MASTERS THESIS

Investigating the role of multi-modal and uni-modal communication in developing trust in human-robot interaction with Furhat Robot

---

Author: Rayane TARKANY

Advisor: Marta ROMEO

A thesis submitted in fulfilment of the requirements
for the degree of MSc Software Engineering
in the
School of Mathematical and Computer Sciences

DECEMBER 2023

Declaration of Authorship

I, Rayane TARKANY, declare that this thesis titled, "Investigating the role of multi-modal and uni-modal communication in developing trust in human-robot interaction with Furhat robot" and the work presented in it is my own. I confirm that this work submitted for assessment is my own and is expressed in my own words. Any uses made within it of the works of other authors in any form (e.g., ideas, equations, figures, text, tables, programs) are properly acknowledged at any point of their use. A list of the references employed is included.

Signed: Rayane TARKANY

Date: 28 November 2023

Abstract

In order to gain insight into the effects of demographics, modes of interaction, shifts in perception and trust, and interaction styles, this study examines human interaction with the Furhat robot in the context of travel planning. During the interaction, a varied group of participants shared their opinions and preferences with the Furhat robot. The findings underscore the significance of modifiable modes of interaction, the function of trust in artificial intelligence systems, and the requirement for improved communication tactics that emulate human-human interactions.

The study also looks at how AI systems like Furhat are integrated into the customer service sector, highlighting the potential for optimization that arises from fusing human empathy with AI efficiency. There is also discussion of ethical issues like data security and privacy.

This study advances our knowledge of human-robot interactions and provides guidance for the creation and use of AI technologies across a range of industries. Future research directions are also indicated, such as examining the effects of cultural variables, demographic influences, and ethical considerations on AI acceptance.

Acknowledgements

I extend my heartfelt appreciation to Marta ROMEO for her invaluable guidance, unwavering support, and encouragement throughout the course of this research. Her profound knowledge, insightful perspectives, and dedicated mentorship have been instrumental in shaping my work. Her inspiration and dedication to academic excellence have profoundly impacted my journey.

Table of contents

List of figures

List of Tables

List of abbreviations

AI	Artificial Intelligence
HRI	Human-robot interaction
TPS-HRI	Trust Perception Scale - HRI

Chapter 1

Introduction

Background of the Study

Human-robot interaction (HRI) has rapidly evolved, becoming increasingly significant in various sectors including healthcare, industry, and personal assistance. As robots integrate more deeply into human environments, the nature and efficacy of their communication with humans become paramount. Traditional robotic systems often rely on uni-modal or basic multi-modal communication systems, which can limit the depth and fluidity of human-robot interactions. The advent of advanced AI language models, such as ChatGPT, presents an opportunity to revolutionize this interaction paradigm.

Problem Statement

Despite the potential of AI-driven communication systems in enhancing HRI, there is a gap in research and practical application, particularly in understanding how these advanced systems influence aspects such as trust and engagement in human users. This gap is especially pertinent when considering robots designed for complex interactions, like the Furhat robot. The integration of a sophisticated language model like ChatGPT with Furhat could potentially enhance its interaction capabilities but requires thorough investigation to understand its impacts fully.

Research Objectives

1. Implement ChatGPT within the Furhat robot platform, creating a more dynamic and naturalistic interaction model.
2. Examine how the integration of ChatGPT affects user trust and engagement during interactions with the Furhat robot.
3. Evaluate the effectiveness of ChatGPT-enhanced communication in HRI through user studies and qualitative feedback.

Significance of the Study

This research offers significant insights into AI-enhanced communication in Human-Robot Interaction (HRI), focusing on integrating ChatGPT with the Furhat robot. It aims to enhance understanding of AI language models in robotic systems, potentially improving human-robot interactions. The findings could inform the design of future robots, making them more responsive to human expressions and needs, thus promoting increased trust and collaboration between humans and robots.

Overview of the Dissertation

This dissertation is organized as follows:

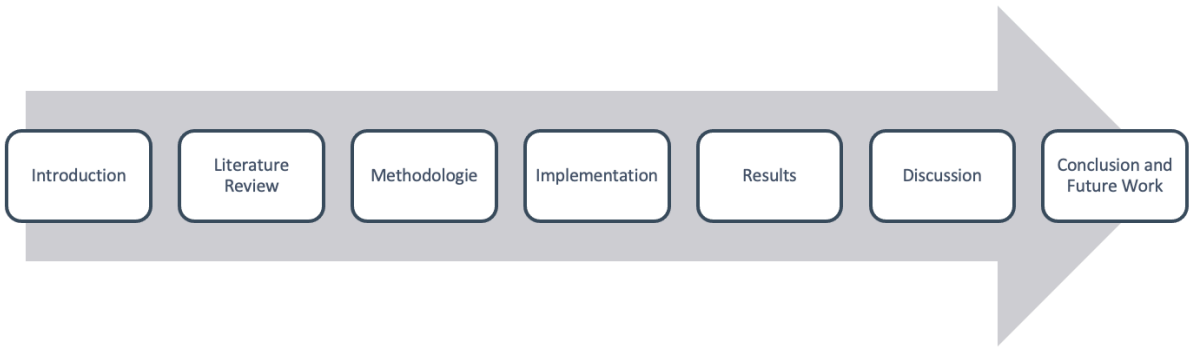

Figure 1.1: Overview of the dissertation

Chapter 2

Literature Review

2.1 Introduction of Literature Review

This review addresses the evolving dynamics in Human-Robot Interaction (HRI), focusing on the pivotal shift from uni-modal to multi-modal communication facilitated by advancements in conversational AI technologies such as ChatGPT. The progression from basic, single-mode interactions to sophisticated, multi-modal systems marks a significant advancement in the field, directly impacting user trust and the quality of interactions in HRI.

2.2 Evolution of Communication Modalities and Advancements in Conversational AI

2.2.1 From Unimodal to Multimodal Interactions

From unimodal interactions reliant on simple command-response mechanisms, HRI has progressed to complex multimodal systems integrating audio, visual, and sometimes tactile elements. This evolution (**Amirova et al. (2021)**) reflects advancements in technical capabilities and user needs. Conversational AI, particularly in speech and language processing (**Li et al. (2023)**), has played a key role in this transition. The integration of emotional intelligence in robotics, through emotion recognition and response capabilities (**Spezialetti**

et al. (2020)), and the incorporation of non-verbal communication like robot motion and gestures (**Deichler et al. (2023)**), have further enhanced the quality of HRI.

2.2.2 Advancements in Conversational AI

The evolution of Human-Robot Interaction (HRI) communication has been significantly shaped by advancements in conversational AI, particularly in three key areas:

- **Speech and Language Processing**: This area has seen remarkable improvements, enabling robots to understand and respond to human language with greater accuracy and nuance (**Li et al. (2023)**).
- **Emotion Recognition and Response**: The development of emotional intelligence in robotics, particularly in recognizing and responding to human emotions, has been crucial for more empathetic and effective communication. **Spezialetti et al. (2020)** emphasizing the role of emotional cues in HRI.
- **Non-Verbal Communication**: The integration of non-verbal cues like robot motion and gestures has augmented the multimodal nature of HRI. (**Deichler et al. (2023)**) provides insights into how these non-verbal elements enhance the naturalness and effectiveness of human-robot interactions.

2.3 Integration of ChatGPT in HRI

The integration of ChatGPT, developed by OpenAI, into Human-Robot Interaction (HRI) represents a significant advancement in the field of robotics. This section explores ChatGPT's role in HRI, its contributions to conversation dynamics, and the implications for trust and interaction quality. The review also encompasses studies that illustrate ChatGPT's integration in robotic systems and its contribution to developing more intelligent robotic applications.

2.3.1 ChatGPT's Role in Robotics

ChatGPT has revolutionized the way humans interact with robots by leveraging its advanced natural language processing capabilities. It accurately understands conversations and generates high-quality responses, making it an indispensable tool in HRI. This proficiency in natural language interaction allows ChatGPT to serve as a sophisticated language-based interface, simplifying the process of providing high-level feedback for various robotic tasks. The study referenced in the ACM Digital Library (**Xie et al. (2023)**) discusses how ChatGPT's integration into robotics enhances human-robot interactions, highlighting its utility in complex communication scenarios.

2.3.2 Strategic Implementation and Versatility of ChatGPT in Robotics Tasks

- **Strategic Implementation for Task Adaptability** ChatGPT's adaptation to diverse robotics tasks involves critical design principles like prompt engineering and a high-level function library. These elements enable its application across various robotic platforms (**Vemprala et al. (2023)**). It underscores the significance of user feedback and design strategies in customizing ChatGPT for different robotics uses.
- **Versatility in Task Execution** ChatGPT's versatility is evident in its capability to handle tasks ranging from basic reasoning to complex activities like aerial navigation. Its proficiency in natural language interaction broadens its applicability in robotics (**Vemprala et al. (2023)**).

In conclusion, the integration of ChatGPT in HRI marks a significant step forward in the field, enhancing the naturalness and effectiveness of human-robot interactions. Its role in robotics, strategic implementation for task adaptability, and versatility in task execution collectively contribute to developing more intelligent, responsive, and user-friendly robotic systems.

2.4 Research Gaps and Conclusion

While advancements in conversational AI, particularly in its integration into Human-Robot Interaction (HRI), have significantly progressed, there remain areas necessitating deeper exploration to further advance this field responsibly and sustainably.

2.4.1 Human Perception of Conversational Agents

:

- **Social Cues and Information Presentation**: Current research emphasizes the need to understand how social cues exhibited by conversational agents affect human perception. Studies are exploring the impact of these cues and the role of information presentation in hybrid conversational systems, indicating a gap in comprehending the full scope of these interactions (**Raaijmakers et al. (2023)**).
- **Anthropomorphization and Ethical Constraints**: There is a growing discourse around the potential anthropomorphization of conversational AI agents and the ethical constraints that might be imposed on them. This area, requires further exploration to understand the implications of attributing human-like characteristics to AI agents (**Kasirzadeh and Gabriel (2023)**).
- **Emotional Bond and Consumer Relationships**: The development of emotional bonds and consumer relationships through Social Companionship (SC) features in conversational agents has garnered interest. A comprehensive understanding of this domain is necessary to guide future research and application (**Chaturvedi et al. (2023)**).

2.4.2 Bias and Fairness in AI

- **Fairness and Bias in Diverse Domains**: Concerns about fairness and bias in AI systems, particularly in sensitive areas like healthcare, employment, and criminal justice, are prevalent. Research points to the need for ongoing investigation into these issues (**Ferrara (2023)**).
- **Lack of Diversity in AI Research**: The diversity in AI 'fairness' research is lacking, which could hinder the development of unbiased AI systems. This gap signifies the need for more inclusive research practices (**Wong**).
- **Ethical and Human Rights Implications**: The ethical and human rights implications of algorithmic bias in AI systems are under investigation (**Min (2023)**). Understanding and addressing these implications is crucial for the ethical integration of AI in society.

In conclusion, while the advancements in conversational AI and its integration into HRI have been remarkable, there are significant research gaps that need to be addressed. These include understanding human perception of AI agents, the societal impact of AI, issues of bias and fairness, and the broader ethical and human rights implications. Addressing these gaps is essential for the development of AI-enhanced robotic systems that are not only technologically advanced but also socially responsible and ethically grounded. This call for further research underscores the need for a multi-disciplinary approach that encompasses technical, social, ethical, and legal perspectives to ensure the sustainable development of AI technologies in harmony with human values and societal needs.

Chapter 3

Methodology

3.1 Furhat Robot: Technical Capabilities

Furhat's technical framework is crucial for its role in our study. It features an advanced motion platform for lifelike head movements, onboard cameras for participant interaction, and microphones optimized for clear audio capture in various environments. Equipped with a projection system for expressive facial animations and a customizable face mask, Furhat can convey a range of emotions and reactions, mirroring human-like expressions.

Figure 3.1: Furhat Robot Overview

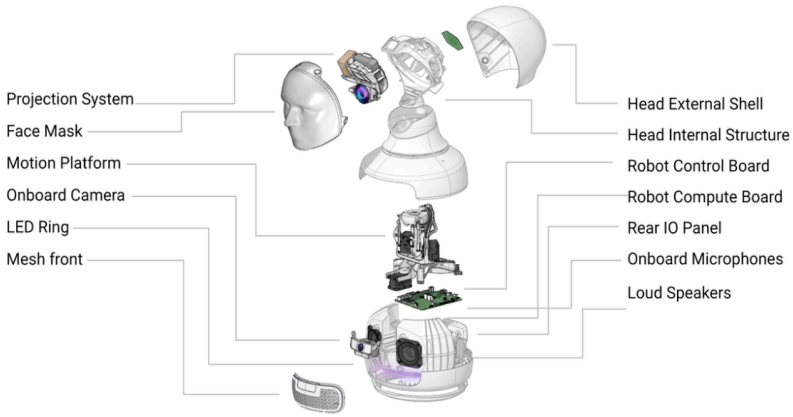

Figure 3.2: Furhat Robot Component Overview

3.2 Scenario Overview and Setup

3.2.1 Objective and Tasks

The primary objective of this study is to assess participants' trust in and engagement with an AI-powered robot, Furhat, during a travel planning process. Participants will engage in a simulated scenario where they are required to plan a multi-destination trip. The tasks include selecting destinations, choosing accommodations, planning activities, and managing a travel budget. This setup aims to mimic real-life travel planning complexities and assess how effectively a robot can assist in such tasks.

3.2.2 Detailed Task Breakdown

The experiment's tasks are designed to engage participants in various aspects of travel planning:

- **Destination Selection**: Participants will choose destinations, considering factors like cultural interests and climate preferences.
- **Accommodation Planning**: Selection of accommodations based on budget, location, and amenities.
- **Activity Scheduling**: Planning of trip activities, balancing time, variety, and personal interests.
- **Budget Management**: Participants will manage a simulated budget, making decisions on allocating funds across different travel aspects.

3.2.3 Physical Setup for the Experiment

The experiment will involve two distinct setups to compare multi-modal and uni-modal interactions with Furhat.

- For the **multi-modal setup**, Furhat will be equipped to communicate using a combination of voice and facial expressions. This setup will emphasize the robot's ability to engage participants with dynamic and expressive communication.
- In contrast, the **uni-modal setup** will involve Furhat communicating solely through voice, without any visual cues, focusing purely on auditory interaction.

Both setups will be conducted in a simple, neutral environment without thematic decorations or background music. This approach ensures that the focus remains solely on the interaction with Furhat, allowing for a clear comparison between the two communication modes. The absence of additional decor and music is intended to minimize external influences on participants' perception and interaction with the robot.

3.2.4 Role of Furhat and Integration of Trust Scales

Scale Application

To rigorously assess trust in human-robot interaction, this study employs established trust scales, specifically adapted to the context of our research. These scales include:

- **HRI Trust Scale**: Adapted to measure participants' trust in Furhat's capabilities and its effectiveness in assisting with travel planning. This scale is particularly relevant for evaluating the nuances of trust in a robot performing complex tasks.
- **Trust in Automation Scale**: Used to gauge participants' perceived reliability of Furhat in providing accurate and helpful information. This scale helps in understanding the participants' confidence in the robot's decision-making abilities.
- **Interpersonal Trust Scale**: Although originally designed for human-human interactions, elements of this scale are adapted to assess participants' willingness to rely on Furhat for important travel decisions. This adaptation is crucial to understanding how human-like trust dynamics transfer to human-robot interactions.

Tailoring Questions

- **Context-Specific Adaptations**: Questions in each scale will be modified to align with the travel planning context. For instance, the HRI Trust Scale will include items that assess trust in Furhat's ability to understand complex travel preferences and provide personalized suggestions.
- **Scenario-Based Questions**: Questions will be designed to reflect scenarios that participants might encounter during the travel planning process, such as choosing between different accommodation options or adjusting plans due to budget constraints.

Timing of Survey Administration

- **Pre-Interaction Survey**: Before interacting with Furhat, participants will complete the adapted scales to establish a baseline of their trust and perceptions (**cf. Section 8.1**).
- **Post-Interaction Survey**: After the interaction, participants will complete the scales again to measure any changes in their trust and perceptions of Furhat. This will help in assessing the impact of the interaction on their trust levels (**cf. Section 8.2**).
- **Strategic Timing**: The timing of these scales is crucial to capture the dynamic nature of trust as it evolves through the interaction with Furhat. Immediate post-interaction administration will capture participants' fresh impressions and any shifts in their trust levels.

3.3 Communication Strategies and User Study Protocol

3.3.1 Communication Strategies

The study will explore different communication strategies employed by Furhat:

- **Uni-modal Communication**: Focuses on verbal interaction, where Furhat communicates using speech only.
- **Multi-modal Communication**: Involves Furhat using a combination of speech and visual cues.

3.3.2 Enhanced Communication Strategies

Adapting to Participant Responses

3.3.3 User Study Protocol

- **Pre-study Preparation**: Participants will complete a baseline survey to gather initial data on their attitudes and experiences with AI and robotics.
- **Interaction Process**: A detailed guide for the interaction process will be provided and how participants are expected to engage with Furhat.
- **Post-interaction Evaluation**: After the interaction, participants will complete the post-interaction survey to provide immediate feedback.

Chapter 4

Experiment Design

4.1 Interaction Design and Implementation

This section outlines the customization of the Furhat robot's software and user interface, as well as the design of the interaction scripts and the flow of dialogue with participants, highlighting the blend of technical preparation and practical interaction design in the experiment.

- **Software Configuration and Script Development**: A critical aspect of the experiment involved customizing Furhat's software to meet the specific requirements of a travel planning dialogue. This entailed programming the robot with a series of scripts designed to anticipate various travel-related inquiries and responses, with a focus on language choice and response timing to emulate a real-life travel advisor. The scripting process was iteratively developed and refined, aiming to provide informative and relevant content while allowing for a dynamic, conversational flow. These scripts incorporated typical travel-related queries and responses one would expect in a travel agency setting.
- **User Interface Adjustments for Enhanced Interaction**: To facilitate an intuitive and engaging interaction, Furhat's user interface was adapted with on-screen prompts and visual cues. These adjustments complemented the verbal interaction, ensuring an easy and natural experience for participants.

- **Semi-Structured Interaction Flow**: The interaction between the participants and Furhat was designed to be semi-structured. It began with Furhat introducing itself as a travel advisor and guiding participants through the process of expressing their travel requirements, including destinations, budget, and preferences. Powered by ChatGPT, Furhat responded with suggestions and engaged in back-and-forth dialogue to refine trip details. While a basic script provided consistency, the interaction was flexible, allowing for natural conversation and adaptability to each participant's unique responses, thus ensuring a realistic and fluid communication experience.

This comprehensive approach to interaction design combines technical software customization with practical script development and user interface adjustments, ensuring that the experiment not only tests the capabilities of the Furhat robot but also closely mimics real-world human-robot interactions in a travel planning context.

4.2 Data Collection

4.2.1 Enhanced Data Collection

Demographic Correlation

- **Analysis of Demographic Influence**: The study will examine how different demographic factors such as age, gender, education level, and occupation influence participants' interactions with Furhat. This analysis will help in understanding if and how these demographic variables impact trust and engagement levels in HRI.
- **Tailored Interaction Analysis**: The data collected will be used to assess whether certain demographics have distinct preferences or reactions to the AI-powered robot. For instance, younger participants might show different levels of comfort or trust in AI technology compared to older participants.
- **Cross-Sectional Insights**: By correlating these demographic details with trust and

perception changes, the study aims to offer cross-sectional insights, potentially identifying unique trends or needs among different user groups.

Expectation vs. Reality

- **Pre-Interaction Expectations**: Participants' initial expectations about their interaction with Furhat will be gathered through the pre-interaction survey. This will include their anticipated level of trust, their curiosity or concerns about specific aspects of the interaction, and their overall interest in the task.
- **Post-Interaction Reality Check**: After the interaction, participants will provide feedback on how their actual experience compared with their initial expectations. This will be crucial in understanding how the real interaction with Furhat aligns with or diverges from their preconceived notions.
- **Comparative Analysis**: By comparing the pre- and post-interaction data, the study will be able to analyze changes in trust, satisfaction, and perception. This comparison will reveal whether the interaction with Furhat enhanced, met, or fell short of participants' expectations, offering valuable insights into the effectiveness and impact of AI-assisted interactions.

This enhanced data collection approach, focusing on demographic correlations and expectation-reality comparisons, is designed to provide a comprehensive understanding of the factors influencing trust and engagement in HRI. It will also allow for the identification of specific areas where the AI-powered robot either excels or needs improvement, based on the diverse experiences and expectations of participants.

4.3 Analysis Strategy and Ethical Considerations

4.3.1 Data Analysis

- **Statistical Methods**: The study will employ a range of statistical methods to analyze survey data. This includes descriptive statistics to summarize the data and inferential statistics, such as t-tests or ANOVA, to compare the pre- and post-interaction survey responses. These methods will help in identifying significant changes in participants' trust and perceptions towards AI-assisted travel planning.
- **Change Detection**: Special attention will be given to detecting any significant changes in trust and perception levels before and after the interaction with Furhat. This involves comparing responses from the pre-interaction and post-interaction surveys to identify patterns or shifts in participants' attitudes towards the AI-powered robot.

4.3.2 In-Depth Data Analysis

- **Demographic Data Analysis**: The analysis will delve into how different demographic factors—such as age, gender, education level, and occupation—affect participants' experiences and trust in HRI. This will involve correlating these demographic variables with changes in trust and perception, using techniques like regression analysis to uncover any significant relationships.
- **Trust Dynamics**: A key focus will be on the dynamics of trust throughout the interaction process. This involves analyzing how participants' trust in Furhat evolves from the beginning to the end of the interaction, and what factors (such as the robot's communication style or the accuracy of its information) most significantly influence this evolution.

4.3.3 Ethical Aspects

- **Informed Consent**: All participants will be provided with clear information about the study's nature, their role, and their rights before participation. This includes informing them that no recording will be conducted. (**cf. Section 8.3**).
- **Privacy and Data Protection**: Stringent measures will be implemented to ensure the confidentiality and security of the data collected. Participants' personal information will be anonymized, and data will be stored securely.

These analytical strategies, combined with a commitment to ethical research practices, will provide a comprehensive understanding of the factors influencing trust and engagement in human-robot interaction. This approach is crucial for ensuring that the study's findings are robust, reliable, and can contribute meaningfully to the field of HRI.

4.4 Conclusion

This chapter presents a structured experimental design for investigating trust in human-robot interaction (HRI), with a specific focus on the effects of uni-modal and multi-modal communication using the Furhat robot. By integrating established trust scales such as the HRI Trust Scale and Trust in Automation Scale, the study employs a validated approach to measure trust, ensuring the results are both reliable and comparable with existing research.

The implementation of two distinct communication strategies with Furhat allows for a comprehensive exploration of how different interaction modes impact participants' trust and engagement. The design carefully considers the nuances of each mode, aiming to provide a clear understanding of their respective influences on the user experience.

Ethical considerations, especially regarding informed consent and data privacy, have been addressed meticulously, reflecting the commitment to conducting responsible and respectful research.

In summary, this research is set to offer valuable insights into trust dynamics in HRI, contributing to the growing body of knowledge in this field. The findings are anticipated to have significant implications for the development and application of AI technologies in contexts where trust is a key component of human-robot interactions.

Chapter 5

Results

The results from your pre-interaction and post-interaction surveys provide valuable insights into the participants' perceptions, experiences, and preferences regarding the Furhat robot interaction in a travel planning context. Here's a brief overview:

5.1 Pre-Interaction Survey Results

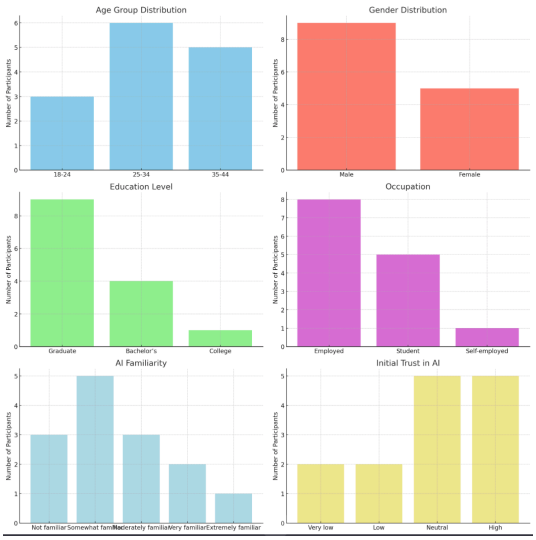

Figure 5.1: Visual representation of the Pre-interaction survey results

5.1.1 Demographic Distribution

- **Age Group**: Majority are aged 25-34 (6 participants), followed by 35-44 (5 participants), and 18-24 (3 participants).
- **Gender**: 9 participants identified as Male, and 5 as Female.
- **Education Level**: Majority hold a Graduate degree (9 participants) followed by Bachelor's degree (4 participants), and 1 participant with a College degree, indicating a higher education level.
- **Occupation**: 8 are Employed (full-time/part-time), 5 are Students, and 1 is Self-employed.

5.1.2 Familiarity with AI and Initial Trust

- **AI Familiarity**: 5 participants are 'Somewhat familiar', 3 are 'Not familiar at all', 3 are 'Moderately familiar', 2 are 'Very familiar', and 1 is 'Extremely familiar'. This suggests a diverse level of exposure to AI technologies among participants.
- **Initial Trust in AI**: 5 rate 'Neutral' trust, 5 rate 'Low trust', 2 rate 'High trust', and 2 rate 'Very low trust' in AI for decision-making. This variation in trust levels could influence their interaction and perception of the Furhat robot.
- **Willingness to Depend on AI**: 5 participants are 'Neutral', 4 are 'Very low', 3 are 'High', and 2 are 'Low' in their willingness to depend on AI in important situations. This could reflect their overall comfort with AI technologies in critical situations.
- **Expectations for Furhat Interaction**: 7 are 'Very interested', 5 are 'Moderately interested', 1 is 'Slightly interested', and 1 is 'Extremely interested'. This indicate varying levels of enthusiasm towards the interaction.
-

5.1.3 Additional Insights

- **Experience with Travel Planning Services**: Most participants have used a travel planning service, which is relevant to the context of the experiment.
- **Satisfaction with Travel Services**: Levels of satisfaction vary, with responses like 'Satisfied', 'Neutral', and some non-responses. This possibly indicating no prior experience.
- **Key Factors in Travel Planning**: Participants prioritize various aspects such as budget, destination variety, accommodation quality, activities, and experiences. This indicates varied personal preferences and considerations in travel planning.

Understanding these nuances is essential for interpreting the post-interaction data, as they offer insights into the varied human factors that play a role in human-robot interactions. This understanding also aids in tailoring more effective and user-friendly robotic systems in the future, by considering the diverse backgrounds and perspectives of potential users. Thus, the demographic distribution and initial perceptions not only provide a context for the experiment but also serve as predictors for the variances observed in the post-interaction responses.

5.2 Post-Interaction Survey Results

5.2.1 Overall Interaction Experience

- **Participant Impressions**: Majority rate 'satisfied' (9 responses), followed by 2 'Very Satisfied', 2 'Neutral' and 1 'Dissatisfied'.
- **Ease of Interaction**: 9 responses reflect effective and straightforward communication ('Easy'), 3 suggest moderate ease of interaction ('Neutral'), 1 indicates some challenges faced by participants ('Very Easy') and 1 shows an especially smooth experience for some participants ('Dificult').

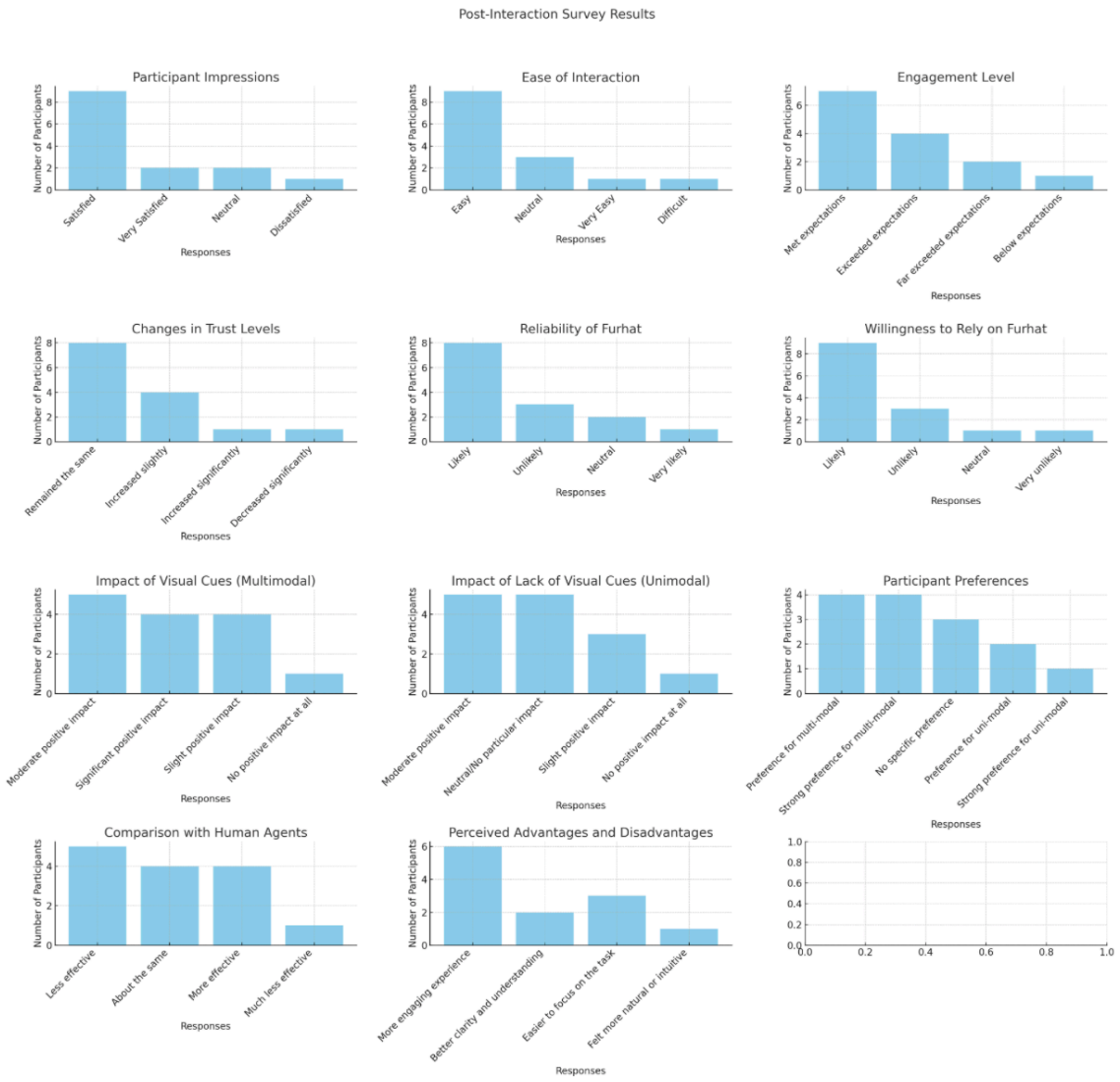

Figure 5.2: Visual representation of the Post-interaction survey results

- **Engagement Level**: 7 responses indicate that the interaction aligned well with what was anticipated ('Met expectations'), 4 suggest a better-than-expected experience ('Exceeded expectations'), 2 indicate a highly positive experience('Far exceeded expectations') and 1 response points to a less satisfactory experience ('Below expectations').

5.2.2 Changes in Trust and Perception

- **Changes in Trust Levels**: 8 responses suggest no significant change in trust levels ('Remained the same'), 4 responses indicate a mild positive shift in trust ('Increased

slightly'), 1 response reflects a substantial increase in trust ('Increased significantly') and 1 response highlights a notable decrease in trust (Decreased significantly).

- **Reliability of Furhat in Providing Information**: 8 responses suggest a favorable perception of Furhat's reliability ('Likely'), 3 responses indicate skepticism about Furhat's reliability ('Unlikely'), 2 responses show an undecided stance ('Neutral') and 1 response indicates a very high level of trust in Furhat's reliability ('Very likely').
- **Willingness to Rely on Furhat for Important Decisions**: 9 responses reflect a readiness to depend on Furhat for critical decisions ('Likely'), 3 responses show reluctance to rely on Furhat ('Unlikely'), 1 response indicates uncertainty ('Neutral') and 1 response shows a strong reluctance to depend on Furhat ('Very unlikely').

5.2.3 Communication Mode Evaluation

- **The impact of visual cues in multimodal interaction**: Moderate positive impact: 5 responses for 'Moderate positive impact', 4 responses for 'Significant positive impact', 4 responses for 'Slight positive impact' and 1 response for 'No positive impact at all'.
- **The impact of the lack of visual cues in unimodal interaction**: 5 responses for 'Moderate positive impact', 5 responses for 'Neutral/No particular impact', 3 responses for 'Slight positive impact' and 1 response for 'No positive impact at all'.

5.2.4 Comparative Analysis with Human Interaction

- **Participant Preferences**: 4 responses suggest a preference for multi-modal interaction. Another 4 responses indicating a strong preference for multi-modal interaction, 3 participants did not express a specific preference, 2 responses indicating a preference for uni-modal interaction and 1 response showing a strong preference for uni-modal interaction.
- **Comparison with Human Agents**: 5 responses suggest Furhat is perceived as 'less

effective' than human agents in some cases, 4 responses indicate a perception of Furhat being on par with human agents ('About the same'), 4 responses show that Furhat is seen as 'more effective' than human agents by some participants, and 1 response indicates a strong perception of Furhat being 'less effective' than human agents.

- **Perceived Advantages and Disadvantages**: 6 responses could indicate that participants see engagement as a key strength of AI interaction ('More engaging experience'), 2 responses suggest that clarity is seen as an advantage in AI interactions over human agents ('Better clarity and understanding'), 3 responses might reflect an area where human interaction is preferred or seen as more effective ('Easier to focus on the task') and 1 response indicates that intuitiveness is an important factor in participants' perception of AI capabilities ('Felt more natural or intuitive').

Chapter 6

Discussion

6.1 Interpretation of Results

6.1.1 Demographics and Initial Perceptions

In examining the demographic distribution and initial perceptions of the participants, several key insights emerge, which are crucial for understanding the interaction dynamics with the Furhat robot in a travel planning context.

Influence of Demographics on Perceptions and Interactions

The age group distribution predominantly in the 25-34 and 35-44 range might have influenced the participants' ease of adapting to and engaging with AI technology. Generally, these age groups are considered to be more tech-savvy and open to technological innovations, which could have contributed to their readiness to interact with Furhat. A study in Nature Machine Intelligence (**Pataranutaporn et al. (2023)**) suggests that perceptions of AI, shaped by mental models, significantly impact trust and engagement levels in AI interactions

The study's demographic composition, skewed towards certain age groups and educational backgrounds, might not fully represent broader population interactions with AI. A scoping review on age-related bias in AI systems, as discussed in Humanities and Social Sciences Communications, shows the prevalence and implications of digital ageism **(Chu et al. (2023))**. Additionally, gender distribution, with a majority of male participants,

raises questions about gender influence on AI interaction preferences. A study on gender considerations in AI from a developing-world perspective indicates significant differences in AI robot perceptions between male and female respondents **(Kumar and Choudhury (2022))**.

Correlation Between Demographics and AI Familiarity

The diversity in AI familiarity across participants indicates a varied baseline from which perceptions of the Furhat robot were formed. Research in the International Journal of Educational Technology in Higher Education (**Seo et al. (2021)**) underlines how AI systems impact communication and support in online learning, reflecting on different perceptions based on user backgrounds.

Initial Expectations from Furhat Interaction

Participants' willingness to depend on AI in important situations, which ranged from very low to high, is a significant indicator of their openness to AI assistance in complex tasks like travel planning. A study published in Frontiers in Psychology explored the factors influencing college students' willingness to accept AI-assisted learning environments. This research, based on the Unified Theory of Acceptance and Use of Technology (UTAUT) and the theory of perceived risk, identified six factors affecting students' willingness to use AI. The findings indicated that effort expectancy, performance expectancy, and social influence positively related to willingness to accept AI, while psychological risk had a significant negative influence (**Wu et al. (2022)**).

In summary, the demographic factors and initial perceptions provide essential context for understanding the nuances of human-robot interaction observed in this study. They highlight the importance of considering individual differences in technology acceptance and trust in AI, which are pivotal in tailoring AI technologies for diverse user groups. This understanding

is critical for designing more effective and user-friendly robotic systems in various practical applications.

6.1.2 Impact of Furhat's Interaction Modes

The study's exploration into the impact of Furhat's uni-modal and multi-modal interaction modes reveals insightful findings regarding participants' experiences and preferences. These interaction modes, differing primarily in the use of visual cues, play a significant role in shaping the user experience and perception of the AI system.

Preference for Multi-modal Interaction

A notable preference for multi-modal interaction was observed among participants. This preference can be attributed to the enriched communication experience that multi-modal interaction provides. Visual cues, such as facial expressions and gestures, often make interactions more engaging and intuitive. The effectiveness of multi-modal interaction aligns with existing research (**Su et al. (2023)**) suggesting that human communication is significantly enhanced by non-verbal cues. In the context of Furhat, the use of expressive facial animations and gestures likely made the interactions more relatable and human-like, enhancing user engagement and satisfaction.

The preference for multi-modal interaction also suggests a potential direction for future AI and robotics development. It underscores the importance of designing robots and AI interfaces that can mimic human-like communication styles to improve user experience.

Uni-modal Interaction and Its Challenges

The uni-modal interaction, which lacked visual cues, received mixed responses. Some participants found it satisfactory, indicating that clear and effective verbal communication can suffice for certain interactions. However, others found it less engaging, suggesting the absence of visual cues might lead to a perceived lack of dynamism and empathy in the interaction.

The varied responses to uni-modal interaction highlight the diverse preferences and needs of users when interacting with AI systems. It suggests that while some users may prefer straightforward, voice-only communication, others might find it lacking in terms of engagement and emotional connection.

The uni-modal interaction, which lacked visual cues, received mixed responses. Some participants found it satisfactory, indicating that clear and effective verbal communication can suffice for certain interactions (**Oviatt (1999)**).However, others found it less engaging, suggesting the absence of visual cues might lead to a perceived lack of dynamism and empathy in the interaction (**Cassell (2000)**). The varied responses to uni-modal interaction highlight the diverse preferences and needs of users when interacting with AI systems (**McTear et al. (2016)**). It suggests that while some users may prefer straightforward, voice-only communication, others might find it lacking in terms of engagement and emotional connection (**Cassell (2000)**).

Implications for AI Interaction Design

These findings have significant implications for the design of AI and robotic systems, especially in customer service settings. They suggest that incorporating multi-modal communication capabilities could lead to higher user satisfaction and engagement. However, the varied responses also point to the necessity of offering customizable interaction modes. This customization allows users to choose their preferred method of interaction, catering to a broader range of needs and preferences.

In conclusion, the impact of Furhat's interaction modes on participants' experiences emphasizes the importance of considering various communication strategies in AI development. The preference for multi-modal interaction suggests that more human-like, engaging communication methods are generally favored. However, the mixed responses to uni-modal interaction underline the need for flexible and adaptable communication options in AI systems,

ensuring a wider appeal and effectiveness across different user groups.

6.1.3 Changes in Trust and Perception

The post-interaction survey results provide crucial insights into how the interaction with the Furhat robot influenced participants' trust and perception of AI in the context of travel planning. These changes are pivotal in understanding the effectiveness of AI-assisted interactions and the factors that shape user trust in AI systems.

Analysis of Trust Dynamics

The data indicated that for a significant number of participants, trust levels remained the same after the interaction. This could suggest that the Furhat robot met the participants' expectations but did not necessarily exceed them to a degree that would significantly shift their trust levels.

However, there were also indications of both positive and negative shifts in trust. The mild positive shift in trust observed in some participants could be attributed to a satisfactory interaction experience, where Furhat effectively assisted in travel planning tasks, demonstrating reliability and competence.

The instances of decreased trust, though minimal, are significant. They might point to areas where Furhat's performance did not align with participants' expectations or needs, such as possible gaps in understanding user queries, response relevance, or the naturalness of the interaction.

Perception of Furhat's Reliability and Decision-Making Assistance

The generally favorable perception of Furhat's reliability in providing information suggests that the robot's responses were mostly accurate and helpful. As **Luger and Sellen (2016)**, accuracy is crucial for building trust in AI systems, where the correctness of information directly impacts user reliance. The positive view of Furhat's reliability indicates its responses

were likely accurate, helping to foster user trust. Getting responses right is key for conversational agents to deliver a satisfactory user experience rather than undermining perceptions with inaccurate information.

Factors Influencing Trust and Perception Changes

Factors Influencing Trust and Perception Changes: These changes in trust and perception can be influenced by several factors, including the quality of interaction (e.g., clarity of communication, response relevance), the robot's ability to understand and cater to individual preferences, and the overall user experience (**Pak et al. (2012)**). The role of initial expectations cannot be overlooked. Participants with higher initial trust or more positive attitudes towards AI might have had their views reinforced, while those with reservations could have been more critical of their interaction experiences (**Lee and See (2004)**).

In summary, the changes observed in trust and perception post-interaction with Furhat reveal a nuanced landscape of user trust in AI. While some users exhibited increased trust and positive perceptions, others showed no change or a decrease in trust. This variability highlights the importance of personalizing AI interactions to cater to diverse user needs and expectations. Furthermore, these findings suggest a need for ongoing improvement in AI systems, focusing on aspects such as naturalness of interaction, understanding of user intent, and the accuracy of information, to build and maintain user trust effectively.

6.1.4 Participants' Preferences and Perceptions

The study reveals a mixed response regarding the preference for multi-modal versus uni-modal interaction when compared to human interactions. Some participants exhibited a preference for multi-modal interaction, likely due to its closer resemblance to human-like communication, which includes both verbal and non-verbal cues. Conversely, the preference for uni-modal interaction in some cases might reflect a comfort with or a perceived efficiency

of straightforward, verbal-only communication, especially in scenarios where visual cues are deemed unnecessary or distracting.

Participants who viewed Furhat as on par with or more effective than human agents highlight the advancements in AI technology, particularly in terms of providing accurate information, maintaining consistency in service, and possibly in managing large volumes of data efficiently (**Go and Sundar (2019)**). Contrastingly, human agents remain indispensable for their unique qualities, primarily empathy and experience. They excel in compassionate messaging, dynamic responsiveness, cultural sensitivity, an experience-driven approach, and understanding human emotional tones, which AI cannot replicate (**Waytz et al. (2010)**). These aspects highlight the challenges AI faces in fully emulating human emotional intelligence and contextual understanding.

The perception of Furhat as less effective than human agents in certain aspects suggests areas where AI technology still lags behind human capabilities, such as empathy, understanding nuanced human expressions, or adapting to unpredictable conversation flows.

6.1.5 Comparison with Human Agents

Participants who viewed Furhat as on par with or more effective than human agents highlight the advancements in AI technology, particularly in terms of providing accurate information, maintaining consistency in service, and possibly in managing large volumes of data efficiently (**Go and Sundar (2019)**). Contrastingly, human agents remain indispensable for their unique qualities, primarily empathy and experience. They excel in compassionate messaging, dynamic responsiveness, cultural sensitivity, an experience-driven approach, and understanding human emotional tones, which AI cannot replicate fully at present (**Waytz et al. (2010)**). These aspects highlight some of the challenges AI faces in fully emulating human emotional intelligence and contextual understanding.

6.1.6 Perceived Advantages and Disadvantages

The responses indicating a more engaging experience with AI interactions could be due to the novelty factor or the precision and speed of information delivery by AI systems. Clarity and understanding in AI interactions over human agents point towards the structured and consistent responses provided by AI, which can be less prone to human error or variability. The preference for human interaction in certain scenarios indicates the intrinsic value of human empathy, intuition, and the ability to handle complex, non-linear conversations. This aspect remains a significant challenge for AI to replicate fully.

In conclusion, the comparative analysis with human interaction highlights both the achievements and the limitations of current AI technology in customer service settings. While AI, as represented by the Furhat robot, excels in consistency, information accuracy, and possibly engagement through novelty, it still faces challenges in areas requiring deep emotional intelligence and nuanced understanding. These insights are crucial for guiding the future development of AI systems, where a balance between AI efficiency and human empathy is essential. Moreover, this analysis underscores the importance of setting realistic expectations for AI capabilities, both in terms of current potential and areas for future improvement.

6.2 Implications of Findings

The study's findings offer several key implications for the future development of AI and robotics, especially in enhancing user experience and building trust.

1. **User-Friendly AI Systems**: The varied preferences for interaction modes emphasize the need for AI systems that are not only technically advanced but also user-friendly. Developers should focus on creating intuitive interfaces that can adapt to different user preferences, possibly by incorporating more natural, human-like communication

methods.

2. **Building Trust in AI**: Trust is a critical factor for effective human-robot interaction. The findings suggest that trust can be enhanced by improving the reliability, transparency, and predictability of AI systems. Efforts should be made to develop AI that can explain its reasoning, provide consistent performance, and reliably handle tasks.
3. **Enhanced Communication Strategies**: The preference for multi-modal communication indicates that future AI developments should include sophisticated non-verbal cues, mimicking human interactions more closely. This could involve advanced gesture recognition, facial expression analysis, and context-aware dialogue systems.
4. **Ethical Considerations**: As AI systems become more integrated into daily life, ethical considerations, including privacy, consent, and data security, become increasingly important. Developers need to ensure these systems adhere to ethical standards and regulations.

In conclusion, the implications of these findings for AI and robotics development and the customer service industry are profound. They highlight the need for user-centered design, trust-building, ethical considerations, and effective integration strategies. As AI technology continues to evolve, these insights will be invaluable in guiding its development and application in various sectors, particularly in enhancing human-robot interactions and customer service experiences.

6.3 Limitations and Future Research

6.3.1 Study Limitations

This research, while providing valuable insights, has certain limitations that must be acknowledged:

1. **Sample Size and Composition**: The study's relatively small sample size may limit

the generalizability of the findings. Additionally, the demographic composition of the participants, skewed towards certain age groups and educational backgrounds, may not fully represent the broader population's interaction with AI.

2. **Controlled Experimental Environment**: Conducting the study in a controlled environment may not accurately reflect the dynamics of real-world human-AI interactions. The lack of environmental variables and real-life distractions might have influenced the participants' engagement and responses.
3. **Focus on Single AI System**: The study centered exclusively on the Furhat robot, which, while advanced, is only one example of AI interaction platforms. Different AI systems with varying capabilities might elicit different user responses.
4. **Short Interaction Duration**: The limited duration of interaction with Furhat may not capture the evolving nature of trust and perception that could develop over longer periods.

6.3.2 Directions for Future Research

Building on these limitations, future research directions can include:

1. Expanding Sample Size and Diversity: Future studies should aim for larger and more diverse participant groups to enhance the representativeness and reliability of the findings.
2. Exploring Various AI Platforms: Comparing different AI interaction platforms can offer a more comprehensive understanding of user experiences across various AI technologies.
3. In-Depth Demographic Analysis: Further research could explore how specific demographic factors influence user interactions with AI, providing insights for more personalized and effective AI designs.
4. Ethical Considerations and User Privacy: As AI becomes more prevalent, studies focusing on the ethical implications, particularly regarding user privacy and data security,

will be crucial.

5. Impact of Cultural Factors: Exploring how cultural backgrounds and societal norms influence AI acceptance and interaction could provide valuable global insights.

In conclusion, while this study lays a foundational understanding of human-robot interaction dynamics, future research in more diverse, long-term, and real-world contexts is essential. These studies will be critical in shaping AI systems that are not only technologically advanced but also ethically sound, user-friendly, and socially accepted.

Chapter 7

Conclusion

This study has provided valuable insights into human interaction with the Furhat robot in a travel planning context, shedding light on the impact of demographics, interaction modes, changes in trust and perception, and a comparison with human interaction. These findings offer important implications for the development of AI and robotics as well as their integration into the customer service industry.

In summary, the key takeaways from this research can be summarized as follows:

- Diverse User Preferences: The study revealed that users have diverse preferences when it comes to interacting with AI systems like Furhat. While some participants favored multi-modal interactions with visual cues, others found uni-modal interactions satisfactory. This highlights the need for customizable interaction modes to cater to individual preferences.
- Building Trust: Trust in AI systems is a complex interplay of factors such as performance, transparency, and user expectations. Developers should focus on enhancing reliability, transparency, and predictability to build and maintain trust in AI technologies.
- Enhancing Communication Strategies: Multi-modal communication, which includes non-verbal cues, was generally preferred by users. Future AI systems should incorporate advanced gesture recognition, facial expression analysis, and context-aware dialogue systems to enhance the user experience.

- Integration in Customer Service: The study suggests that AI systems like Furhat can be effectively integrated into customer service settings to handle routine queries and provide consistent service. This dual approach, combining AI efficiency with human empathy, can optimize customer experiences.
- Ethical Considerations: As AI technology becomes more integrated into daily life, ethical considerations, including privacy, consent, and data security, should be a priority in AI system development and deployment.
- Future Research: While this study provides valuable insights, future research should focus on larger and more diverse samples, explore different AI platforms, conduct in-depth demographic analyses, and delve into ethical and cultural factors that influence AI acceptance.

In conclusion, this study underscores the complexity of human-robot interactions and the importance of user-centered design, trust-building, and ethical considerations in the development of AI and robotics. As technology continues to advance, understanding and addressing user preferences and concerns will be pivotal in creating AI systems that enhance human experiences across various domains.

Chapter 8

Annexes

8.1 Survey: Pre-Interaction Survey for HRI Experiment

Participant Information

Age:

- ☐ Under 18
- ☐ 18-24
- ☐ 25-34
- ☐ 35-44
- ☐ 45-54
- ☐ 55-64
- ☐ 65 or older

Gender:

- ☐ Male
- ☐ Female
- ☐ Prefer not to say

Education Level:

☐ High school or lower

☐ College degree

☐ Bachelor's degree

☐ Graduate degree

Occupation:

☐ Student

☐ Employed (full-time/part-time)

☐ Self-employed

☐ Unemployed

☐ Retired

Other: _____

Familiarity with AI (Trust in Automation Scale adaptation):

☐ Not familiar at all

☐ Somewhat familiar

☐ Moderately familiar

☐ Very familiar

☐ Extremely familiar

Initial Trust and Expectations

[HRI Trust Scale adaptation] How reliable do you consider AI technology in decision-making?

☐ Very low trust

☐ Low trust

☐ Neutral

☐ High trust

☐ Very high trust

[Interpersonal Trust Scale adaptation] To what extent are you willing to depend on AI in important situations?

☐ Very low

☐ Low

☐ Neutral

☐ High

☐ Very high

What are your expectations for the interaction with the Furhat robot?

☐ Not at all interested

☐ Slightly interested

☐ Moderately interested

☐ Very interested

☐ Extremely interested

Please indicate your level of curiosity or concern about each of the following aspects of interacting with the Furhat robot.

Rate each aspect on a scale where 1 indicates 'Not curious or concerned at all', 2 'Slightly Curious or Concerned', 3 'Moderately Curious or Concerned', 4 'Very Curious or Concerned'

Rate	1	2	3	4	5
Understanding Furhat's Responses					
Furhat's Ability to Understand Your Queries					
Naturalness of Furhat's Communication					
Furhat's Response Time					
Relevance of Furhat's Suggestions					

Table 8.1: Level of curiosity or concern

and 5 indicates 'Extremely curious or concerned'.

Travel Planning Experience

Have you ever used a service (online or offline) for travel planning?

☐ Yes

☐ No

If yes, how satisfied were you with the service?

☐ Very dissatisfied

☐ Dissatisfied

☐ Neutral

☐ Satisfied

☐ Very satisfied

What are the key factors you consider important in travel planning? (Select all that apply)

☐ Budget

☐ Destination variety

☐ Accommodation quality

☐ Activities and experiences

☐ Travel convenience

Other: _____

8.2 Survey: Post-Interaction Survey for HRI Experiment

Overall Experience

How satisfied were you with your interaction with the Furhat robot?

- ☐ Very dissatisfied
- ☐ Dissatisfied
- ☐ Neutral
- ☐ Satisfied
- ☐ Very satisfied

How would you rate the ease and clarity of communication with Furhat, including understanding its responses and expressing your needs?

- ☐ Very difficult
- ☐ Difficult
- ☐ Neutral
- ☐ Easy
- ☐ Very easy

Did the interaction with Furhat meet your expectations?

- ☐ Far below expectations
- ☐ Below expectations
- ☐ Met expectations
- ☐ Exceeded expectations
- ☐ Far exceeded expectations

Changes in Trust and Perception

[HRI Trust Scale adaptation] How much do you trust Furhat's capabilities post-interaction?

- ☐ Decreased significantly
- ☐ Decreased slightly
- ☐ Remained the same
- ☐ Increased slightly
- ☐ Increased significantly

[Trust in Automation Scale adaptation] How reliable did you find Furhat in providing information?

- ☐ Very unlikely
- ☐ Unlikely
- ☐ Neutral
- ☐ Likely
- ☐ Very likely

[Interpersonal Trust Scale adaptation] Would you rely on Furhat for important travel decisions?

- ☐ Very unlikely
- ☐ Unlikely
- ☐ Neutral
- ☐ Likely
- ☐ Very likely

Communication Mode Experience and Naturalness of Interaction

How would you rate your overall experience in communicating your travel requirements and navigating the travel planning process with Furhat?

- ☐ Very Easy
- ☐ Easy
- ☐ Neutral
- ☐ Difficult
- ☐ Very Difficult

How did the presence of visual cues in the multimodal interaction affect your experience?

- ☐ Significant positive impact
- ☐ Moderate positive impact
- ☐ Neutral / No particular impact
- ☐ Slight positive impact
- ☐ No positive impact at all

How did the lack of visual cues in the unimodal interaction affect your experience?

- ☐ Significant positive impact
- ☐ Moderate positive impact
- ☐ Neutral / No particular impact
- ☐ Slight positive impact
- ☐ No positive impact at all

Comparative and Reflective Questions

Compared to a human travel agent, how effective do you think Furhat is in assisting with travel planning?

- ☐ Much less effective
- ☐ Less effective
- ☐ About the same
- ☐ More effective
- ☐ Much more effective

Which mode of interaction did you find more effective?

- ☐ Strongly prefer uni-modal
- ☐ Prefer uni-modal
- ☐ No preference
- ☐ Prefer multi-modal
- ☐ Strongly prefer multi-modal

What is the primary reason for your preference?

- ☐ Better clarity and understanding
- ☐ More engaging experience
- ☐ Felt more natural or intuitive
- ☐ Easier to focus on the task

8.3 Informed Consent

Introduction

You are being invited to participate in a research study that investigates trust in human-robot interaction, focusing on communication strategies using the Furhat Robot integrated with ChatGPT. This consent form will provide information to help you decide whether to participate.

Purpose of the Study

This study aims to understand how different communication modes impact trust in interactions with a robot. Your participation will involve interacting with the Furhat robot in a simulated travel planning scenario.

Procedures

During the study, you will be asked to plan a trip with the Furhat robot. The session will last approximately 30 minutes.

Voluntary Participation

Your participation in this study is completely voluntary. You have the right to withdraw at any time without any negative consequences.

Confidentiality

All information collected during this study will remain confidential. Data will be stored securely and only accessible to the research team. Recordings and transcripts will be anonymized.

Risks and Benefits

There are no significant risks associated with participating in this study. While there are no direct benefits to you, your participation will contribute valuable information to the field of human-robot interaction.

Questions

If you have any questions or concerns about the study, please contact: rt2039@hw.ac.uk.

Consent

By signing below, you acknowledge that you have read and understood this consent form, and agree to participate in the research study.

Participant's Name: _____

Participant's Signature: _____

Date: _____

Bibliography

Amirova, Aida, Nazerke Rakhymbayeva, Elmira Yadollahi, Anara Sandygulova and Wafa Johal. 2021. 10 years of human-nao interaction research: A scoping review. *Frontiers in Robotics and AI* 8. 744526.

Cassell, Justine. 2000. Embodied conversational interface agents. *Communications of the ACM* 43(4). 70–78.

Chaturvedi, Rijul, Sanjeev Verma, Ronnie Das and Yogesh K Dwivedi. 2023. Social companionship with artificial intelligence: Recent trends and future avenues. *Technological Forecasting and Social Change* 193. 122634.

Chu, Charlene H, Simon Donato-Woodger, Shehroz S Khan, Rune Nyrup, Kathleen Leslie, Alexandra Lyn, Tianyu Shi, Andria Bianchi, Samira Abbasgholizadeh Rahimi and Amanda Grenier. 2023. Age-related bias and artificial intelligence: a scoping review. *Humanities and Social Sciences Communications* 10(1). 1–17.

Deichler, Anna, Siyang Wang, Simon Alexanderson and Jonas Beskow. 2023. Learning to generate pointing gestures in situated embodied conversational agents. *Frontiers in Robotics and AI* 10. 1110534.

Ferrara, Emilio. 2023. Fairness and bias in artificial intelligence: A brief survey of sources, impacts, and mitigation strategies. *arXiv preprint arXiv:2304.07683* .

Go, Eun and S Shyam Sundar. 2019. Humanizing chatbots: The effects of visual, identity and conversational cues on humanness perceptions. *Computers in Human Behavior* 97. 304–316.

Kasirzadeh, Atoosa and Iason Gabriel. 2023. In conversation with artificial intelligence:

aligning language models with human values. *Philosophy & Technology* 36(2). 1–24.

Kumar, Shailendra and Sanghamitra Choudhury. 2022. Gender and feminist considerations in artificial intelligence from a developing-world perspective, with india as a case study. *Humanities and Social Sciences Communications* 9(1). 1–9.

Lee, John D and Katrina A See. 2004. Trust in automation: Designing for appropriate reliance. *Human factors* 46(1). 50–80.

Li, Chen, Dimitris Chrysostomou and Hongji Yang. 2023. A speech-enabled virtual assistant for efficient human–robot interaction in industrial environments. *Journal of Systems and Software* 205. 111818.

Luger, Ewa and Abigail Sellen. 2016. " like having a really bad pa" the gulf between user expectation and experience of conversational agents. In *Proceedings of the 2016 chi conference on human factors in computing systems*, 5286–5297.

McTear, Michael Frederick, Zoraida Callejas and David Griol. 2016. *The conversational interface*, vol. 6, 94. Springer.

Min, Alfonso. 2023. Artificial intelligence and bias: Challenges, implications, and remedies. doi:10.13140/RG.2.2.36029.03049.

Oviatt, Sharon. 1999. Ten myths of multimodal interaction. *Communications of the ACM* 42(11). 74–81.

Pak, Richard, Nicole Fink, Margaux Price, Brock Bass and Lindsay Sturre. 2012. Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. *Ergonomics* 55(9). 1059–1072.

Pataranutaporn, Pat, Ruby Liu, Ed Finn and Pattie Maes. 2023. Influencing human–ai interaction by priming beliefs about ai can increase perceived trustworthiness, empathy and effectiveness. *Nature Machine Intelligence* 1–11.

Raaijmakers, Stephan, Anita Cremers, Emiel Krahmer and Matthijs Westera. 2023. Con-

versational ai. *Frontiers in Artificial Intelligence* 6. 1203910.

Seo, Kyoungwon, Joice Tang, Ido Roll, Sidney Fels and Dongwook Yoon. 2021. The impact of artificial intelligence on learner–instructor interaction in online learning. *International journal of educational technology in higher education* 18(1). 1–23.

Spezialetti, Matteo, Giuseppe Placidi and Silvia Rossi. 2020. Emotion recognition for human-robot interaction: Recent advances and future perspectives. *Frontiers in Robotics and AI* 145.

Su, Hang, Wen Qi, Jiahao Chen, Chenguang Yang, Juan Sandoval and Med Amine Laribi. 2023. Recent advancements in multimodal human–robot interaction. *Frontiers in Neurorobotics* 17. 1084000.

Vemprala, Sai, Rogerio Bonatti, Arthur Bucker and Ashish Kapoor. 2023. Chatgpt for robotics: Design principles and model abilities. *Microsoft Auton. Syst. Robot. Res* 2. 20.

Waytz, Adam, Kurt Gray, Nicholas Epley and Daniel M Wegner. 2010. Causes and consequences of mind perception. *Trends in cognitive sciences* 14(8). 383–388.

Wong, Carissa. ???? Ai'fairness' research held back by lack of diversity. *Nature* .

Wu, Wentao, Ben Zhang, Shuting Li and Hehai Liu. 2022. Exploring factors of the willingness to accept ai-assisted learning environments: An empirical investigation based on the utaut model and perceived risk theory. *Frontiers in Psychology* 13. 870777.

Xie, Bing, Xiangming Xi, Xinan Zhao, Yuhan Wang, Wei Song, Jianjun Gu and Shiqiang Zhu. 2023. Chatgpt for robotics: A new approach to human-robot interaction and task planning. In *International conference on intelligent robotics and applications*, 365–376. Springer.