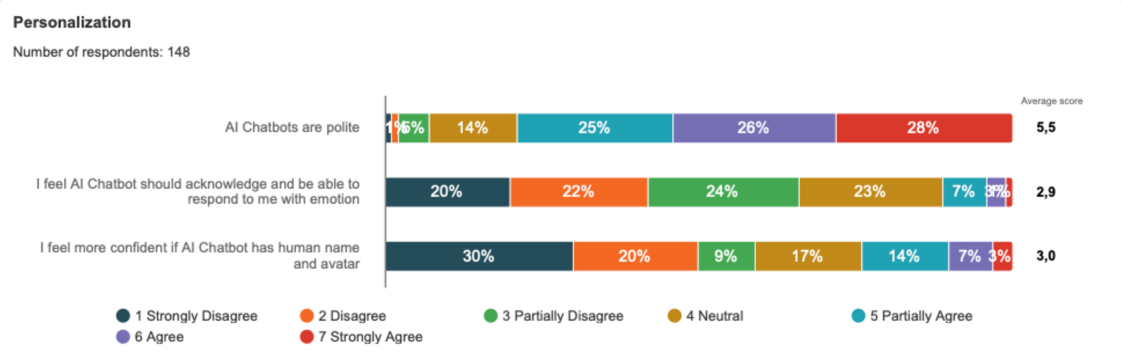
Figure 17 Reliability

The first question measures, in general, the feeling of reliability, and it scored an average of 3,9 and a mean of 4,0, giving the overall impression of a neutral answer. The question did not raise that many strong opinions in either way. Results were more critical when it came to the accuracy and correctness of the answers. The average was 3,4, and the mean was 3,0, partially disagreeing. 27% of the answers strongly disagreed or disagreed with the statement, but only 6,1% strongly agreed or agreed. Overall, the most popular answer was partially disagreeing 26,4% or neutral 24,3%.

The other two questions were more related to data security and the ethics of companies using AI chatbots. The first measured the feeling of securely using personal data, and it received neutral answers averaging 4,1 with a mean of 4,0. Answers were distributed more or less equally, focusing mostly on the neutral options. The same sort of neutral answers received the feeling of companies using the AI Chatbots to do it ethically, averaging 4,2 with a mean of 4,0. 41,9% considered the ethical behaviour to be at least partially agreeable when at disagreeing parties, whereas only 28,4% did. To conclude, the perception towards ethical behaviour and use are neutral and somewhat trusted.

4.5 Personalisation

The category personalisation aimed to gain deeper insight into the effects of how AI Chatbots communicate and how human-like these behave. These questions will try to give more insight into the research model. The results of this section are presented in Figure 17.



	1 Strongly Disagree	2 Disagree	3 Partially Disagree	4 Neutral	5 Partially Agree	6 Agree	7 Strongly Agree	Average	Median
AI Chatbots are polite	0,7%	1,4%	4,7%	14,2%	25,0%	26,3%	27,7%	5,5	6,0
I feel AI Chatbot should acknowledge and be able to respond to me with emotion	19,6%	21,6%	24,3%	23,0%	6,8%	3,4%	1,3%	2,9	3,0
I feel more confident if AI Chatbot has human name and avatar	29,7%	19,6%	9,5%	16,9%	13,5%	7,4%	3,4%	3,0	3,0
Total	**16,7%**	**14,2%**	**12,8%**	**18,0%**	**15,1%**	**12,4%**	**10,8%**	**3,8**	**4,0**

Figure 18 Personalisation

The general opinion on the politeness of the AI chatbots averaged 5,5, with a mean of 6,0 agree. This indicates that, in total, participants found the AI Chatbots to use language that can be seen as polite. 79% of the participants agreed with the neutral answer option, and only 6,8% disagreed below the neutral. To conclude, AI chatbots are widely considered to be polite overall.

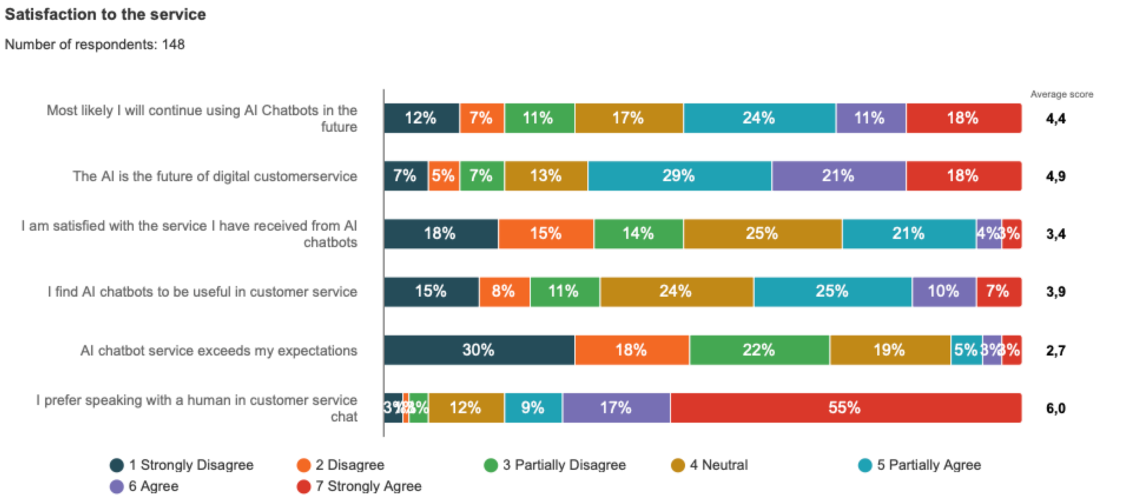
The last two questions in this category mapped the anthropomorphism of the AI Chatbots. How emotionally connecting to the customer and having human-like attributes affect the experience. Overall, these results were surprising considering the theoretical background. The theory suggested that the ability to react to emotions would increase the satisfaction of using the service, but in this research, the findings were the opposite (Adam et al., 2021). The feeling of Ai Chatbot should acknowledge and be able to respond with emotion averaged 2,9 with a mean of 3,0 partially disagreeing. 65,5% of the participants responded below the neural option and only 11,5% above the neutral. This indicates that the ability to respond and react to emotions is not considered to be a factor contributing to the service.

A similar result yielded the question of whether the person feels more confident if the AI Chatbot has a human name and avatar. This way of appearing more human-like did not get much positive feedback. The average and mean were 3,0, partially disagreeing. 58,8% answered the below-neutral option, and only 24,3% agreed with the above-neutral option.

To conclude, it appears that participants did not expect, require or value the more human-like design of the AI Chatbots.

4.6 Satisfaction with the service

This category focused on the aspects of satisfaction with the AI Chatbots. Satisfaction is based on the future perspective and the perception of usefulness and likelihood of the use. The total results are presented in Figure 18.



	1 Strongly Disagree	2 Disagree	3 Partially Disagree	4 Neutral	5 Partially Agree	6 Agree	7 Strongly Agree	Average	Median
Most likely I will continue using AI Chatbots in the future	12,2%	6,7%	10,8%	17,6%	24,3%	10,8%	17,6%	4,4	5,0
The AI is the future of digital customerservice	6,8%	5,4%	7,4%	12,8%	29,1%	20,9%	17,6%	4,9	5,0
I am satisfied with the service I have received from AI chatbots	17,6%	15,5%	14,2%	25,0%	20,9%	4,1%	2,7%	3,4	4,0
I find AI chatbots to be useful in customer service	14,9%	8,1%	10,8%	24,3%	25,0%	9,5%	7,4%	3,9	4,0
AI chatbot service exceeds my expectations	29,7%	17,6%	22,3%	18,9%	5,4%	3,4%	2,7%	2,7	3,0
I prefer speaking with a human in customer service chat	2,7%	0,7%	3,4%	11,5%	9,4%	16,9%	55,4%	6,0	7,0
Total	**14,0%**	**9,0%**	**11,5%**	**18,4%**	**19,0%**	**10,9%**	**17,2%**	**4,2**	**4,0**

Figure 19 Satisfaction with the service

The first question asked if the participants were most likely to continue using AI chatbots in the future. The average was 4,4, and the mean 5,0 partially agree. 52,7% of answers were above the neutral option, while 29,7% were below the neutral. This question itself has problem at its own since the likeliness of the future use cannot be factored straight into satisfaction of use – but rather some may consider it as a necessity that there are no other options than to rely on AI Chatbots.

The second measure was more about the perception of the concept for the future. Will AI be the future of digital customer service? The questions received an average of 4,9 and a mean of 5,0, indicating that most people partially agree on that future development. The most 67,6% of the answers were above the average. Only 19,6% disagreed with the statement below the neutral answer option. Expectations seem to be more partially agreeing that, in some instances, AI will play a part in the future of customer service on digital platforms.

Satisfaction with the service received from AI chatbots received a neutral score with an average of 3,4 and a median of 4,0. 47,3% answered below the neutral option, yet only 2,7% strongly agreed that the service was satisfactory. Similar results came from the general question of how useful participants find AI chatbots in customer service. The result was an average score of 3,9 with a median of 4,0. A neutral result, but it got stronger opinions opposing 14,9% strongly disagreed, but only 7,4% strongly agreed on the usefulness.

Considering the usefulness and exceeding the expected service quality, it is rare that people perceive the service to be better than expected. Most of the people, 29,7%, answered strongly disagreeing with the question if AI Chatbot service exceeds expectations. The average was 2,7, and the mean was 3,0, falling towards partially disagreeing in total. Only 2,7% of the participants felt that AI chatbots exceeded their expectations. Even only 11,5% of the participants answered above the neutral option.
The strongest result in this research came from the question of whether the participant would rather prefer to speak with a human in customer service. 55,4% of the answers strongly agreed with this that they would rather talk with a human in customer service. The average was 6,0 agreeing, and the median was 7,0 strongly agreeing. In total, 81,7% of the answers agreed more than the neutral answer option. Only 6,8% of the answers here below the 11,5% neutral option.

4.7 Research Model

Based on the previous literature, this paper suggests the following model in Figure 19, to evaluate "*What factors are perceived to influence the adoption of new AI technology, create value, and affect the service quality in digital platforms*".

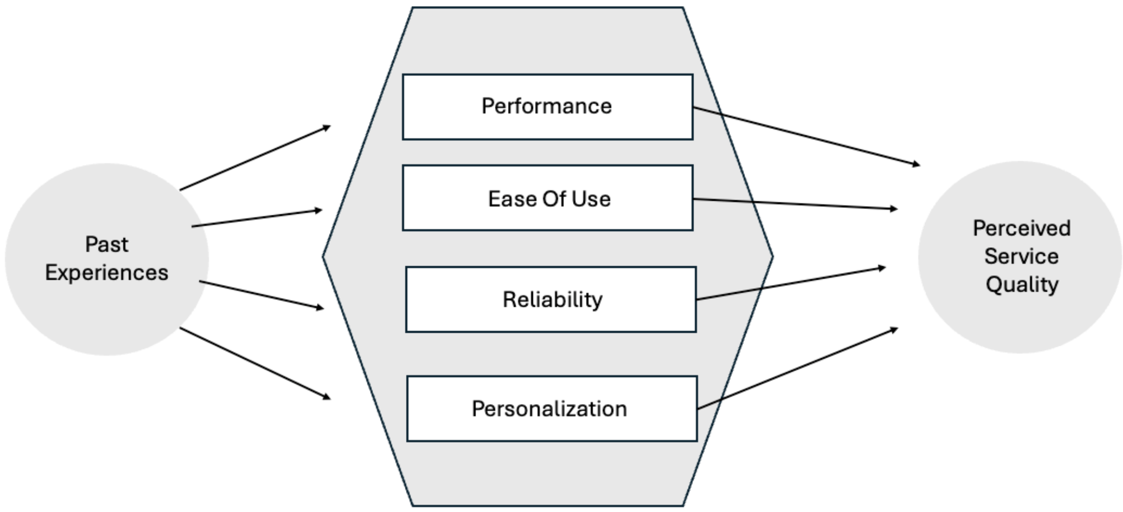


Figure 20 Research Model for the hypothesis

Considering the previously introduced models TRA and TBH, previous experiences create expectations and beliefs toward similar future behaviour. The evaluation of potential future behaviour, attitudes, subjective norms, and perceived behavioural control are some of the attributes contributing to the actual behaviour. (Ajzen, 1991). Also, social peer pressure and cultural context have been shown to affect the behavioural belief system and expectations. (Parasuraman et al., 2005, 2005)

To conclude, this paper makes the hypothesis that past experiences, needs, beliefs, and attitudes form expectations for service and its (Parasuraman et al., 2005; Santos, 2003). The actual service quality is experienced through how well these factors are met. Different customers can experience the same service and different quality (Parasuraman et al., 2005; Santos, 2003)

This study focuses on better understanding customers' perceptions of factors that contribute to creating quality service with AI chatbots. The research model presented in Figure 9 is based on the concepts introduced in literature and previous studies. The model is constructed to specify theory in the context of AI chatbots in creating quality service. The focus is on customers' perception, not necessarily on the factual contributors, but on factors that customers consider to give value to the service. The model can be seen as a lens that reflects the perceptions through the factors creating the experience service quality.

4.8 Valuating the research model

For the purpose of the study, the relevance is in correlations of the remaining categories towards the perceived service quality. The personalisation and past experiences categories had to be discarded due to inconsistent answers and poor Cronbach alpha's. This indicates heavily the factor's relevance to the participant's perception factors importance to total satisfaction on using the AI Chatbots.

Evaluating the total scored results, the correlation matrix is used to understand the dependencies between categories. Due to the nature of 7-step Likert chart and ordinary values, the most accurate metric is Spearman's rho. The correlation matrix is presented in Table 6.

Correlations

			Total_satisfaction	Total_reliable	Total_performance	Total_easeofuse
Spearman's rho	Total_satisfaction	Correlation Coefficient	1.000	.514**	.823**	.609**
		Sig. (2-tailed)	.	<,001	<,001	<,001
		N	148	148	148	148
	Total_reliable	Correlation Coefficient	.514**	1.000	.435**	.541**
		Sig. (2-tailed)	<,001	.	<,001	<,001
		N	148	148	148	148
	Total_performance	Correlation Coefficient	.823**	.435**	1.000	.569**
		Sig. (2-tailed)	<,001	<,001	.	<,001
		N	148	148	148	148
	Total_easeofuse	Correlation Coefficient	.609**	.541**	.569**	1.000
		Sig. (2-tailed)	<,001	<,001	<,001	.
		N	148	148	148	148

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6 The correlation matrix of factors

The results from Table 6 are the most relevant compared to the total satisfaction. The most significant finding is that the performance factor is the most correlated (0.823) with satisfaction and perceived overall service quality. Also, the ease of use can be seen to correlate with satisfaction with a value of 0.609. Also, reliability is correlated with 0.514, but not as strongly as the other factors. In conclusion, all the attributes reflect correlation and could be seen to affect the perceived service quality of AI Chatbots. The performance rises to be the most correlated. All the correlations are somewhat correlated with each other, which might indicate that all the attributes have an effect. There are no negative correlations that could be indices otherwise.

Based on the correlation analysis, the model can be simplified into a verified model presented in Figure 20.

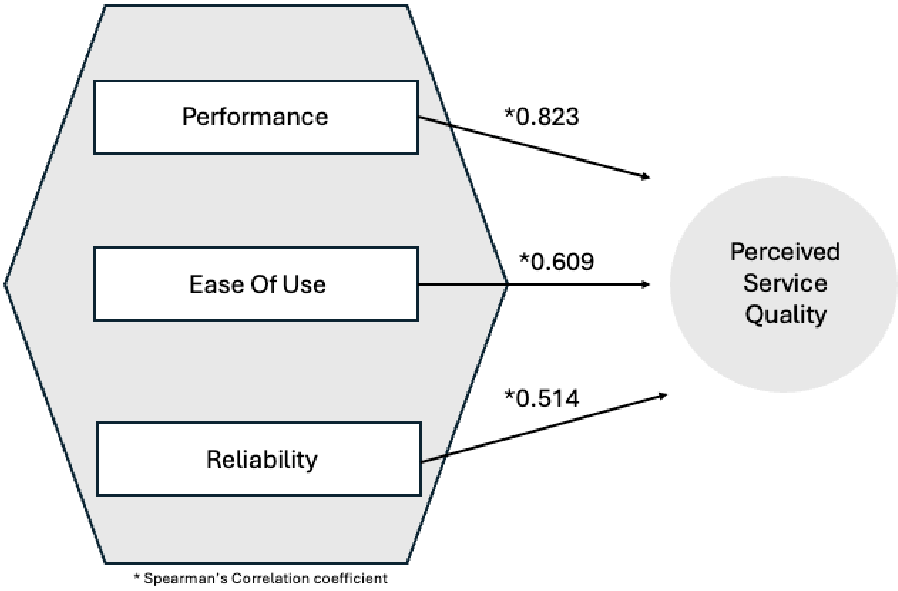


Figure 21 Simplified research model with correlation coefficients

The correlation alone does not yet establish a causal relation between the factors and satisfaction. Linear regression analysis must be used to determine if there is a causal relation.

In a linear regression with performance as a predictor and satisfaction as a dependent variable, the R-squired is 0.707. This can be interpreted as 70,7% of the variance in satisfaction scores being explained by performance variations. This result indicated a strong relation with these factors (Table 7).

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.841[a]	0.707	0.705	4.11883

a. Predictors: (Constant), Total_performance
b. Dependent Variable: Total_satisfaction

Table 7 Performance and Satisfaction R-squired

Table 8 shows that ease of use variation explains 41.4% of the satisfaction variability. This is not as strong as with performance, but the relation is still significant.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.643[a]	0.414	0.410	5.82865

a. Predictors: (Constant), Total_easeofuse
b. Dependent Variable: Total_satisfaction

Table 8 Ease of use and satisfaction R-squired

Reliability shows that 35.3% of the variability in satisfaction is accounted for by reliability (Table 9). This result is also meaningful and can be interpreted as indicating a significant portion of causality in explaining the phenomenon.

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.594[a]	0.353	0.348	6.12298

a. Predictors: (Constant), Total_reliable
b. Dependent Variable: Total_satisfaction

Table 9 Reliability and satisfaction R-squired

To validate the research model, the whole model was evaluated with linear regression, with satisfaction as the dependent variable. The model factors explain 75,9% of the variation in Satisfaction (Table 10).

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.871[a]	0.759	0.754	3.75877

a. Predictors: (Constant), Total_reliable, Total_performance, Total_easeofuse

b. Dependent Variable: Total_satisfaction

Table 10 All factors and satisfaction R-squired

To summarize the findings, this thesis concludes that performance, ease of use and reliability are correlated with satisfaction, and there is a strong causality with the factors towards satisfaction. Thus, the model can be validated to be accurate in explaining satisfaction and quality service. Now the model, Figure X, establishes that there is a strong causal relationship. This model is in line with the factors in TAM and UTAUT models, where the performance and ease of use have been seen to affect the actual use of the service (Davis et al., 1989; Venkatesh et al., 2003). Reliability and previous factors combined has seen to be correlated and affect with the service quality as part of SERVIQUAL model (Parasuraman et al., 2005)

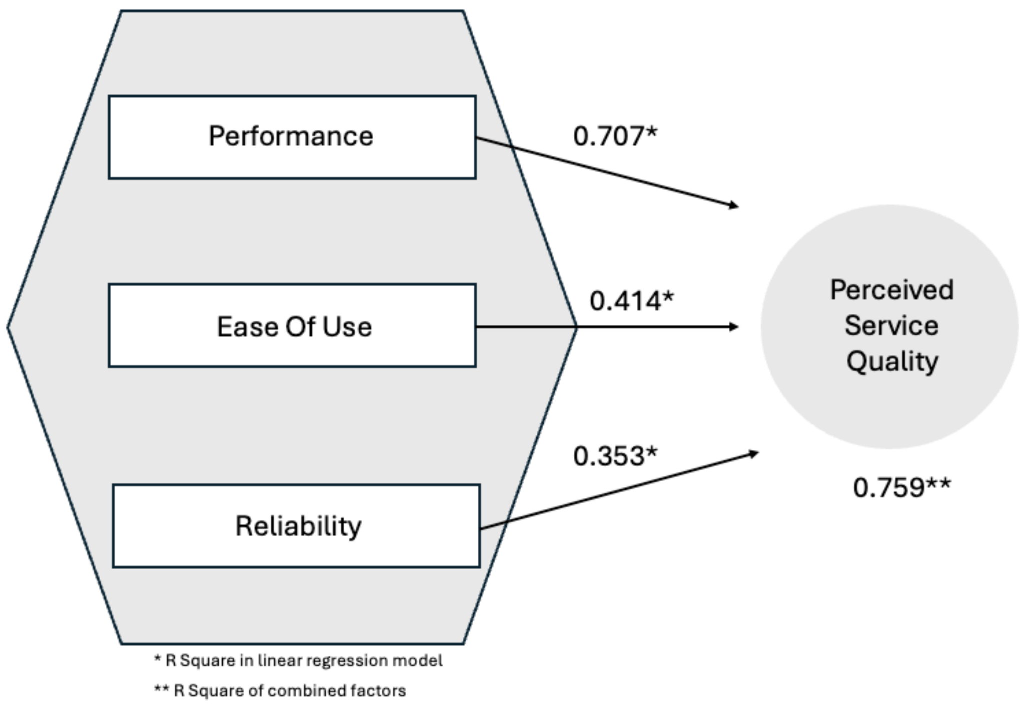


Figure 22 Linear regression model explains causality in research model

Next, this thesis analyses individual questions and then the open-ended questions to gain deeper insights to complement the model.

4.9 Further statistical analysis of the results

Based on the survey question this research aimed to gain more insights on the perception of different factors effect on customer satisfaction on service provided by AI Chatbots. The study implemented a few statistical methods to find significant results.

To mention, it is important to determine if there are differences in satisfaction with AI chatbots between different demographics. For example, the age was crossable with the question "*I am satisfied with the service I have received from AI chatbots*", Table 11.

Age * I am satisfied with the service I have received from AI chatbots										
			I am satisfied with the service I have received from AI chatbots							Total
			1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree	
Age	18-24	Count	7	6	4	10	6	2	4	39
		Expected Count	6.9	6.1	5.5	9.8	8.2	1.6	1.1	39.0
		% within Age	17.9%	15.4%	10.3%	25.6%	15.4%	5.1%	10.3%	100.0%
	25-34	Count	5	6	9	10	15	1	0	46
		Expected Count	8.1	7.1	6.5	11.5	9.6	1.9	1.2	46.0
		% within Age	10.9%	13.0%	19.6%	21.7%	32.6%	2.2%	0.0%	100.0%
	35-44	Count	4	2	5	6	5	2	0	24
		Expected Count	4.2	3.7	3.4	6.0	5.0	1.0	0.6	24.0
		% within Age	16.7%	8.3%	20.8%	25.0%	20.8%	8.3%	0.0%	100.0%
	45-54	Count	8	7	1	5	4	0	0	25
		Expected Count	4.4	3.9	3.5	6.3	5.2	1.0	0.7	25.0
		% within Age	32.0%	28.0%	4.0%	20.0%	16.0%	0.0%	0.0%	100.0%
	55-64	Count	1	2	2	5	1	1	0	12
		Expected Count	2.1	1.9	1.7	3.0	2.5	0.5	0.3	12.0
		% within Age	8.3%	16.7%	16.7%	41.7%	8.3%	8.3%	0.0%	100.0%
	Over 64	Count	1	0	0	1	0	0	0	2
		Expected Count	0.4	0.3	0.3	0.5	0.4	0.1	0.1	2.0
		% within Age	50.0%	0.0%	0.0%	50.0%	0.0%	0.0%	0.0%	100.0%
Total		Count	26	23	21	37	31	6	4	148
		Expected Count	26.0	23.0	21.0	37.0	31.0	6.0	4.0	148.0
		% within Age	17.6%	15.5%	14.2%	25.0%	20.9%	4.1%	2.7%	100.0%

Table 11 Cross table: Age & Satisfaction sub question

Then, methods such as ANOVA and Pearson's Chi-Squire, value 0,212, where used to determine if there could be found age groups where the results deviated from the greater population.

The only statistic that indicated statistical significance between demographics and factors was age towards ease of use. ANOVA indicated 0,049 that there could be found significant differences among the age groups for the dependent variables but evaluating this more with Post Hoc test Tukey HSD analysis – it's not possible to find individual comparisons that are significant.

This study did not find any significant results between any demographics with relevant questions. Reasons for this could be the sample size of 148 participants, homogeneity of variances indicating the same kind of results or the spread of means. The total perception towards the AI Chatbots was critical.

Even though this research did not yield meaningful results in this area, the correlation analysis in the previous chapter and the qualitative answers discussed in the next chapter could give more insights into the topic.

4.10 Previous satisfaction and dissatisfaction factors

The question Q1 was, *"Why have you previously been satisfied or dissatisfied with customer service AI Chatbots?".* The question, in total, gained 95 answers from 148 participants. The below, Table 8, is presented the number of mentions in the answers about the specified topic.

Q1	Why have you previously been satisfied or dissatisfied with customer service AI chatbots?
n	95
Number of mentions	
62	Inability to answer more complex questions & limited ability to understand
14	Lacking intergration and implementation
17	Generic and repetitive answers
13	Quick answers for basic questions

Table 11 Q1 Categorized comments.

As Table 8 demonstrates, the most commented issue was the inability to answer more complex questions or the ability to understand the questions properly. It gained 62 mentions in comments. Many comments sympathised with the problems where the AI Chatbot is either incapable of providing satisfactory answers or does not even understand the question. This has led to some frustration in some participants and made them contact human customer service instead.

The same kind of issue is when the answers are perceived as generic or even repetitive – not really solving the problem the customer is experiencing. These answers are experienced as irrelevant and does not address the actual user's problem. Many of these felt that AI Chatbots just gives hyperlinks to webpages where the asked topic is discussed but yet still won't solve the problem. This was also mentioned in 17 comments.

The execution of the AI Chatbot, which lacked integration and implementation, was also mentioned in 14 comments. The main issue was on the user-experience side, where the pop-up chatbots were seen as disturbing, the chatbots were seen as annoying, or the structure of the chatbot had been implemented poorly so that it was irrelevant to the customer.

In more positive perspective, the most good experiences related to the availability and quickness to the AI Chatbots. Topic was discussed in 13 comments, where the chatbots got positive feedback on being able to answer simple basic questions or at least to guide user to the information they were seeking. Also the ability to not wait on line for human customer service was discussed.

Some other answers emphasised that AI chatbots may be better at customer service in the future. The comparison was made with OpenAI´s ChatGPT, and most of the customer service chatbots are not that good. This could indicate a more advanced level of AI based on levels of intelligence could be beneficial for satisfaction (Hollebeek et al., 2021; Huang & Rust, 2018).

Based on these comments, the overall sentiment leans towards dissatisfaction; however, recognising the potential benefits of quickness and the ability to guide towards the right way could be addressed by the companies.

Limitations are regarded mostly as the ability to handle complex questions or understand the issue itself. Even though the comments were mostly negative and related to performance issues or irrelevant answers, the executive's summary might be based on these answers, and if these answers could be addressed, the future of customer service chatbots could look brighter. Or at least these comments would be noted in future development since the topics were discussed in comments concisely and often.

4.11 Benefits of AI Chatbot versus Human-driven Customer Service

The second question Q2 was, "*In some situations, is AI chatbot a better way to contact for issue resolution than conversing with a human? In what situation?". The q*uestion received 79 written answers from 148 participants. To better evaluate the content to get excess information, these answers are grouped by the issue that arose from the comment. The Table 9 is presented the distribution of answers. The simple irrelevant responses are left out, such as no comment or simple yes or no answers.

Q2	In some situations, is AI chatbot a better way to contact for issue resolution than conversing with a human? In what situation?
n	79
Number of mentions	
16	Effective in simple questions
19	As a preliminary step easy access service, not in complex issues
6	Quick answers and fast
8	24/7 availability and efficiency
13	Frustration and performance problems

Table 12 Q2 Categorized comments

The main advantages that arose from the comments were related to the speed of AI Chatbots and their ability to answer simple questions, mainly related to finding resources, for example, from company websites. These comments sympathised with the role of AI Chatbots as a preliminary step to get help and support. If the capability to answer questions was not enough, participants felt the need to wait for human assistance. Yet, it is also important to notice that many participants, 13 in this sample, feel the same issues with performance as in Q1. This can even lead to annoyance and frustration.

These results reflect the performance, reliability, and ease of use standards confirmed in TAUT and SERVIQUAL models. (Davis et al., 1989; Parasuraman et al., 2005; Venkatesh et al., 2003)

To conclude the answers, there were a lot of similarities to the Q1 answers. The participants saw some benefits, mostly related to positioning the AI Chatbot as a first line of contact that can answer simple questions fast. There were also opposing opinions that considered this just to be annoying. The majority of participants still considered the AI Chatbot in this role to be efficient, even though the performance and accuracy of the answers remain the major issues.

4.12 Preferences for future development

The last question, Q3 in the survey, asked an open-ended question, *"I would be more satisfied with customer service AI chatbots if…".* It got 58 answers that were also grouped into categories of the most occurring answers for easier analysis. Other answers did not provide insight. The grouped answers are in Table 10.

Q3	I would be more satisfied with customer service AI chatbots if...
n	58
Number of mentions	
36	Improvements of understanding and responce quality
8	Technically more advanced into language models
4	Quick escalation to humanled service

Table 13 Q3 Categorized comments

In the ideas of future development, the same theme continues as in the Q1 and Q2 questions. Most of the ideas related to improving AI chatbots for the future were related to performance improvements. Some commented more on technical advancements towards more OpenAI ChatGPT language models. Participants were united on the future goals of improving their understanding, accuracy, and ability to answer the questions efficiently. The tone appeared accepting towards future development even though some answers did not like the concept of AI Chatbots at all or just wanted to talk with a human.

5 DISCUSSION AND CONCLUSIONS

This research aimed to answer the main research problem: "*What factors are perceived to influence the adoption of new AI technology into use, create satisfaction, and affect the service quality in digital platforms?*".

Based on the results, this research concludes that performance, ease of use, and reliability factors in AI Chatbots are perceived to affect service quality and satisfaction. There is a strong correlation and a strong causal relation between factors and perception of satisfactory quality service. The model factors explain 75,9% of the variation in satisfaction based on the linear regression model. Open-ended questions provided qualitative insights into customer perceptions, indicating and establishing the same kind of results as the quantitative part did. Performance is seen as the most crucial factor for customers to perceive quality service with AI chatbots in digital service, and it received a vast amount of mentions in open-ended questions.

This research confirmed and established the results that are in line with TAM, UTAUT, and SERQUAL models when it comes to implementing theories in the context of AI Chatbots in customer service. Stating that these factors affect the customers to adopt the solution into actual use (Davis et al., 1989; Parasuraman, 2000; Venkatesh et al., 2003). The study found performance to be the most correlated and causally effective factor towards customer satisfaction, but ease of use and reliability also affected the result.

Considering the ECT model and the expectations of customers based on this thesis has confirmed that regarding AI chatbots, the performance, ease of use, and reliability factors need to be addressed so that customers can confirm the usefulness and continue using the solution in the future (Bhattacherjee, 2001). The TAM and UTAUT models recognised these factors to be key to keeping customers using the solutions(Davis et al., 1989; Venkatesh et al., 2003), and the SERVIQUAL model recognised these attributes connected to perceived quality of service (Parasuraman et al., 2005; Santos, 2003). This study found the similar results.

The overall perception of satisfaction in the current state of AI chatbots was not good since the majority of people did not find the solution effective. Most people prefer human customer service agents. Most issues arose from the perceived performance and accuracy of the output the chatbots provide. There was no statistical difference between the different demographics concerning the perceptions on the topic.

In future research, it would be useful to explore the application of these theories in the context of AI chatbots. This study did not investigate the impact of anthropomorphism or past experiences on customer satisfaction, which is an area that could benefit from further research. Connecting the effect of more human-like AI chatbots to the frameworks of customer satisfaction. Additionally, it is important to validate the effects of personalisation as one of the contributing factors.

Future research would be interesting from the perspective of advanced language models, when the technology advances, and how these would change people's perceptions of satisfaction.

5.1 Managerial implications

From a business perspective, the overall attitudes and perceptions towards AI Chatbots in customer service are negative. Most of the issues people are unsatisfied with are related to performance issues. The chatbot does not appear to understand and answer correctly. Also, the generic answers are seen as frustrating or annoying. The majority of people prefer talking with the customer service agent. From the business perspective it could be beneficial to consider business case with AI Chatbots costs, costs savings and effect on customer satisfaction.

For the future implementation and development, the ability to address problems with performance could help the overall satisfaction of AI Chatbots. Currently, the main benefit appears to be the ability to get quick answers and guidance from the chatbots.

The vast majority of people still prefer humans in chat or at least the option to talk with a human. The advantage is the ability to resolve complex problems and get all problems addressed at the same time. From the business case perspective, the recommendation could remain in a hybrid model, where AI Chatbots are a preliminary step towards customer service agents. Recommendations for the KPIs to measure when implementing the AI Chatbot into customer service would be the %-of resolved issues, accuracy of the answers received and metrics to determine how many resources are saved when AI Chatbots handle part of the incoming service requests.

Also, the ability to measure customer satisfaction with the AI Chatbots is important since, it appears these could be seen as frustrating and annoying – especially if the problems are not solved or the chatbots wastes customers time.
One consideration would be the availability of customer service since AI Chatbots can operate 24/7 and resolve at least some problems when human customer service is not open.

Adopting AI Chatbots in customer service should be considered from different business case perspectives to evaluate the suitability of the company’s strategic goals and brand strategy.

6 REFERENCES

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, *31*(2), 427–445. https://doi.org/10.1007/s12525-020-00414-7

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Bhattacherjee, A. (2001). Understanding Information Systems Continuance: An Expectation-Confirmation Model. *MIS Quarterly*, *25*(3), 351. https://doi.org/10.2307/3250921

Camilleri, M. A. (2024). Factors affecting performance expectancy and intentions to use ChatGPT: Using SmartPLS to advance an information technology acceptance framework. *Technological Forecasting and Social Change*, *201*, 123247. https://doi.org/10.1016/j.techfore.2024.123247

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, *35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

Dwivedi, Y. K., Rana, N. P., Jeyaraj, A., Clement, M., & Williams, M. D. (2019). Re-examining the Unified Theory of Acceptance and Use of Technology (UTAUT): Towards a Revised Theoretical Model. *Information Systems Frontiers*, *21*(3), 719–734. https://doi.org/10.1007/s10796-017-9774-y

Frank, D.-A., Jacobsen, L. F., Søndergaard, H. A., & Otterbring, T. (2023). In companies we trust: Consumer adoption of artificial intelligence services and the role of trust in companies and AI autonomy. *Information Technology & People*, *36*(8), 155–173. https://doi.org/10.1108/ITP-09-2022-0721

Gerlich, M. (2023). Perceptions and Acceptance of Artificial Intelligence: A Multi-Dimensional Study. *Social Sciences*, *12*(9), 502. https://doi.org/10.3390/socsci12090502

Hollebeek, L. D., Sprott, D. E., & Brady, M. K. (2021). Rise of the Machines? Customer Engagement in Automated Service Interactions. *Journal of Service Research*, *24*(1), 3–8. https://doi.org/10.1177/1094670520975110

Huang, M.-H., & Rust, R. T. (2018). Artificial Intelligence in Service. *Journal of Service Research*, *21*(2), 155–172. https://doi.org/10.1177/1094670517752459

Khan, A. (2021). *Public Opinion Toward Artificial Intelligence*. https://doi.org/10.31219/osf.io/284sm

Klein, K., & Martinez, L. F. (2023). The impact of anthropomorphism on customer satisfaction in chatbot commerce: An experimental study in the food sector. *Electronic Commerce Research*, *23*(4), 2789–2825. https://doi.org/10.1007/s10660-022-09562-8

Lee, Y., Kozar, K. A., & Larsen, K. R. T. (2003). The Technology Acceptance Model: Past, Present, and Future. *Communications of the Association for Information Systems*, *12*. https://doi.org/10.17705/1CAIS.01250

Nass, C., Steuer, J., & Siminoff, E. (1994). *Computer are social actors*. 204. https://doi.org/10.1145/259963.260288

Parasuraman, A. (2000). Technology Readiness Index (Tri): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies. *Journal of Service Research*, *2*(4), 307–320. https://doi.org/10.1177/109467050024001

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. *Journal of Marketing*, *49*(4), 41–50. https://doi.org/10.1177/002224298504900403

Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). E-S-QUAL: A Multiple-Item Scale for Assessing Electronic Service Quality. *Journal of Service Research*, *7*(3), 213–233. https://doi.org/10.1177/1094670504271156

Santos, J. (2003). E-service quality: A model of virtual service quality dimensions. *Managing Service Quality: An International Journal*, *13*(3), 233–246. https://doi.org/10.1108/09604520310476490

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Zavareh, F. B., Ariff, M. S. M., Jusoh, A., Zakuan, N., Bahari, A. Z., & Ashourian, M. (2012). E-Service Quality Dimensions and Their Effects on E-Customer Satisfaction in Internet Banking Services. *Procedia - Social and Behavioral Sciences*, *40*, 441–445. https://doi.org/10.1016/j.sbspro.2012.03.213

7 APPENDIX

7.1 Appendix 1

AI Chatbots in customer service on digital platforms

☐ Mandatory questions are marked with a star (*)

The questions are targeted towards opinions on various levels of artificial intelligence chatbots (AI ChatBots). These are used in customer service, online stores, and support services, among others. The questions concern general attitudes towards the solution of automating customer service.

This survey is part of a Master's thesis at the University of Vaasa in School of Management. All results will be handled confidentially without personal data.

Open-ended questions are not mandatory.

1. Age *

- ◯ 18-24
- ◯ 25-34
- ◯ 35-44
- ◯ 45-54
- ◯ 55-64
- ◯ Over 64

2. Gender *

- ◯ Female
- ◯ Male
- ◯ Other
- ◯ I do not wish to answer

3. Education *

- ◯ Primary school
- ◯ High school/Vocational school
- ◯ Bachelors
- ◯ Masters
- ◯ Doctorate

4. Annual income (Gross) *

- ◯ 0-9 999€
- ◯ 10 000-19 000€
- ◯ 20 000-29 999€
- ◯ 30 000-39 999€
- ◯ 40 000-49 999€
- ◯ 50 000-59 999€
- ◯ 60 000-69 999€
- ◯ 70 000-79 000€
- ◯ 80 000-89 999€
- ◯ 90 000-99 999€
- ◯ Over 100 000€
- ◯ I do not wish to answer

5. Past Experience *

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
I regularly use AI Chatbots	◯	◯	◯	◯	◯	◯	◯
I have basic understanding how AI and Customer service Chatbots work	◯	◯	◯	◯	◯	◯	◯
I am comfortable and enjoy using different AI and Chatbot services.	◯	◯	◯	◯	◯	◯	◯
I have encountered issues and problems with AI chatbots in the past	◯	◯	◯	◯	◯	◯	◯

6. Why have you previously been satisfied or dissatisfied with customer service AI chatbots?

7. Usefulness and performance *

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
AI Chatbots provide useful assistance to my problems with the service	○	○	○	○	○	○	○
AI Chatbots provide quick assistance to my problems	○	○	○	○	○	○	○
AI Chatbots make the customer service more efficient	○	○	○	○	○	○	○

8. In some situations, is AI chatbot a better way to contact for issue resolution than conversing with a human? In what situation?

9. Ease of use *

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
It is easy to use AI Chatbots	○	○	○	○	○	○	○
AI Chatbots answer me clearly and understandably	○	○	○	○	○	○	○
I feel that AI Chatbots user-friendly	○	○	○	○	○	○	○

10. Reliability *

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
I feel AI Chatbots are reliable	○	○	○	○	○	○	○
Assistance that AI Chatbots give is correct and answers my question	○	○	○	○	○	○	○
I feel that AI Chatbots use my personal data securely	○	○	○	○	○	○	○

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
I feel companies use AI Chatbots ethically	○	○	○	○	○	○	○

11. Personalization *

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
AI Chatbots are polite	○	○	○	○	○	○	○
I feel AI Chatbot should acknowledge and be able to respond to me with emotion	○	○	○	○	○	○	○
I feel more confident if AI Chatbot has human name and avatar	○	○	○	○	○	○	○

12. Satisfaction to the service *

	1 Strongly Disagree	2	3	4	5	6	7 Strongly Agree
Most likely I will continue using AI Chatbots in the future	○	○	○	○	○	○	○
The AI is the future of digital customerservice	○	○	○	○	○	○	○
I am satisfied with the service I have received from AI chatbots	○	○	○	○	○	○	○
I find AI chatbots to be useful in customer service	○	○	○	○	○	○	○
AI chatbot service exceeds my expectations	○	○	○	○	○	○	○
I prefer speaking with a human in customer service chat	○	○	○	○	○	○	○

13. I would be more satisfied with customer service AI chatbots if...