Person Transfers Between Multiple Service Robots

Xiang Zhi Tan

CMU-RI-TR-22-01

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Robotics.

The Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

January 2022

Thesis Committee:

Aaron Steinfeld (Robotics Institute, Carnegie Mellon University)
Reid Simmons (Robotics Institute, Carnegie Mellon University)
Jodi Forlizzi (Human-Computer Interaction Institute, Carnegie Mellon University)
Selma Šabanović (School of Informatics and Computing, Indiana University Bloomington)

Abstract

As more service robots are deployed in the world, human-robot interaction will not be limited to one-to-one interactions between users and robots. Instead, users will likely have to interact with multiple robots, simultaneously or sequentially, throughout their day to receive services and complete different tasks. In this dissertation, I describe work, in collaboration with my colleagues, broadened the knowledge on a crucial aspect of multi-robot human interaction: person transfer, or the act of transferring users between multiple service robots. We first investigated rationales for transfer and important aspects of transferring users. We then explored how person transfers should be designed and implemented in laboratory and field settings. We used a combination of design, behavioral, and technical methods to understand the challenges and nuances in realizing person transfers. Our research consisted of (1) A collection of Research through Design workshops to chart out the space of person transfers; (2) A lab study to understand how people perceive social interaction between robots and the flow of information in a person transfer scenario; (3) A description of an interactive system that implemented realistic person transfers in both laboratory and field settings; (4) A lab study that evaluated different robot joining strategies and people's spatial behavioral responses during person transfers; and (5) An in-the-field demonstration of person transfers. Our work seeks to increase our understanding of this crucial phase and inform developers and designers about appropriate robot behaviors when a human is being transferred from one robot to another.

Acknowledgements

This dissertation is the culmination of inspiration, help, and heart from multiple people who helped me along the way.

All this would not have been possible without the help and guidance of my advisor, **Aaron**. Thank you for giving me the space and time to explore different topics and helping me find a research topic that we were both excited about. I always have been a person with a thousand half-baked ideas, and your guidance and support have helped me focus and concentrate on what's important and gave me important tools for my career ahead. Thank you.

To my committee members, thank you for being part of this journey. To **Jodi** for guiding me on the design research process and providing valuable insights that strengthen my work. To **Reid** for asking the hard questions that helped me crystallize and better form the rationale behind the choices I made in the dissertation. To **Selma** for listening to my ideas and dispensing valuable feedback.

To **Liz** Carter, thank you for letting me be part of the bystander intervention project that eventually led to so much more collaborations. Also, these chapters would not flow as smoothly without you.

To **Michal** Luria, thank you for showing me what design research is and for making the design workshops in this dissertation possible. Your research and process have always been an inspiration.

To everyone in TBD (Transportation, Bots, and Disability) Lab, past and present, these six and a half years would have been tougher without all the chats, games, social events, and help from you all – **Marynel** Vázquez, Lynn **Kirabo**, **Ceci** Morales, **Sarthak** Ahuja, **Abhijat** Biswas, **Lucile** Callebert, **Tesca** Fitzgerald, **Álvaro** Castro González, **Mary** Hatfalvi, **Oscar** J. Romero, **Kate** Shih, **Allan** Wang, and **Qian** Yang – thank you for being here.

To the current and former undergraduates that I was fortunate to work with – **Obi** Adubor, **Joe** Connolly, **Prithu** Pareek, **Fern** Limprayoon, **Amal** Nanavati, and **Clara** Wang – thank you for showing me the joy of mentoring and giving me the opportunity to be part of your experience here at CMU.

To **Reuben** Aronson, thank you for all the chats, discussion, and help leading the HRI reading group for the last six years. Also, for explaining to me all the math and engineering concepts I lacked as a computer science student.

To **Devin** Schwab, thank you for not only all the tips and tricks about everything computational but also being a wonderful housemate. All the boardgames and food definitely fuelled a large part of this journey.

To members of HARP Lab, RASL Lab, and CMU's HRI community, – **Henny** Admoni, **Maggie** Collier, **Roshni** Kaushik, **Pallavi** Koppol, **Nik** Martelaro, **Ben** Newman, **Mike** Lee, **Ada** Taylor, **Stephanie** Valencia-Valencia, **Michelle** Zhao – thank you for all the reading group discussion, hallway chats, and help along the way. All of these have worked their way into this dissertation.

To my office mates, – **Tejas** Mathai, **Haewon** Jeong, **Tabitha** Lee, and **Terri** Kent – thank you for being accommodating and supportive and for opening the office doors when I accidentally locked myself out. May Charizard look over you.

To others in my cohort and program – **Nick** Gisolfi, **Micah** Corah, **Rick** Goldstein, **Jingyan** Wang, **Xuning** Yang – all of the games, chats, PSRs, and lunches were priceless.

To **Sean** Andrist, **Dan** Bohus, and others who I was fortunate to work with at Microsoft Research on my internship, thank you for showing me how industry research works and the beautiful Pacific Northwest. The work we did got me thinking a lot about multi-modal situated interactions and inspired part of my work in this dissertation. Also, this dissertation would have taken longer if it has not been for \Psi .

To **Peggy** Martin and **Suzanne** Muth – thank you for helping me with all the administrative tasks and making this a smooth PhD journey.

To **Bilge**, **Allie**, **Dan**, **Chien-Ming**, **Irene**, **Steven**, and **Tomislav**: I would not have started this journey if I wasn't fortunate enough to be part of the HCI lab. Even today, my research has been influenced by what I learned from you all. Thank you for showing me the ropes.

To my family, thank you for instilling in me the sense of curiosity and adventure. I guess those broken computers in my childhood were worth it for a PhD.

To all my friends, no matter the distance or whether there are oceans that separate us, thank you for all the wonderful times together.

There are also countless others at Carnegie Mellon, Pittsburgh, at home, and all around the world who supported my journey and made an impact, however small, throughout my time as a PhD student.

Lastly, **Sam**, thank you for everything along the way. You not only answered the countless random research questions I had throughout the day but also gave me the confidence that maybe I am doing something interesting. Thank you for making sure I made time outside of research. Even if we both are children posing as PhD students, we both are doing great because we believe in each other.

Research projects reported in this dissertation were funded by National Science Foundation grants (CBET-1317989 & IIS-1552256) and National Institute on Disability, Independent Living, and Rehabilitation Research grant (90DPGE0003).

Contents

Contents v

1. Introduction 1
1.1. The Cambrian Explosion of Service Robots 1
1.1.1. Difference From Single Service Robot Human Interaction 1
1.1.2. Future of Multiple Robot Interaction 2
1.2. Our Research Agenda 3
1.3. Our Research Approach 4

Part 1: Background 7

2. Precursor Work 9
2.1. Categorizing Multi-Robot Human Interaction 9
2.1.1. Exploration of the Connectivity Spectrum 10
2.1.2. Summary 11
2.2. Relationship to Prior Work 11
2.2.1. Emergence of the Need to Transfer Between Multiple Robots 11
2.2.2. Importance of Spatial Positioning & Proxemics in Multi-Robot Systems 12
2.2.3. Nuances in Robot Behaviors 13
2.2.4. Bringing the Lab into the Real World 13
2.3. Conclusion 14

3. Related Work 15
3.1. Multiple Intelligent Systems 15
3.2. Multiple Robot Interaction in HRI Literature 16
3.2.1. Human Responses to Multiple Robots 16
3.2.2. Robot-to-Robot Communication in HRI 17
3.2.3. HRI Findings in Multi-Robot Applications 18
3.3. Spatial Formations & Proxemics 18
3.3.1. Proxemics 19
3.3.2. Spatial Formations 19

Part 2: Understanding Person Transfer 21

4. Design Space for Multiple Robots And Person Transfer 23
4.1. Overview 23
4.1.1. Chapter Summary 24
4.2. Series 1: Scoping Workshop 24
4.2.1. Overview 24
4.2.2. Method 24
4.2.3. Analysis 25
4.2.4. Findings 25
4.2.5. Discussion 30
4.3. Ideation Workshop 31
4.3.1. Overview 31

4.4. Method 31
4.4.1. Analysis 32
4.5. Findings 32
4.5.1. Overview of Findings 33
4.5.2. Types of Person Transfers 34
4.5.3. Rationale Behind Person Transfer 35
4.5.4. Design of Person Transfers 36
4.5.5. Information Shared 38
4.6. Utilizing the Taxonomy 39
4.6.1. Interaction Design 39
4.6.2. Researchers 40
4.6.3. In this Dissertation 41
4.7. Discussion 41
4.7.1. Transfer Trade-Offs 41
4.7.2. Complexity in Transfer Interaction Design 42
4.8. Limitations 42
4.9. Conclusion & Contributions 43

5. Inter-Robot Communication & Information Transmission In Person Transfer 45
5.1. Overview 45
5.1.1. Chapter Summary 46
5.2. Method 46
5.2.1. Study Design 46
5.2.2. Hypotheses 47
5.2.3. System and Study Setup 49
5.2.4. Procedure 49
5.2.5. Measures 51
5.2.6. Other Data 52
5.2.7. Participants 52
5.3. Results 53
5.3.1. Measure Reliability and Confounds 53
5.3.2. Perception of Mobile Robot 53
5.3.3. Perception of Stationary Robot 54
5.3.4. Robot-Robot Relationship 55
5.3.5. User Preference 55
5.3.6. Other Findings 56
5.4. Discussion 57
5.4.1. Hypotheses Support 57
5.4.2. Social Interaction Between Robots 58
5.5. Limitations 58
5.6. Conclusion & Contributions 59

Part 3: Realizing Person Transfer 61

6. Interactive System 63
6.1. Overview 63
6.1.1. Chapter Summary 64

6.2. Related Work 64
6.2.1. Building Interactive Systems in HRI 64
6.2.2. Our underlying frameworks 64
6.3. Overall Design 65
6.4. Sensory Layer 68
6.5. Execution Layer 69
6.5.1. Interaction Scene Generation 69
6.5.2. Behavior Machine 69
6.5.3. Behavior Machine in Practice 70
6.6. Behavior Layer 70
6.6.1. Coordinators 71
6.6.2. Actuators 71
6.7. Storage Layer & Analysis 72
6.7.1. Post-Processing 72
6.8. Discussion 73
6.9. Conclusion & Contributions 73

7. Spatial Formation in Person Transfers **75**
7.1. Overview 75
7.1.1. Chapter Summary 76
7.2. Robot Group-Joining Strategies 76
7.2.1. Operationalizing Algorithm 77
7.2.2. Execution & Trajectory 78
7.3. Method 78
7.3.1. Study Design 78
7.3.2. Service Scenarios 78
7.3.3. Hypotheses & Evaluation 80
7.3.4. Study Physical Environment 80
7.3.5. Interaction System 81
7.3.6. Study Procedure 81
7.3.7. Measures 82
7.3.8. State of Study Environment & Participants 83
7.4. Results 83
7.4.1. Validation 83
7.4.2. Subjective Responses 83
7.4.3. Participant Spatial Behaviors 85
7.4.4. Qualitative Responses 88
7.5. Exploratory Analysis 90
7.5.1. Spatial Formation and Social Cues (RQ1) 90
7.5.2. Effects of Physical Action (RQ2) 91
7.6. Discussion 91
7.6.1. Hypotheses Support 91
7.6.2. Strategy Effectiveness 92
7.6.3. Movement & Tolerance 92
7.6.4. Effects of the Environment, Embodiment, & Context 92
7.6.5. Variability in Complex Interaction System 93
7.7. Conclusion & Contribution 93

8. Person Transfers in the Field **95**
8.1. Overview . . . 95
8.1.1. Chapter Summary . . . 95
8.2. Method . . . 95
8.2.1. Study Environment & System . . . 96
8.2.2. Robot Joining Strategy . . . 96
8.2.3. Scenario . . . 97
8.2.4. Participants & Recruitment . . . 98
8.2.5. Technical Issues & Implications . . . 99
8.3. Findings & Discussion . . . 99
8.3.1. Overall Impressions . . . 99
8.3.2. Trust in Person Transfers . . . 99
8.3.3. Mobile Robot Handover . . . 100
8.3.4. Change in Group Membership . . . 100
8.3.5. Mobile Robot Joining Position . . . 101
8.3.6. Effects of Keyboard Inputs & Failures of Speech-to-Text . . . 102
8.3.7. Robot-Robot Communication . . . 102
8.3.8. Other People in the Scene . . . 102
8.4. Limitations & Future Work . . . 103
8.5. Conclusion & Contributions . . . 104

Part 4: Final Words **105**

9. Future Work & Real World Considerations **107**
9.1. Unanswered Questions from this Dissertation . . . 107
9.1.1. Composition of Robot Teams . . . 107
9.1.2. Domain of Transfer . . . 108
9.1.3. Scheduling Person Transfers . . . 108
9.2. Other Interesting Areas in Multi-Robot Interaction . . . 108
9.2.1. Person Transfers in a Multi Intelligent System Future . . . 109
9.3. Ethical Considerations . . . 109

10. Conclusion **111**
10.1. Research Contribution . . . 111
10.1.1. Design Findings . . . 111
10.1.2. Behavioral Findings . . . 111
10.1.3. Technical Findings . . . 112
10.2. Closing Remarks . . . 112

Appendices **115**

A. Questionnaires for Chapter 5 **117**

B. Multi Depth Camera Calibration **129**
B.1. Procedure . . . 129

C. Social Navigation System **131**
C.1. Human Costmap 131
C.2. Interaction Space Costmap 131
C.3. Limitations & Future Work 132

D. Behavior Machine **133**
D.1. State 133
D.2. Information Passing Between States 133
D.3. State Transitions & Machine 134
D.4. Interruption 134
D.5. Behavior Trees Pattern 134
D.6. Illustrative Example 135

E. Questionnaires for Chapter 7 **137**

Bibliography **151**

List of Figures

2.1. Spectrum of perceived connectivity in multi-robot systems 10

4.1. Some of the cards and tools used in the scoping workshop. 24
4.2. Overview of the findings from the scoping workshop. 26
4.3. Participants engaging in the Ideation phase of the workshop. 31
4.4. Overview of the taxonomy. 33

5.1. Stationary and mobile robots' dialogue for each condition. 48
5.2. Illustration of top down view of the study location. 49
5.3. Effect of Stationary Robot Behavior on Participants' Perception of Mobile Robot Warmth & Competence. 53
5.4. Effect of Stationary Robot Behavior on Participants' Perception of Mobile Robot Likability 54
5.5. Effect of Stationary Robot Behavior on Participants' Perception of Stationary Robot Warmth & Competence. 54
5.6. Effect of Stationary Robot Behavior on Participants' Perception of Stationary Robot Likability . . . 54
5.7. Effect of Stationary Robot Behavior and Information Transfer on Participants' Perception of Robot Relationship. 55
5.8. Participant preferences for Stationary Robot Behavior. 55

6.1. Overview of our system. 67
6.2. Screenshot of the textinput screen. 68

7.1. Example F-formation in Chapter 5. Left to right: (1) Participant faces the stationary robot, (2) Mobile robot arrives, (3) Stationary robot turns to talk to the mobile robot, (4) Participant shifts her orientation to ensure a shared space. 75
7.2. Illustration of the Circular strategy. 76
7.3. Illustration of the Line strategy. 76
7.4. Illustration of the Improper strategy. 77
7.5. Top down illustration of the study location. 80
7.6. Illustrations of the stages of the interaction. 81
7.7. Participants' responses to 4 subjective questions by scenario. 84
7.8. Participant responses to social appropriateness and whether the mobile robot position was "too far" by strategy. *** means $p < 0.001$. 85
7.9. Participants' aggregated distances to the stationary robot and mobile robot in the group interaction stage. The shaded region shows the range of values, the region inside the dotted line is the interquantile range, and the solid line is the mean. As the length of the interaction varied across participants, the shown aggregated value ends when half of the participants completed the stage. 86
7.10. Participant's distance to the stationary robot by stage for each robot joining strategy. *** means $p < 0.001$. 87
7.11. The distance ratio by stage for each robot joining strategy. *** means $p < 0.001$. 88
7.12. The distance to the centroid by stage and mobile robot joining strategy. *** means $p < 0.001$. 88

8.1. The scene of the study. 95
8.2. An illustration of the study scene layout 96
8.3. Illustration of the group circular strategy. 96

8.4. A screenshot from simulated human data. In the figure, three people (green) stood in front of Tank and were recognized to be in a group. The system first calculated the average midpoint (black) and generated a circle of candidate points around it (red). The system then picked the best point (blue) based on the criteria. The area inside the black outline is the valid region which accounted for the mobile robot footprint. 97
8.5. A participant reacting to the arrival of the mobile robot. 99
8.6. A comic strip showing participants leaving after Tank informed them it had run out of stickers and coming back after the mobile robot showed up. (1) A group of participants interacting with Tank. (2) The group walked away after Tank told them it ran out of stickers. (3) One of the participants (in blue) saw the mobile robot, pointed at it, and called for the group. (4 & 5) The group came back and surrounded the mobile robot to get the stickers. 99
8.7. A participant taking the stickers from the mobile robot as the mobile robot finished its movement. 100
8.8. A comic strip showing a participant (P4, blue backpack) joining halfway. From left to right: (1) P4 joined two other participants who were interacting with the robot. (2) The mobile robot was summoned and moved towards them. (3) P4 stepped back as the robot approached them. (4) P4 observed the interaction between the mobile robot and the other two participants. (5) P4 stepped further away as the interaction progressed. 101
8.9. An example showing a participant first attempting to use speech to interact with Tank (top) and eventually using the keyboard (bottom). 102
8.10. A comic strip of a group of people walking through the group interactions. From left to right: The participant (red shirt) observed the greeting between robots. A group of people simply walked between the robot and the participants. After the group passed, the participant approached the mobile robot to get the sticker. 103
8.11. The top down annotated point cloud view of the scene. We annotated the position of the participants (circle), the walls (white edges), the mobile robot (square), and the bystanders moving through (triangle). 103

9.1. Location of the disclosure signs in the study area. 109

B.1. The Charuco board used for calibration. 129
B.2. The combination of the pointclouds from all three Azure Kinect cameras in the hallway after calibration. 130

C.1. An illustration of the human ellipse model with our default parameter. In our studies, the value of a was set to 0.4. The ellipse ratio and ellipse offsets were configurable by RQT reconfig. 131

List of Tables

5.1. Participant demographics (36 valid sessions) 52
5.2. Coded perceived robots' relationship by Stationary Robot Behavior. 55

6.1. List of actuators for robotic platforms used in our studies. 72

7.1. Participant demographics (36 valid sessions) 83
7.2. Factor Loading. Items marked with (R) are reverse-coded. 85

Introduction | 1.

1.1 The Cambrian Explosion of Service Robots 1
1.2 Our Research Agenda . . . 3
1.3 Our Research Approach . . 4

1.1. The Cambrian Explosion of Service Robots

Today, robotic systems have transitioned from working behind the scenes in factories to providing services directly to users. While some robots today function as novelty tokens to attract customers [1, 2], more robots are starting to provide useful services that are truly beneficial to both service owners and customers [3]. Future interactions with robots are unlikely to be limited to only one robot at a time. Multiple robots could be deployed to simultaneously serve multiple users and accomplish different tasks based on their form factors and physical capabilities, or even provide redundancies to counter failures. They may also be owned by different services or users.

The evolution from a unicellular organism to a multi-cellular organism was a big leap in the evolution of life on Earth[1]. The synergy among multiple cells allowed for the emergence of the complex life that exists today. Similar to evolution, the cooperation and interactions among multiple robots opens up not only research questions about how they should work together, but also possibilities for new interactions and designs that were unimaginable before.

1: It took about 3 billion years for life to make that leap; hopefully, it will be shorter for robots.

As the field of Human-Robot Interaction (HRI) is still developing[2], many in the community are still focusing on the design and development of algorithms for single-robot-single-human interactions and attempting to understand their effects in diverse scenarios. However, as more researchers recognize the potential of group interactions, they have started to explore scenarios involving multiple entities. We have now started to see pioneering work in situations involving a single robot that interacts with groups of people [4, 5], one or more users managing multiple robots [6, 7], multiple robots interacting with multiple users [8, 9], and using multiple robots for interesting services and applications [10, 11]. Similar to the evolution of organisms, the increase in the number of robots has greatly increased the number of research questions that can be asked about the way the entities operate. We are now starting to ask how we should envision, design, and program these robots to not only work with users, but also to work well together in an ecosystem of multiple robots.

2: The main HRI conference started in 2006, which is relatively recent compared to ICRA (1984), NeurIPS (1987) and CHI (1982). This is not to say that no HRI research happened before 2006, but we are using the establishment of a dedicated high quality venue as a measure of maturity of the sub-field.

1.1.1. Difference From Single Service Robot Human Interaction

One might ask how adding more robots changes existing findings about one-to-one human-robot interactions. Prior work has shown the difference comes in three forms:

(1) Introduction of new forms of interaction The addition of another robot creates new ways for the robots to interact with users. Robots do not have the same limitations as humans and can do things that go beyond human capabilities. For instance, they have the ability to communicate vast amounts of information to each other in mere milliseconds and out of sight. This capability has been shown in prior work [12] to be undesirable as users found it creepy[3]. Another example of non-human capability is the appearance of a robot agent transferring from one body to another [13, 14] or a single agent simultaneously controlling multiple robot bodies.

3: Our work explored a similar concept in the context of robot information transfer and found similar results, further validating this work.

(2) Human behavioral changes due to an additional agent The inclusion of a second robot also transforms the dyadic interaction into a group interaction. Prior work has demonstrated how group interaction changes ways people interact and perceive the robots. Fraune et al. [15] showed how the number and appearance of robots changes whether people perceive the robots as threatening.

(3) Complexity in algorithms and implementation Besides change in human responses, the addition of another robot not only exponentially increases the state space of possible robot actions in existing algorithms, but also demands the creation of new systems to support new types of interactions. Sellner et al. [6] described a planning algorithm that reasons about the human's capability to determine when and which robot the human should teleoperate in a multi-robot construction setting. Khandelwal et al. [16] presented a formulation and search algorithm to determine the optimal actions and number of robots to guide people while minimizing service disruption. These questions and complexities are not present when users can only interact with a single agent in a single robot body.

1.1.2. Future of Multiple Robot Interaction

While it is hard to predict the advancement of robots, we can make some reasonable guesses about how they might evolve in the future. It is unlikely that we will see any general multipurpose robot in the near future ; instead, we will continue to see specialized robots that can complete just one or two tasks efficiently (e.g., Roomba). We are now seeing these robots integrating with smart voice assistants. It is not unimaginable that the trend will continue with deeper integration beyond simple commands to start a cleaning routine. Integration is also unlikely to stop with voice assistants, and soon will involve other robots that may or may not be owned by the same service. Even if we do successfully develop general multipurpose robots, these robots are unlikely to live in isolation as there are still both practical and social reasons to use more than one robot[4]. The eventual integration and collaboration among humans and multiple robots will undoubtedly be a part of the future of human-robot interaction.

4: In one of precursor work [14], participants expressed that they will be bored if they only interact with one agent throughout the day.

As people interact with multiple robots and digital systems, the group interaction and the formation of such groups does not have to be instantaneous. People can interact with one robot before the arrival of the second or subsequent robots. While the research community has examined both sides of this interaction – the one-to-one human-robot interaction at the beginning, and subsequent multi-robot human interaction at the ending – there has been sparse work examining this phase shift to group interaction. For example, a human user may interact with a kiosk robot that provides information about the venue and subsequently summons a mobile robot to guide the user to their destination. Upon the summoning, how should the robot join and interact with the other group members? This crucial phase of the interaction is the primary focus of this dissertation. While this appears simple, this dissertation will demonstrate the surprising deepness of this phase shift and how this work opens a vast new area of research.

1.2. Our Research Agenda

This dissertation examined the crucial phase shift between single robot human-robot interaction and multi-robot human-robot interaction as service robots transfer users between themselves. To further narrow the scope in this vast research area, we focused on the service context. We choose the service context since this is an area where robotic technologies are most likely to have an impact in the immediate future and sidesteps the questions about ownership and user performance [5].

5: These are important questions, as demonstrated by work from Luria et al. [17] .

In particular, this work addresses the following research question:

> How should **service robots** behave as they **transfer a human user from one robot to another**?

The question encapsulates the following crucial concepts:

Service Robots – This dissertation examines robots that are designed to provide services to users. Services are defined as interactions where the robot provides value to the user's experience at the location.

A Human User – We focus on a single individual to define and explore the nuances of the research question. While we defer extending to groups of people to later work, we also observed some group interactions in our field study.

Transfer – We focus on situations that begin with a one-to-one interaction, followed by the arrival of a second robot, and may include the eventual departure of one of the robots. There are four different phases in the interaction: "1-to-1 interaction", "Arrival", "Group", and "Departure".

One Robot To Another – We limited the interaction to two robots and a transfer of the user from robot A to robot B[6]. This setting reflects how existing human-to-human service transfers occur, and also simplifies the problem to its core components.

6: A transfer does not mean the first robot ceases to be part of the interaction. It simply means that the presence of the second robot is needed for the continuation of the service. Using a human service analogy, when a worker summons the manager, the arrival of the manager does not automatically lead to the worker disengaging and might even require both to complete the task.

Our goal is not only to understand the problem space, but also to explore how person transfer should be implemented and realized in the real world.

1.3. Our Research Approach

To answer this research question, we took an integrated and iterative approach that used a combination of design, behavioral, and technical methods. We used design methods to scope and probe the space where there is sparse prior research and generate novel ideas and insight in our work. We built systems and artifacts that create new opportunities to see how people interact with and react to our system and made technical contributions to realize such systems along the way. Lastly, we used behavioral research methods to examine how the behaviors of multiple robots could influence user behavior and perspectives on transfers. This all culminated in a simple yet meaningful implementation of a working person transfer system in the field. While simplistic, this is the first step in understanding the complex space of person transfers.

Here, we summarize the chapters of this dissertation, explain how they address our research questions, and note their key insights.

Part 1: Background

Chapter 2 (Precursor Work)
This chapter provides background material on the topic and situates our work in existing multi-robot human interaction literature. This chapter also provides case studies from our prior work on how existing one-to-one interactions could benefit from the use of multiple robots.

Chapter 3 (Related Work)
This chapter provides a literature review on prior work in multi-robot human interactions, spatial formation of robots, and how our work relates to existing work on transfers in Human-Computer Interaction and Human-Robot Interaction literature.

Part 2: Understanding Person Transfer

Chapter 4 (Design Space for Multiple Robots And Person Transfer)
This chapter describes our work in probing the space of multi-robot/agent service systems and exploring the context in which person transfers may happen in the future using Research through Design methods.

Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer)
This chapter describes a user study that examined how people perceived verbal information transfer between robots and how to frame this interaction between robots.

Part 3: Realizing Person Transfer

Chapter 6 (Interactive System)
This chapter details the underlying systems used in realizing person transfers in both laboratory and field settings.

Chapter 7 (Spatial Formation in Person Transfers)
This chapter addresses how a second robot should approach and join an existing human-robot interaction and how different strategies could change human behavior and perceptions.

Chapter 8 (Person Transfers in the Field)
This chapter reports on a small observational study of how people reacted and behaved when they experienced person transfers in the field.

Part 4: Final Words

Chapter 9 (Future Work & Real World Considerations)
This chapter lists the limitations of this dissertation and potential research questions worth exploring further.

Chapter 10 (Conclusion)
This chapter summarizes our work and contributions.

Part 1: Background

Precursor Work 2.

2.1. Categorizing Multi-Robot Human Interaction

2.1 Categorizing Multi-Robot Human Interaction 9
2.2 Relationship to Prior Work 11
2.3 Conclusion 14

We categorize multi-robot human interaction as any interaction where users interact with multiple different robots. The user could interact with either homogeneous (same type) or heterogeneous (different types) robots [18][1]. Yanco et al. [18] created a human-robot team taxonomy to describe 8 different variations of human-robot teams. While their work focused on the control in human-robot teams, it provided a framework that we can use to reason about the composition of multi-robot human interactions. Among the 8 different team types, the two most relevant configurations are *one human, robot team* and *one human, multiple robots*. In the *one human, robot team* group, the controller sends one command to all robots (robot team) and the robots figure out which robot is best suited to execute the action according to certain rules. This is different from *one human, multiple robots* teams, in which the user manually commands each robot.

1: There could be an additional distinction between (1) robots that are exactly the same type and (2) robots that appear visually the same, but different in internal components and capabilities.

Social interaction with a robot is different from the operation and control of multiple robots. The robots that interact with a human user could be simultaneously interacting with the user and communicating with other robots in the background to execute the task in a way that provides the best experience for the user. There is also a distinction between (1) the implementation and technical details about the robot team and (2) the perceived agency and relationships between the robots. Multiple physical robots with different personalities, functionalities, and demeanors could be working together while being controlled by the same program[2]. In contrast, a well-programmed, coordinated, and decentralized multi-robot system could give the impression that all of the robots are part of a single program when they are not. From the perspective of the user, the control scheme of the robot is masked by the appearance of the interaction. To reason about the desired behaviors for the robots, we need to focus on the user's perceptions of agency and interaction as well as their mental models of the robots.

2: This is called a "centralized" control scheme in the multi-robot literature.

To help reason about how users perceive multiple robots, we envision these systems as existing on a spectrum of perceived connectivity that is anchored by two mental models. On one side is *Isolated Units*. This is the mental model where each robot is unique, and all of the data is isolated in each robot's own physical embodiment. On the other end of the spectrum is *One for all* [14]. This is the mental model wherein all of the physical embodiments are part of one bigger unit. In this model, users would assume that each embodiment that is part of the larger unit is controlled by the same agent[3].

3: This is also known as the "hive mind" in science fiction. Examples are "The Borg" in *Star Trek* where Borg drones are connected to each other and "HAL" in *2001 Space Odyssey*, where one intelligence controls all devices on the spaceship.

Along the spectrum, we have the following modes:

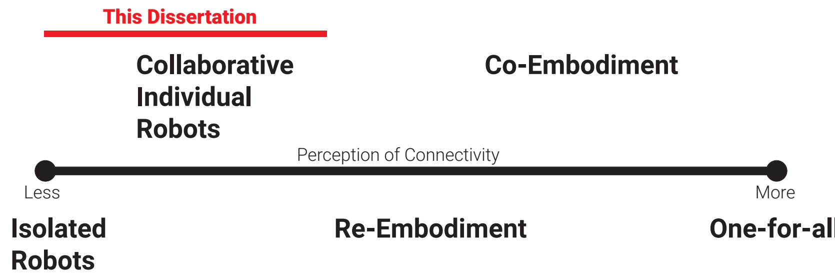


Figure 2.1.: Spectrum of perceived connectivity in multi-robot systems

1. *Isolated Units* – These are decentralized robots where the only way for the robots to communicate to each other is through the same medium that humans use (speech, vision, etc.).
2. *Collaborative Individual Units* – These are robots that, while they appear decentralized, still maintain connections that allow for unseen communication and coordination.
3. *Re-Embodiment* – This is the scenario where the robot's intelligence can transfer between multiple different robot bodies.
4. *Co-Embodiment* – Here, multiple agents live in the same robot body [14].
5. *One-for-All* – The user perceives a single entity controlling multiple different robot bodies.

All of these different paradigms come with their own research questions, such as whether and how people will perceive these systems and how we can implement them technologically. We have conducted some initial explorations into this space.

2.1.1. Exploration of the Connectivity Spectrum

To understand how people think about different forms of multi-robot and cross-device systems, we designed a User Enactments study where participants experienced different low fidelity multi-robot scenarios and reflected on their experiences[4]. In the study, we designed four different scenarios where participants experienced multi-robot scenarios that simulated systems at different points along the spectrum in a wide variety of domains. The participants experienced both Isolated Units and Re-Embodiment in a DMV scenario, Re-Embodiment and One-for-All in a dinner party planning scenario, Re-Embodiment again (this time through physical token that a user used to control agent transfer) in a healthcare scenario, and, finally, Co-Embodiment in an autonomous car scenario[5]. The settings were all low fidelity mock-ups created using props. The robots' verbal behaviors were controlled by a hidden experimenter through a Wizard-of-Oz methodology. The robots were physically moved around during each scenario by an experimenter. In the study, we recruited 18 participants who were more than 25 years old. They experienced these scenarios in different orders. At the end of each scenario, participants were interviewed about their thoughts on their experiences.

4: This work, published in [14], was led by Michal Luria, with contributions by Samantha Reig, myself, Aaron Steinfeld, Jodi Forlizzi, and John Zimmerman. I took part in the ideation, execution, analysis and write up of the study.

5: For more details about the methods and scenarios, please refer to the paper [14].

Here, we would like to highlight some important results and their relationship to this dissertation. Participants were accepting of re-embodiment and did not find it weird. One participant drew an analogy to how Siri could be activated in both the watch and the phone, which they gave as an example of re-embodiment in real life[6]. Participants responded negatively to re-embodiment in cases where there was a perceived need for expertise. Some participants did not want the robot who guided them in the hospital to be the same as the robot who completed a medical exam. This suggests that there are cases where *Collaborative Individual Units* are required to present different robots with different capabilities and expertise. For example, instead of a robot guiding a patient and also performing a medical examination, the better design would be for a guide robot to transfer a patient to another robot for a medical examination. In addition to those considerations, some participants also wanted to have separate agents for different contexts to create clear boundaries between different domains (for example, work and home).

6: The example given by the participant was probably more of a *One-for-All* example as one could activate Siri simultaneously on both the phone and the watch.

Another interesting finding from the study was that in the co-embodiment scenarios, participants found the verbal conversation between agents to be unnecessary[7]. Some participants felt that the agents were ignoring them and that such conversations should be done behind the scenes. This finding differs from our results presented in Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer) and prior work [12] where people reacted negatively to silent communication between robots. One possible hypothesis is that this difference is due to the content of the conversations.

7: The agents were conversing between themselves about what kind of windscreen wiper the home agent should purchase after receiving confirmation from the participant to purchase a replacement windscreen wiper.

2.1.2. Summary

This dissertation focuses on scenarios where the robots are *Collaborative Individual Units* and present themselves to users as separate entities[8]. We wanted to further understand how two collaborative individual robots should interact with each other while they provide a service to users.

8: An interesting, unanswered question is, "Where on the spectrum do people currently perceive two robots that work together?"

2.2. Relationship to Prior Work

Our prior work during the last few years has contributed to the different ideas in this dissertation.

2.2.1. Emergence of the Need to Transfer Between Multiple Robots

One of our earliest studies explored how a deployed stationary robot in a public place could assist people who are blind[9]. Robots, both stationary and mobile, have been identified as useful tools to help people who are blind navigate complex indoor spaces [20]. We developed methods for the stationary robot that were inspired by a technique used by Orientation &

9: This work, published in [19], was done in collaboration with Elizabeth Carter, Samantha Reig, and Aaron Steinfeld.

Mobility Experts where they draw paths on a person's palm while giving verbal navigation instructions. While our user study participants who were blind responded positively to our system, they pointed out that because our robot is stationary, an initial challenge would be to locate it before being able to receive the directions. Participants discussed ways that they envisioned finding the robot: they could be guided by sounds or transferred to it by other people or even another robot. One could imagine a scenario where a self-driving car drops off a user and directs them to a stationary robot to continue their journey.

In addition to providing route instructions, we have explored other robotic technologies to assist blind people in navigation. In [21], we developed a method for a handheld commercial spherical robot to generate kinesthetic haptic feedback which we used to provide instructions to users. In [22], we created a geometrical model to estimate a coupled user (the user was holding onto the robot) position while the robots is moving. Each of these platforms has different capabilities and its own set of limitations. For example, both the spherical robot and mobile robot have no verbal or audio capabilities. This characteristic highlights the value of collaborative, heterogeneous robots working together. To fully take advantage of these robots for different tasks, we need to ensure smooth transfers between them. To this end, we have also ventured beyond robots and explored how mobile robot should rendezvous with a blind user after being summoned using a smartphone [23].

2.2.2. Importance of Spatial Positioning & Proxemics in Multi-Robot Systems

In [22], we developed Markovian models for how people holding on to a mobile guide robot would move and shift locations as the robot moves[10]. The deterministic models were created based on the observations that people always tried to move as little as possible and slowly converged back to the center as they followed the robot. While our models outperformed a simplistic model that assumed that the person always stood behind the robot, our models did not account for obstacles or the presence of other robots.

10: This work [22] was led by Amal Nanavanti with contributions from me, Joe Connolly and Aaron Steinfeld. I helped mentor Amal and Joe on this project and assisted in the simulation and data analysis.

The behaviors of the robots should also depend on the spatial distance to reference objects. Prior work by Sauppé et al. [24] demonstrated that the clarity of a robot's gesture is affected by the distance to the referred object. We have explored this space and built a prototype interactive system that reasons about the distance of objects and user attention and selects the most suitable behavior to redirect user attention[11] [25].

11: This work was published in [25] and was done in collaboration with Sean Andrist, Dan Bohus, and Eric Horvitz while I was an intern at Microsoft Research.

The ability to reason about spatial formations will be important in HRI applications [5]. Adding additional robots to an existing spatial formation is different from adding an inanimate object because people will shift their position in response to a moving robot. It also adds another point of control for designers/developers to change spatial dynamics. It is crucial for effective HRI for us to better understand the spatial dynamics of robots in multi-robot human interaction scenarios.

2.2.3. Nuances in Robot Behaviors

Depending on the scenario, the behavior of a robot can change how humans interpret its intention. In our work[12], we sought to understand how robots can protect themselves from "abuse"[13] in the field, which was shown to be a real concern in earlier studies [27], news reports [28], and our own field study (Subsection 8.3.8 (Other People in the Scene)). Drawing from knowledge about humans in groups, we explored how a robot could use different behaviors to induce human bystander intervention during a human's "abuse" of a robot. Our hypothesis was that strong reactions from the robot, such as angry and sad behaviors, would induce more frequent responses. Instead, we found that an indirect behavior, such as shutting down, led to a stronger reaction from the participants. In the post-study interview, participants mentioned that they interpreted the strong reaction from the robot to be indicative of it "playing" with the abuser (a confederate in our study).

12: This work, published in [26], was done in collaboration with Marynel Vázquez, Elizabeth Carter, Cecilia Morales, and Aaron Steinfeld.

13: We used the word "abuse" to describe a wide range of behaviors from curiously testing the robot's limits all the way to malicious destruction.

In other related work, we investigated what kinds of robot behaviors influence user perceptions and actions after failure[14]. We were interested in how different types of failures (personal vs. property damage), levels of severity (low vs. high), order of mistakes, and presence of social features (face vs. no face) influenced trust in the robot as well as whether people would help and trust a robot after observing failures. While we did not find that the presence of social features influenced the likelihood of people assisting the robot, some of the 32 participants who saw the robot's face reported confusion about the robot's intent, and they were not sure if the robot was angry at them or felt sad because it made a mistake.

14: This work [29] was led by Cecilia Morales in collaboration with me, Elizabeth Carter and Aaron Steinfeld. I assisted in the development of the system and analysis of the data.

These works highlight the importance of well-designed robot behaviors. We hope to capture people's priorities and preferences during person transfers through a combination of design workshops, lab studies, and in-the-field observations to inform designers and developers about what robot behaviors are most appropriate in these scenarios.

2.2.4. Bringing the Lab into the Real World

All of our precursor work was done in a controlled lab environment. This was problematic for some studies. For example, some participants in the bystander intervention study [26] reported that they chose not to act because they knew it was a study and expected the experimenter to intervene. Furthermore, real-world studies are seen as more valuable in the field of HRI because they can uncover aspects of interactions that are missed in controlled studies. As part of this dissertation, we also bought part of our person transfer system into the field (Chapter 8 (Person Transfers in the Field)).

2.3. Conclusion

Our precursor work highlights the benefits of each robotic system individually, but these do not live in isolation. The HRI community and literature will benefit from explorations of how people might transfer between these robots. Situating the robots in real world scenarios can also help uncover issues we missed.

Related Work 3.

3.1 Multiple Intelligent Systems 15
3.2 Multiple Robot Interaction in HRI Literature 16
3.3 Spatial Formations & Proxemics 18

In this chapter, we first present a brief review of prior research on cross-device systems and how that work transfers to robots working with other intelligent systems. We then examine different facets of multi-robot human interaction. Lastly, we talk about proxemics, spatial configurations in multiparty interactions, and their relevant applications in multi-robot systems.

3.1. Human Interaction with Multiple Intelligent Systems

Since the advent of personal computers and, more recently, the arrival of smartphones and tablets, researchers have looked into how to enable interactions across devices and their surroundings [30]. Brudy et al. [31] developed a taxonomy to classify cross-device interactions into 6 domains: temporal, configuration, relationship, scale, dynamics, and space. We can categorize the transfer between service robots using the same taxonomy. Our interaction is primarily a sequential, asynchronous interaction (temporal domain) where the user interacts with one robot followed by the other. We focus on interactions involving 1 person and 2 robots (relationship domain) in a collocated environment (space domain)[1]. An similar example of cross-device interaction is work by Ghiani et al. [32] which describes a system that transfers part of a web page from a screen to a mobile device for continuous viewing.

1: As discussed in Chapter 4 (Design Space for Multiple Robots And Person Transfer), there are some variation in the space and temporal domains.

Throughout the evolution of computing, a large number of ethnographic reports have examined how people use multiple devices. Dearman et al. [33] explored how professionals balance the use of personal and business devices and found it was hard for users to clearly separate the intended use of the devices. Santosa et al. [34] conducted a field study on how people use multiple devices with the advent of smartphone, tablets, and cloud applications. At that time, they noted the need for better awareness of each device and easier ways to transfer information from one device to another. Researchers have also looked into newer cross-device technologies, such as Continuity by Apple, where users can start a task on one device before completing it on another. Raptis et al. [35] collected online responses to Apple's Continuity feature and identified pain points in the system. They highlighted the problem of privacy when data from one device pops up on others and noted that users would like additional control over what happens and what is shown on their devices.

However, robots[2] are fundamentally different from the computational devices described above: they are physically embodied, have more autonomy, are able to act on the physical world, and provoke different emotional responses [36]. While cross-device interaction provides some insights on

2: How a robot is defined remains a debate. In this dissertation, we define a robot as a computer with decision-making processes that can manipulate physical objects in the world. This definition is not perfect because it could include very smart dishwashers.

user behavior across multiple physical systems, we people are likely to act differently when interacting with robots.

People already interact simultaneously with robots and other devices in many different kinds of settings. Teleoperation of robots often involves interacting with the robots through another device, such as a screen [37–40], controllers [41, 42] or even virtual reality [43–45]. These digital devices have also been used to share knowledge or additional information about the interaction. Walker et al. [46] demonstrated an augmented reality system that informs users about what actions a flying robot will take next in a task scenario. There also has been a long history of research in which robots work together with other intelligent systems. Scassellati et al. [47] showed how an embodied robot's gaze could be used to guide deaf infants' gaze to a screen with a virtual agent.

3.2. Multiple Robot Interaction in HRI Literature

3.2.1. Human Responses to Multiple Robots

People's reactions to and perceptions of robots changes when they interact with multiple robots compared to a single robot. Prior work on human perception of robots has explored how the robot group composition changes human perception. Fraune et al. [48] explored how appearance of the robots (anthropomorphic, zoomorphic, and mechanomorphic) and number of robots affect the humans' perceptions of robots in videos. Participants perceived a group of mechanomorphic robots more negatively but a group of anthropomorphic robots more positively compared to other types. The lead author conducted a similar experiment with an in-person study in both Japan and USA to see whether the culture of the participants influenced their perception of robot behavior and number of robots [8]. In their second study, they found opposite results from their prior work. Their results showed that users responded more positively to a single social robot and a group of functional robots than to a group of social robots and a single functional robot. The authors attribute the differences to the familiarity of robots at the study country and the fact that users may have been responding to the human-likeness of the behavior instead of the appearance. When combining different types of robots, Fraune et al. [15] found that robots were perceived more negatively when presented in a group of identical robots than when alone or in a diverse group of robots.

Besides changing perceptions of the robots, the use of multiple robots can also influence human behavior. Salomons et al. [9] showed that people may follow the opinions of a group of robots in ambiguous tasks due to their trust in the robots' opinion. This study was a variation of Asch's conformity experiment [49] where 36.8% of participants conformed to a factually incorrect answer to a question after seeing several other people (confederates) give that answer. Other prior research was unable to directly replicate the conformity experiment with robots [50, 51]. Other work has explored how the number of robots and participants' genders influence conformity in ambiguous tasks [52].

Connolly et al. [53] described a study that explored whether a sad reaction by a robot in response to a confederate abusing another robot led to more prosocial action by human observers. They found participants were more likely to intervene when the second robot expressed a sad reaction to the abuse.

Admoni et al. [54] found that participants were less accurate in determining which robot was looking at them among random movements when they had to keep track of multiple robots. Oliveira et al. [55] observed that when interacting simultaneously with two robots, participants gazed more often at a cooperative robot than at a competitive robot.

Fraune et al. [56] explored how a single person and groups of people reacted to and interacted with different numbers of robots in a competitive game. They found that participants' reactions to robots differed from what is expected according to human social psychology literature, where individuals were more fearful of and had strong negative emotions when interacting with a group and that groups of humans will be more greedy towards individuals. Participants' emotion, greed and competitiveness only increased when interacting with the same number of robots.

Reig et al. [57] investigated how different robot behaviors impacted trust for an operator after a robot experienced a failure. Among the strategies they tested were *call* (the failed robot called a second robot for help) and *sense* (a second robot arrived after sensing the first robot failed). They found that people preferred for a robot to announce that it had updated itself and claim that it would not encounter the same failure over summoning another robot.

3.2.2. Robot-to-Robot Communication in HRI

An early exploration of human perception of multi-robot interaction investigated how people interpreted a conversation between two robots [58]. Kanda et al. [58] described a system where the robot communicated with another robot using verbal and non-verbal behaviors to not only provide context for the human observer about the information transfer between robots, but also to demonstrate the robots' communication capabilities. They showed that after observing from afar, participants were able to infer the robots' verbal and nonverbal communicative capabilities. They also responded to a robot's nonverbal behavior by following its pointing gesture.[3] Later, Hayashi et al. [60] expanded upon the idea of human users inferring content from a conversation and came up with the idea of "Robot-Manzai" [4]. In this scenario, the two robots acted as passive social media where they communicated with each other in front of bystanders with the goal of conveying information to bystanders. They showed success in using "Robot-Manzai" to acquire bystander attention in science museums [61] and train station [62] scenarios.

3: A similar study [59] showed that a conversation using emotional expressions by two robots could lead to an observer being able to understand the context of the interaction and help the robots.

4: Manzai is a type of Japanese stand-up comedy where two characters exchange jokes and commentaries.

Another application for robot-to-robot communication is to smooth conversation and reconcile information when encountering recognition errors. Iio et al. [63] developed a two-robot question-and-answer dialog structure where one robot asked questions and the other robot replied with the answer that the user provided. The dialog was designed such that it sounded like mimicking when the recognition was successful, and like the robot expressing a different opinion if the interpretation of the user input was incorrect. The research found that participants perceived the conversation to be more coherent and the robots more friendly when they utilized this strategy.

Prior work has also looked at how two robots should communicate with each other. Fraune et al. [64] examined how different types of inter-robot communication (none, loud, and silent) between basic functional robots affected the attitude of a bystander. The study did not find any significant differences across the conditions. The authors hypothesized that the result arose from participants assigning *groupness* to the robots rather than treating them as individual social entities. Williams et al. [12] explored how different robot-to-robot communication affected user perceptions. In a simulated nuclear disaster scenario, 56 participants issued commands for two different robots in a search task. The robots communicated the user commands and their results either verbally aloud or covertly to each other. Participants described the covert communication between the robots as creepy.

Since the publication of Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer), other works have built upon our work and explored other aspects of the space. An online video study demonstrated even a one-way expression of warmth in robot-to-robot communication increased perceived warmth of all involved robots [65].

3.2.3. HRI Findings in Multi-Robot Applications

Previous research has explored applications of multiple-robot systems. For example, Shiomi et al. [66] described a field study where two networked robots roamed a shopping mall and performed a variety of tasks. Among other skills, the robots had a coordinated activity where one robot would lead the guest to another robot that then welcomed the guest to a store. In Leite et al. [67] , two Keepon robots interacted with either a single child or a group of children during storytelling tasks where the robots played different roles from the story. Guinness et al. [11] created a system where multiple small tangible robots, together with a touchscreen, created a changeable tangible tactile interface to display information for people with visual impairments. Vázquez et al. [68] designed a robot ("Chester") with a sidekick robot ("Blink") attached to it. The authors found the addition of the sidekick robot increased children's attention without changing the proxemics and spatial formation with the robot.

3.3. Spatial Formations & Proxemics

The physical positions and orientations of individuals encode and inform information about the interactions and relationships between interactors.

3.3.1. Proxemics

One of the first spatial properties to be studied was the distance between interactors, i.e., proxemics. In his seminal work, *The Hidden Dimension* ([69]), Hall categorized how people stand relative to each other in four categories based on distance: "Intimate", "Personal", "Social" and "Public" (page 108). These distances mirror the boundaries of human sensory inputs (e.g., interactors can touch each other in the intimate space and clearly see each other's faces and in the personal space). As noted by Hall and other related work, the boundaries of these zones and actual distances in practice can be influenced by other factors such as culture [70] and environment.

Prior HRI research has shown that people observe a similar distance pattern when interacting with a social robot. Vázquez et al. [68] found that children interacting with a robot also observe the different categorizes of distances, but with only three categories where "Intimate" and "Personal" are combined. As with human groups, these distances can be influenced by multiple factors. For example, Walters et al. [71] found that participants' "proactiveness" decreased their social distance from the robot. Additionally, Walters et al. [72] observed that participants maintained greater distances from a robot that used a machine-like synthesized voice compared to ones that used more natural voices. Takayama et al. [73] conducted a study to understand how participants' familiarity with robots and where the robot looked at a participant changed their comfort with how close they were to the robot. They found that female participants had a higher minimum comfortable distance when the robots looked at them but the opposite for male participants. Mumm et al. [74] found that people maintained a large distance from a robot if they disliked the robot and the robot gazed at the person.

3.3.2. Spatial Formations

Beyond observing the occurrences of these distance selections, HRI researchers have also attempted to describe and influence physical group formations to improve interaction quality. Humans form an F-formation when in a group, a term coined by Kendon [75] , where in humans leave different spaces (O-space) between each other to facilitate the group activity (e.g., space to see others in a conversation). Kendon [75] and others (e.g., [76]) also observed common spatial arrangements and patterns that people maintained, such as circular, vis-a-vis (face-to-face), and rectangle (multiple people facing one person). Kuzuoka et al. [77] demonstrated that robot physical behaviors can change a group's F-formation, and that changes in robot body orientation were more effective than changes in robot head movement. Work has also been done on robot group reasoning. Vázquez et al. [78] conducted a simulated study to find the optimal policy for a robot to orient its body in an F-formation. Bohus et al. [79] created a heuristic system that categorized different spatial layouts of people who interacted with robots and attempted to modify people's positions through dialog.

Spatial Formation & Proxemics for multiple robots

As we add more robots into human-robot interactions, the spatial formations of the systems change. Joining robots could also influence where the first robot should stand and the users' perceptions of the relationship between the robots. Some preliminary work has been done in this area. Matsumoto et al. [80] found that putting a wide distance between two conversational tabletop robots when interacting with elderly participants increased their mental load. Also, Vázquez et al. [68] found the addition of an interactive sidekick robot on a robot did not change the proxemics profile of the children interacting with the robot.

Part 2: Understanding Person Transfer

Design Space for Multiple Robots And Person Transfer 4.

4.1 Overview 23
4.2 Series 1: Scoping Workshop 24
4.3 Ideation Workshop 31
4.4 Method 31
4.5 Findings 32
4.6 Utilizing the Taxonomy . 39
4.7 Discussion 41
4.8 Limitations 42
4.9 Conclusion & Contributions 43

4.1. Overview

Research in this chapter was conducted in collaboration with Michal Luria, Jodi Forlizzi, and Aaron Steinfeld.

Parts of this chapter are adapted from the following publications:
Xiang Zhi Tan, Michal Luria, Aaron Steinfeld, and Jodi Forlizzi. 'Charting Sequential Person Transfers Between Devices, Agents, and Robots'. In: Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction. *2021*

Xiang Zhi Tan, Michal Luria, and Aaron Steinfeld. 'Defining Transfers Between Multiple Service Robots'. In: Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction. *HRI '20. Cambridge, United Kingdom: Association for Computing Machinery, 2020. ISBN: 9781450370578*

As pointed out in the Chapter 3 (Related Work), adding an additional robot into human-robot interaction fundamentally changes the way people interact with robots. Issues such as the composition of human-robot teams, what jobs and abilities robots have, and the modalities robots use all play a role in how people perceive these interactions. As one step in our understanding, we wanted to learn what kinds of scenarios person transfer between service robots might occur, explore potential variations on the concept of person transfer, how they might be situated in the larger service context, and identify important factors in person transfer.

As these interactions are mostly set in the future, we lacked an understanding of how to design them and the social norms around them. Design methods are well suited to tackle this gap: they allow us to extrapolate knowledge from existing interactions and explore ways in which novel interactions may play out in the future. The use of design methodology to generate knowledge is commonly called "Research through Design". Research through Design [83] is a research philosophy that uses design techniques to create the "right" artifact[1] to transform the present to a preferred future. In other words, it uses design methods to generate knowledge about potential future systems. Our goal is exploratory in the sense that we are trying to envision preferred futures rather than trying to find a specific use case. We choose to use design methods as ways to probe these spaces to better understand how different people see the future of multiple robot interactions.

1: According to Zimmerman et al. [83], artifacts can be anything from models and guidelines to products and technologies.

Design research has been used in the past to better understand different aspect of HRI. For example, Rule et al. [7] used design methods to identify crucial factors when designing teleoperation interfaces for multi-user, multi-robot systems. Participatory design has been used to explore how robots can help older adults [84] and people with visual impairments [20].

4.1.1. Chapter Summary

In this chapter, we present two design workshops that probed the design space of how we might envision the future of multiple robots, and what we should pay attention to. We first conducted a *scoping* workshop to explore the space of human and multiple robot interactions. This workshop identified ways in which people might interact with multiple robots. We highlighted person transfer to be one of the three types of sequential interactions that people can encounter. We then followed with an *ideation* workshop that focused on exploring sequential person transfers through the lens of existing person transfers between robots, devices, and human staff. This workshop yielded details on the components of person transfers. We then combined findings from the two workshops and identified emerging themes to construct a 4-dimensional taxonomy that breaks down sequential person transfers and highlights research opportunities for the future within this space. We end this chapter by discussing our findings, how the taxonomy can be applied, and its value to interaction designers and researchers.

4.2. Series 1: Scoping Workshop

4.2.1. Overview

Our first step was to understand the possible roles of robots in sequential interactions. We conducted a scoping workshop that was structured around brainstorming future human and multiple robot interactions. For this workshop, we recruited people who were familiar with robots and cutting edge technologies. Participants were told to generate as many descriptions of transfer scenarios as possible.

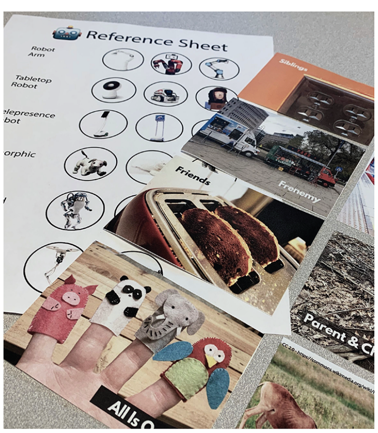


Figure 4.1.: Some of the cards and tools used in the scoping workshop.

4.2.2. Method

The scoping workshop consisted of three separate events with a total of 8 different participants and was conducted in September and October, 2019. Each event had 2 or 3 participants (excluding the facilitator) and took about 2 hours. All events were conducted in a meeting room at Carnegie Mellon University's Pittsburgh Campus. Participants were recruited through email lists and word of mouth, and they were compensated 20 USD for their participation. Participants were graduate students who studied robotics or human-computer interaction. We chose graduate students as they have high familiarity with current robot and device capabilities. The workshop was approved by our university's Institutional Review Board.

After warm-up activities with Alternative Use and New Metaphors [85] exercises, participants were asked to generate ideas about how they envision multiple robots working together to provide services to people in the future. We chose not to limit ideation to sequential interactions in order to capture the natural occurrence of sequential interactions within the broader pool of ideas. Each event consisted of multiple idea generation rounds. In each round, participants were given a specific location or context and provided with ideation cards. These were printed cards with words or images. Participants were then given a few minutes to write down as many ideas as possible on a worksheet.

To maximize the number of unique ideas and improve idea generation, we modified the procedures and tools between each event. In events 1 and 2, participants were given cards that specified the types of robots involved in an interaction, such as humanoid robot, flying robot, etc. Theses cards were removed in event 3. For events 2 and 3, participants were also given cards with prescribed relationship, e.g., sidekicks, swarms, or friends, to inspire new forms of interactions. After each round, participants shared and discussed some of their favorite ideas. We also wrote down new ideas that arose during the discussion phase.

4.2.3. Analysis

All the data were transcribed into digital form and grammatical errors were corrected. We then proceeded to analyze the data through Affinity Diagramming [86]. We extracted the different themes and ideas from the data and sorted them based on their role in the environment to understand how our participants imagined multi-robot systems.

4.2.4. Findings

Participants generated a total of 132 ideas. As we chose not to correct or limit workshop participants in their ideation of multi-robot interactions, 17 ideas were invalid (e.g., only have one robot, unparseable). Furthermore, the majority of the ideas (59 ideas) involving multiple robots fell into the category of *simultaneous* multi-robot interaction (e.g., "*A physical service robot helps you go to [a] doctor's office while [another] robot carries your stuff.*") or interactions where the two robots can be perceived as part of the same entity (e.g., "*[A robot] has a robotic arm attached [to it] to guide visitors & point to exhibits while giving a tour.*"). We categorized 28 of the ideas as *sequential* interactions. The imbalance in categories reflects the difficulty of envisioning interactions across time, and the current tendency to conceptualize multiple robot applications as swarms or simultaneous interactions.

In the following paragraphs, we briefly cover the types of multi-robot interactions in service interactions. All our scenarios involved the user being in a **multiple robot environment**. However, this did not mean that the person receiving the service would necessarily directly interact with multiple robots. We then separated and categorized scenarios where people

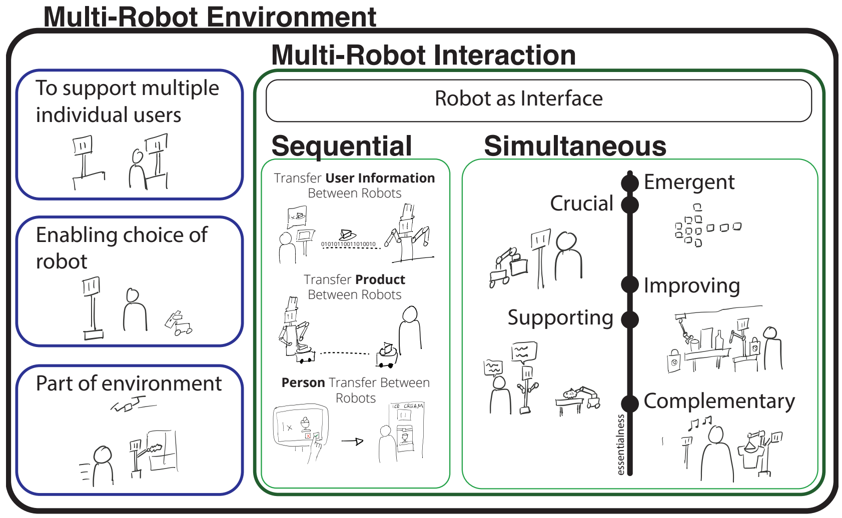


Figure 4.2.: Overview of the findings from the scoping workshop.

directly interacted with multiple robots (**multi-robot interaction**). In *multi-robot interaction* scenarios, we found three main categories. An overview of our findings is shown in Figure 4.2.

Multi-Robot Environment

The prerequisite of a person transfer is that it has to be an environment with multiple robots.

To support multiple individual users
One of the main use cases for multiple robots is to simultaneously provide the same experience to multiple different people. In these cases, the user interaction with multiple robots might be minimal and passive. Users of the service might see other robots but they only interact with one. For example,

> *"[Embodied Agent] in all tables in the restaurant and customers enter their orders."* - P4

In the example above, each table has its own embodied agent that serves the customers, and the people at each table presumably do not interact with other agents at other tables. Others imagined how an interaction with a single robot could contribute to a bigger ecosystem of robots. For instance,

> *"Hospital bed as robot: They measure sleep quality of patient, have a leader board to compete, also self-report patient data to show comfort"* - P6

This participant imagined linking individual hospital robots to allow hospital patients to compare how they are doing with other patients through the robots. The robots allow users to indirectly interact with other users through them.

Part of the Environment
Participants also generated multiple scenarios where the role of the robots was to be part of the environment, whether it was to create ambiance for

the environment or complete backstage tasks for the service. One example was to make an airport more entertaining,

> *"Airports are boring, fill it with realistic zoomorphic robots to make it more entertaining."* - P3

Others focused on using multiple robots to complete background tasks.

> *"Humanoid re-organizes the display section, mobile robot follows while carrying books."* - P1

Sometimes, robots in the background also served to attract customers or improve user experience.

> *"2 robot arms acts as chefs of a mom and pop, home cooked meal vibe cafeteria."* - P2

> *"Laundry bin (dog) & janitor robot (owner) collect laundry throughout the building and entertain people (play fetch with sheets)."* - P7

In both scenarios, robots display behaviors that are independent of their tasks (cooking & cleaning) for the benefit of users.

The defining feature that separates these scenarios from the user-focused scenarios is that these interactions are part of the service environment and would likely happen with or without the presence of the users. It is visible for both the users of the service and also for bystanders who pass by.

Enabling choice of robots
Having multiple robots in the environment also creates opportunity for people to choose the robots they want to interact with. For example,

> *"Humanoid receptionist and zoomorphic puppy rush to welcome children in the pediatric lounge"* - P4

> *"The humanoid robot and social service robot to go everywhere to check if anything weird is happening"* - P2

In the examples above, the use can choose the robot they are most comfortable interacting with. While this could be seen as a "simultaneous" interaction, the choice of which robot to use could also be perceived as a feature of the environment.

Multi-Robot Interaction

Multiple robots could also be part of the user's service experience. They can experience the robots simultaneously or sequentially. We also included a special category where the robots act as a mediator or interface to other robots.

Robots as Interfaces
Participants also talked about how a robot could be an interface to other robots. For instance,

> *"[robots] can act as guides around the mall while smart speaker on the [robot] can receive commands"* - P1

In the scenario above, the smart speaker acts as an input device and control for the other robot. There could also be complex scenarios where a user commands robots through a telepresence robot.

> *"Telepresence robot as doctor, tells physical service robot to perform certain checkups / administers needles, etc"* - P1

There could also be cases where a single robot acts a leader or manager of multiple other robots.

> *"The restaurant manager robot keeps track of every worker robot in the restaurant, overseeing how they perform, coordinate"* - P6

This example describes how people who are in managerial roles, such as restaurant owners, may want to coordinate all the actions and tasks of all worker robots through one single robot.

Simultaneous interaction
When a user simultaneously interacts with multiple different robots, we observe a scale in terms of how crucial the second robot is to the goal of the task. To gauge the importance of the second robot, we analyzed the scenarios to determine if the main goal of the task could be completed with only one of the robots. We only considered the robot with its stated capabilities (e.g. A telepresence robot has a screen but no manipulators) as any tasks could be completed with only one omnipresent, super capable robot. Though the notion of the second robot's importance is somewhat subjective[2], it can provide a framework to understand the role of robots in simultaneous interactions.

2: Some examples, such as providing distractions for patients, might be seen as complementary to some users but crucial for others. The examples were analyzed by two research team members. We chose not to rigorously categorize them as we believe these to be inherently ambiguous and subjective, depending strongly on the service requirement and tasks.

Crucial In the tasks envisioned by our participants where both robots are crucial, the scenarios often have the two robots completing different sub-tasks where one robot might not be able to complete each one. For instance,

> *"The social service robot checks you in at the gate while the humanoid robot checks if your luggage meets their constraints"* - P2

In the scenario, the robot lacks the manipulation capabilities to complete the tasks. Others might have the interaction with competing modalities where one robot cannot do everything such as

> *"Zoomorphic robot acts out specific scenes from history while the smart speaker narrates."* - P3

The zoomorphic robot cannot self-narrate the scene without breaking character. Lastly, the mere presence of more robots, even when not practically necessary, may be useful in some interactions for its own sake.

> *"VIP service: Pack of service robot/humanoid robot surrounds guests, one carries clothes, one carries a padlock, acting as security, one being the tour guide"* - P5

In this scenario, having the large number of robots is part of the experience for the user.

Emergent Behavior Participants also envisioned traditional robotic swarm applications where multiple robots come together to provide a service to a user.

> *"Hundreds of robots that are small in size can be summoned to create arrows for museum visitors to know where to go."* - P8

Participants mentioned large numbers of robots coming together to generate a single action. This category differs from "Crucial" in that not every single robot is needed for the interactions and likely able to tolerate a few failures or missing robot. However, it likely requires a critical mass of robots to be able to perform its role.

Improving The second robot could also speed up or improve the interaction, e.g., by parallelizing sub-tasks. Instead of one robot doing both tasks sequentially, a second robot (or more) could do the tasks and speed up the interaction.

> *"Humanoid robot reads to kids and mobile robot can go around to distribute supplies."* - P1

The addition of the second system could also provide redundancy to the interaction.

> *"Tabletop Robot and [smart speaker] act as welcome reception at each exhibit - these are said in chorus."* - P4

Supporting In some scenarios, the second robot played a supporting role for the user. For example, one robot could distract users while another robot completes the main task.

> *"One robot entertains the guest while they wait outside so a second robot has time to figure out what tables to assign."* - P8

A second robot could also be a supporter or an advocate for the user.

> *"The physically assistive robots checks you in for the appointment while the zoomorphic robot pretends to have the same problem as you."* - P2

Complementary Lastly, there are cases where additional robots or systems are included to smooth out an interaction or for aesthetic purposes such as adding sound and music to an existing interaction.

> *"The robot arm and the smart room collaborate to show you a dress with lights/musics."* - P2

Sequential interaction

Within the sequential human-multi-robot interactions generated by participants, we identified three categories. These are not mutually exclusive categories; multiple types of transfer are likely to be present in an individual human-multi-robot interaction. For example, a transfer of product (food) is likely preceded by a transfers of information (order).

Transfer of Product Between Robots In the first category, multiple robots transfer physical objects between them before engaging people.

> *"Robot up very high lowering things down to [robots] below for service."* - P7

For a more concrete example, consider an interaction involving food preparation. First, a robot in the kitchen prepares the food. Then, a second, mobile robot retrieves the food and delivers it to the customer. In these direct robot-to-robot transfers, the customer may not even be aware of the presence of the first robot. Alternatively, the customer might know about the interaction between the two robots but only directly interact with one of them. Furthermore, in a ubiquitous computing future with blurred lines between machines and objects, a robot might collaborate with a robot that the customer may perceive as not a robot. For an example, a humanoid robot might pick out an object from a mobile robot shelf and pass it to the user.

Transfer of User Information Between Robots For this category, *information* is transferred between robots, rather than physical objects.

> *"Cheap flying robot acts as sidekicks for the social service (more expensive) robot waiters that gauge how "needy" a table is (i.e. how much they want to get the waiters attention)."* - P3

Here, the flying robot serves a backstage role and might never be seen by the customer, so they may not even realize that multiple robots were involved. The user might also be aware of both robots: for example, a tabletop robot in a restaurant might monitor a table and summon another robot to deliver dessert when it detects the diners are ready.

Person Transfer Between Robots In this category, *the user* interacts first with one robot and then transferred to continue the interaction with another.

> *"[A robot] receives people at the counter for check-in, then directs [a] mobile robot to guide people to their respective waiting rooms."* - P3

In the scenario above, the customer initially interacts with one robot and is then transferred to a mobile robot to complete the task as the first robot is incapable of completing the task due to its service need to be at the counter. Both robots serve as *user-facing* robots. In other scenarios, a transition occurs between two robots, both capable of completing the task, due to the fact that one robot is simply better suited for it. For example, a user approaches a humanoid robot to request help finding an object in a warehouse store, but instead of the humanoid robot itself performing the task, a closer and faster mobile robot retrieves the object. More details about the features and rationale of person transfer are discussed in the following section.

4.2.5. Discussion

In both *product* and *information* transfers, a person might indirectly interact with the "backstage" robot. In the store example, the flying robot's role is solely to transfer information to the other robot when a customer needs assistance. Even if the person does not interact with the flying robot, they might notice that the interaction involved multiple robots. In these types of scenarios, it is worth considering whether these "backstage" robots should be visible or have any direct contact with people.

In contrast to *product* and *information* transfers, a *person* who is transferred will necessarily interact directly with multiple robots. This kind of transfer can happen under a variety of configurations which will influence how designers create the corresponding interactions. For example, a transfer where both robots are present in the same space as the person raises different design considerations than does a transfer in which the person needs to independently travel from a robot receptionist to a screen-based interface in another location.

Given this broad range of configurations and our research interest, more insights were needed to build understanding of the design space and interaction nuances of *person transfers* with multiple robots. We collected 13 examples of person transfers. These few examples generated in the scoping workshop encouraged us to conduct a second workshop in which we could focus specifically on this topic. As the service landscape of today already involves a variety of person transfers between human staff and digital systems, we also sought to capture characteristics of these existing transfer scenarios in the second workshop, and learn how future human-robot interactions may map onto them.

4.3. Ideation Workshop

4.3.1. Overview

The Ideation workshop build upon the work from the Scoping workshop to better understand the features and properties of Person Transfer Between Robots.

4.4. Method

The workshop was a single, 3-hour workshop with 7 designers as participants, and was conducted in early February, 2020. Participants were recruited through email lists and word of mouth and compensated 30 USD for their participation. The workshop was conducted in a classroom at Carnegie Mellon University's Pittsburgh Campus. We chose designers as they are accustomed to design processes and are capable of generating many high-quality ideas in a short amount of time. Furthermore, our goal did not require familiarity with special technologies. The workshop was approved by our university's Institutional Review Board.

The workshop had two parts: ideation and idea clustering.

Ideation After a warm-up exercise, participants took part in multiple rounds of idea generation. Similar to the scoping workshop, participants were shown a location or context in each round. Participants were given "Entity" cards labeled "Human", "Machine", or "Robot". The goal was to use these cards to define transfers between the selected entities. For example, participants were asked to draw 2 cards from the deck and generate ideas based on that combination. For example, a participant who drew "Human" followed by "Machine"

Figure 4.3.: Participants engaging in the Ideation phase of the workshop.

would be asked to generate ideas involving a transfer from a human to a machine. Using custom cards allowed us to constrain some of the design space and encourage participants to creatively think of new transfer forms. Each round lasted about 5 minutes and was followed by a discussion that allowed participants to share ideas and inspire new ones.

Idea Clustering Participants were split into three groups. Each group analyzed its own data and clustered it accordingly. This allowed participants to further elaborate on the different aspects of their ideas. Their clustering served as the inspiration and starting point for our analysis and taxonomy.

4.4.1. Analysis

The generated scenarios were analyzed by the research team using Affinity Diagramming [86]. We used this established method because we aimed to identify emerging themes rather than formulate a new theory. Instead of a coding scheme, this method relies on team members working together to reveal insights within the collected data [86]. Affinity diagramming was conducted iteratively through discussion over multiple sessions until the research team came to an agreement about the placement of all ideas. The clustering activity done by participants served as the starting point for our identification of important factors of sequential person transfers. This process was also supported by analysis of selected scenarios using FAST diagrams, which decompose services into their most basic actions by different entities [87].

4.5. Findings

In the ideation workshop, participants generated 287 different "one-line" scenarios for a range of service contexts and domains (e.g., airports, medical centers). Given our goal to capture existing person transfers in services today in addition to imagining future situations, we asked participants to think about existing interactions and service transfers they were familiar with and also extrapolated interactions if suitable.

As participants were asked to generate quantity over quality (as is common in brainstorming activities), some ideas did not involve person transfers, but instead focused on the transfer of product or information between systems and people (as in the previous workshop). For our analysis, we only included scenarios that fit the definition of *Person Transfers Between Robots*. Among the 287 generated ideas, 68 (23.6 %) fit this definition. These were combined with the 13 scenarios that matched the definition from the scoping workshop for a total of 81 scenarios that we used in our analysis.

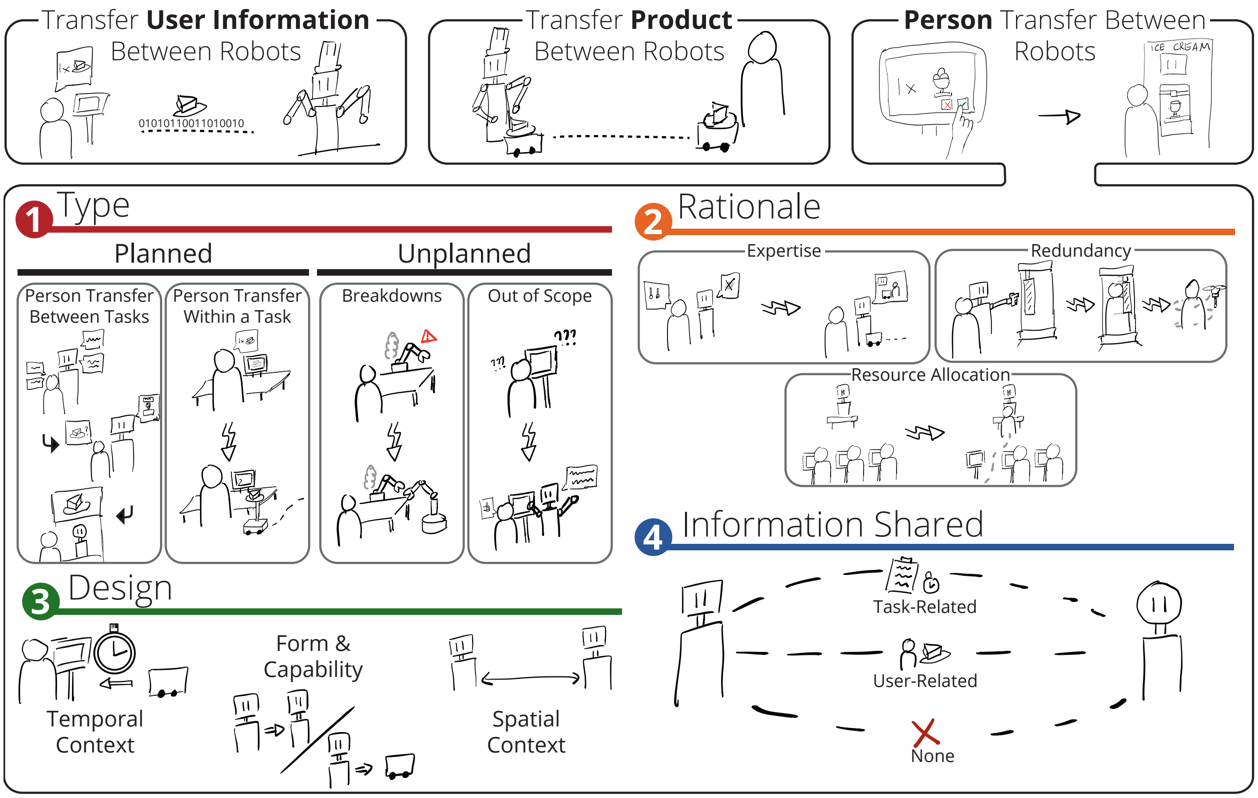


Figure 4.4.: Overview of the taxonomy.

4.5.1. Overview of Findings

Through affinity diagramming, we created a taxonomy for *person transfers between robots*. An overview of the taxonomy is shown in Figure 4.4. The taxonomy includes four dimensions:

Type of Person Transfers The service journey touchpoints where person transfers occur.

Rationale behind Person Transfer The interaction and service needs that drive the inclusion of person transfers in the service.

Design of Person Transfers The implementation variables of the person transfer throughout the service (time, space, and form).

Information Shared The different information shared and used between entities in person transfers.

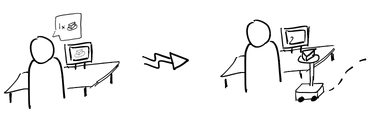
While some categories within these four dimensions are not mutually exclusive, considering each separately can contribute to a fuller understanding of person transfers and how to design for them. Rather than a comprehensive taxonomy, we present a set of prominent dimensions that should be considered for person transfers. We also included scenarios from the workshops to help motivate each category.

4.5.2. Types of Person Transfers

The first dimension looks at the moment in the person's service experience in which they experience a person transfer. The transfer can be either *planned* or *unplanned*. A planned person transfer is a transfer that always occurs and is required as part of the service experience. An unplanned transfer only occurs in some cases, deviates from the normal interaction, and is not intended to be experienced by the majority of people.

Planned Person Transfer

A planned person transfer can be further separated into a **Person Transfer between Tasks** and a **Person Transfer within a Task**. The distinction between tasks and subtasks is often ambiguous. For example, the experience of "getting on a flight" can be further decomposed into sub-tasks such as *checking in* and *going through security*. Each sub-task can be further subdivided into smaller actions; e.g., *going through security* may involve moving between the luggage scanner and human security staff. For our taxonomy, we define a task as a collection of compulsory actions that are **mandatory** to complete. In the airport security example, the person will fail the task if they choose to only use the luggage scanner but not the person scanner.

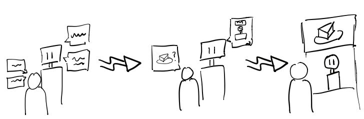

Person Transfer Within a Task

Person Transfer within a Task

These are scenarios where person transfers are an integral part of the interaction, and connect **two actions within the same task**.

> *"You print your boarding pass and luggage tag, then place [your] luggage on the conveyor belt."* - P10

In this case, if the person wants to drop off their luggage, they need to transfer between different systems to complete the task.


Person Transfer Between Tasks

Person Transfer between Tasks

This category describes situations in which people might do two **different tasks** in sequence—each task is independent, and the transfer only occurs if the person would like to engage in both parts of a service. This is likely to happen as part of a bundled service where different but complementary services are combined to increase customer convenience and value [88].

> *"Patient receives physical therapy before acupuncture treatment in the same center."* - P9

While both treatments are part of the same service, they are not dependent and can be completed separately and in any order.

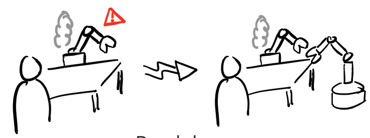
> *"A person checks in electronically and is directed to a waiting room, then [they are] summoned by a screen."* - P10

Transfers between tasks also include scenarios where people are transferred to a waiting area or given a notification device. We consider the waiting as a separate task, as it can potentially be skipped or even avoided if the service is optimized.

Unplanned Person Transfers

Unplanned person transfers are person transfers that are not intended to be part of the standard user experience. We identified two types:

Breakdowns & Failures



Breakdowns

These are person transfers caused by machine malfunction or human limitations, which may prevent them from being able to continue the task.

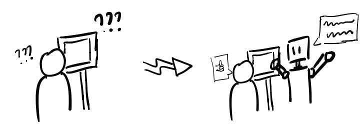
> *"A ticket machine breaks down, [it] calls over maintenance."* - P13

These breakdowns are not limited to machines, a human staff member may encounter challenges that prevent them from completing a task as well. This category also includes scenarios where the intended interaction fails due to extraneous circumstances.

> *"A tour guide cannot be heard by all tour members, so they refer tour members to an audio guide."* - P10

In this example, the tour group is too large for everyone to be able to hear the guide, resulting in a transfer from the tour guide to audio guides.

Out of Scope



Out of Scope

The second reason for unplanned transfers can be that a requested service is not part of the first robot's job scope or capabilities. While these transfers are likely to be pre-planned by the service, they are not requested by all users and often serve as a catch-all for any situation that robots might not understand or be able to address.

> *"[You] see the menu on an iPad, [then you] ask for clarification about ingredients [from] a waiter."* - P10

The device lacks information about the person's intent, which force them to transfer the interaction to another entity (a human employee, in this case) to retrieve the needed information.

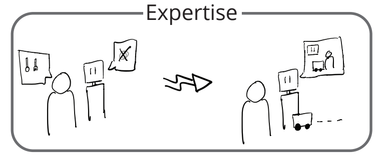
> *"You use an ATM to take out cash and then ask a bartender to break the large bills."* - P10

Similarly, the customer transfers from the machine to a human staff due to the machine's inability to complete the task.

4.5.3. Rationale Behind Person Transfer

The second dimension looks at the reasons why incorporating person transfers may be valuable to a service.

Expertise

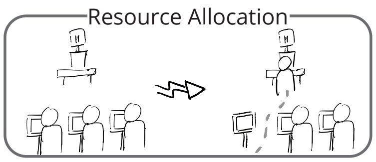


Expertise

In some instances, the first and second entities have different capabilities, both of which are required to complete the task. For example, the first entity that interacts with a person may be a good "user-facing" entity but lack the capability to fulfill their request, requiring it to rely on another entity.

> *"Tablets stationed for people to enter where they want to go, [for a robot] to guide them."* - P1

The tablet lacks the ability to move and guide users, therefore it would need to rely on the mobile robot to complete any navigational guidance tasks.

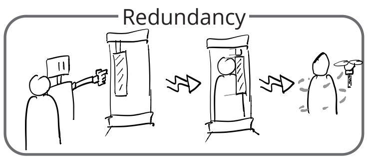


Resource Allocation

Using multiple entities may better allocate a service's resources and allow more people to receive service simultaneously. Instead of using a single entity, a service can be split to smaller parts and assigned to separate entities, similar to an "assembly line". Breaking down a task also allows some entities to engage "when needed"—they do not need to pay attention to every customer at every given moment.

> *"At food courts: tablet to order food, mobile robot delivers."* - P1

In this example, multiple people can order food simultaneously through the tablet and only a few robots deliver all of the food when it is ready.



Redundancy

Multiple entities take part in the same task to ensure safety by cross-checking or serving as a backup.

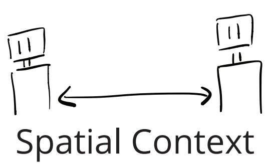
> *"Metal detectors [do] the first pass at the front of the museum, then someone goes through your bag."* - P14

Here, the person interacts with two entities, both of which exist for the same goal of detecting forbidden objects. Repeated interactions with several entities help mitigate potential errors.

4.5.4. Design of Person Transfers

The third dimension in the taxonomy is how person transfers are implemented and what actions a service provider might expect from people. This dimension covers both the physical and interactional aspects of person transfers.

Spatial Context



A person transfer can happen in a variety of spatial contexts. For example, both entities could be together (**co-located**), or they may never be co-present and thus require the person to move independently to complete the interaction (**remote**).

Co-Located

These are person transfers that occur when both entities are present in the same space as the person.

> *"[A robot] receives people at the counter for check-in, then directs a mobile robot to guide people to their respective waiting rooms."* - P3

The person interacts with the first robot, who then summons the mobile robot. At one point in the interaction (after the mobile robot appears), both robots are co-located in a single place. This allows both robots to interact with the person and each other.

Remote

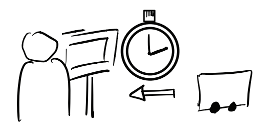
The alternative is to have two entities in different locations, which requires the person to move from one location to another to complete the task.

> *"Ticket machine directs visitors to a navigation computer."* - P12

From the person's point of view, there is no visible interaction between the entities. Furthermore, the interaction may differ based on *how* remote the entities are from each other. For example, they can be in view of each other, or apart in different rooms, requiring more instructions on how to locate the other entity.

Temporal Context

In addition to location changes, a temporal gap may also exist during person transfers. This gap can be caused either by the time needed for the second entity to join the interaction or by the time needed for the person to reach the second entity. Similar to spatial contexts, timing can also range from occurring **immediately** after the initiation of a transfer to occurring after some **time gap** between the two interactions.



Temporal Context

Immediate

These are situations where the second encounter happens nearly instantaneously after the first.

> *"Bouncer checks [you] before you go in, machine scans ID [and] unlocks door."* - P11

This kind of interaction is often coupled with *co-located* entities.

Gaps

Alternatively, transfers might take a while, as the person moves to the second entity or as they wait for the second part of the interaction.

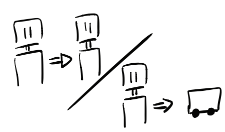
> *"Bartender asks you about what kind of drink you feel like getting, then directs you to a cocktail machine where you do it yourself."* - P10

The need to travel to the next machine creates a *gap* in time.

> *"Ordering food via a screen on flight. Flight attendant has to bring it to you."* - P13

Some gaps can also be caused by the time needed for the second entity to arrive after the person initiates the request.

Form & Capability

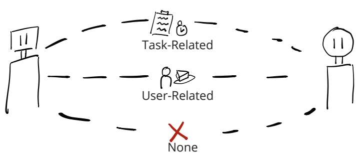


Form & Capability

As this is an exploration of transfers based on existing interactions between robots, devices, and human staff members, entities can take on different forms with vastly different capabilities. A person might be transferred between two entities with a **similar** form (e.g., human to human) or **different** forms (e.g., human to robot). The form is not a simple binary category, but instead a spectrum of similarity. Two humans wearing similar uniforms will have more similarity than two humans who are not. This is important in transfers because visual similarity provides some indication about with whom people should interact next. These cues can be designed for non-human entities, for example by using comparable branding or color palettes.

Form is also important because it shapes the entity's perceived capabilities [89, 90] and trustworthiness [91]. A robot with a visible face and mouth is more likely to be perceived as able to verbally communicate compared to a flying drone. The degree of difference between forms in person transfers can alter the amount of introduction and explanation that is required as part of the interaction.

4.5.5. Information Shared



The final dimension in our taxonomy concerns the information that is being shared between entities during person transfers. As digital systems have the ability to transfer and process large amounts of information, they can easily share information with the next system when person transfers occur. We identified three types of information sharing:

None

Information sharing does not always happen during person transfers. In some cases, the person might be asked to re-convey the information to the second entity to complete the task.

> *"Audio guide tells the visitor to ask a docent for additional information."* - P13

In this situation, it is likely that the docent will not be aware of the visitor's intent and will require them to explain their needs.

User-Related

The second type sets out to transfer newly learned information from the user, for example, what they hope to achieve.

> *"[There is an] iPad on a restaurant table for ordering food and calling staff to deliver the check and pay."* - P12

The patron communicates their intent through an iPad—the staff is likely to be aware of their intent upon arrival and already have their bill ready to go. This can also include non-task related personal information that the person disclose.

Task-Related

The most common information shared in our findings is regarding the state of the task. Task information about what had happened is generated by the first entity and used to inform the next entity and help its decision process. This differs from user-related in which they are generated by the first entity about the task on-hand not by the user. This information could be as simple as audio notifications or as complex as information about the state of the task.

> *"[An] X-ray security machine beeps so [the] human has to do a pat down."* - P13

> *"Self checkout machine where you get your blood pressure. It talks to you to say "I sent your information to the nurse, she will come in a minute.""* - P9

Furthermore, the sharing can be bidirectional, e.g., the second entity can send back information on their status and arrival time.

4.6. Utilizing the Taxonomy

In this section, we detail ways in which our taxonomy can be used for interaction design and research.

4.6.1. Interaction Design

To highlight the potential of our taxonomy for interaction design, we position it within a hypothetical design scenario—the design process for *Tri-bot*, an emergency room triage robot that helps with basic tasks such as measuring patients' vitals, interviewing them about their conditions, and leading them to different examination rooms accordingly. Besides its primary tasks of escorting and collecting patient's vitals and information, Tri-bot has to receive and transfer patients to other humans and machines for various treatments. We discuss how our taxonomy may be of use to Tri-bot's designers and to the hospital in various stages of the design process.

Need finding: Using the taxonomy, we can identify the primary role the robot should play. We can also use vocabulary provided by the taxonomy to help designers and stakeholders on the team discuss the different interactions between the robot and the users, and the range of transfers it would need to make.

*Tri-bot's primary role is to help with **resource allocation** by taking on basic tasks and allowing nurses to focus on more critical tasks. Most of Tri-bot's transfers will happen **within a task**, as the patients move between healthcare professionals.*

Robot behavior design: The taxonomy can help designers plan and design the robot's behavior during transfers, depending on the transfer type and the different aspects it involves.

*As patients are led around the hospital, nearly all transfers will be **co-located**. However, these transfers can happen **immediately** or after a while(**temporal gap**), depending on the availability of machines and clinicians. Therefore, Tri-bot needs to adapt its response based on the duration of the temporal gap and perhaps provide additional attention if the gap is too long.*

Physical design: The design of the robot's form and capabilities can also be informed by the types of transfers it will need to perform (identified in the previous two phases).

*The hospital may want Tri-bot to serve as a bridge between the initial triage nurse and other healthcare team members. In the beginning, Tri-bot's physical appearance is **different** from a typical nurse, and it is likely that a new patient will not have a mental model of Tri-bot's role. This will require introductions*

> *and guides for patients. Throughout the service experience, patients are likely to interact with **similar** Tri-bots and become familiar with how they look and work. Over time, less introduction and explanation will be needed.*

Inter-System Connection: The taxonomy can also assist in reflecting on what kind of information will be shared across robots and other service entities within a service. Designers should also determine how to disclose and be transparent about the information that is being shared.

> *As a task-driven robot, Tri-bot is most likely to share **task-related** information with other team members. Patients are likely to expect health information and vitals to be shared with their doctors; therefore, these transfers will not require patient consent. However, the service provider needs to determine how much **user-related** information is shared between systems. For example, making use of personal information disclosed during prior interactions might improve rapport with the patient, but it could also be off-putting.*

4.6.2. Researchers

In addition to designers, our taxonomy can also be useful for researchers in various methodologies. For example, it could be used during field work and observations to categorize and describe person transfers. An observation report may note:

> *The patient was transferred **between tasks** when they were directed to the checkout machine. The transfer was **remote** because the patient was given directions to go to the checkout desk by Tri-bot. Unfortunately, they got lost because the robot failed to....*

For evaluative user studies, the taxonomy can be used to isolate specific factors of person transfers to be further studied. For instance, a researcher interested in the effect of spatial context on the perceptions of robots could use this taxonomy to ensure other related factors are kept constant in their study.

> *This research explored how spatial and temporal contexts in a user transfer scenario affect user experience. In this 2x2 user study, we ....*

We can also analyze prior work according to this taxonomy. For instance, in [66], researchers demonstrated a *co-located*, *within-task* person transfer of two field robots with the same *form*. In [16], researchers showed a system that determines whether a user undergoes a *co-located* or *remote* person transfer based on efficiency.

Finally, this taxonomy has identified under-explored aspects of person transfers, thereby revealing high-priority research needs in these less known but important areas. Some examples include:

Types How does transfer type influence the information shared? How can we improve seamless transfers without sacrificing privacy?

Rationale What are the downsides of person transfers in different contexts, compared to using a single agent?

Design What are the trade-offs between different design choices, and how can interactions mitigate them? How can remote person transfers be better supported through design?

Information How should entities handle privacy and informational disclosure during person transfers? How might users perceive sharing of user-related information?

4.6.3. In this Dissertation

We can also use the taxonomy to categorize the other research thrust in this dissertation.

Chapter 5 This study exploring how different interactions between robots with *different forms* affect robot perception in a *co-located*, *immediate*, and *within-task* setting.

Chapter 7 This study explored the mobile robot joining behavior with robots of *different form* in a *co-located*, *immediate*, and *within-task* setting.

Chapter 7 This study explored the person transfers in the field with robots of *different form* in a *co-located*, *gap*, and *within-task* setting.

As an initial foray into this research topic, we chose to focus on similar types of person transfers with small differences. This allowed us to explore parts of the person transfer without worrying they were influenced by other factors.

4.7. Discussion

4.7.1. Transfer Trade-Offs

Sequential person transfers are beneficial in that they can create a more efficient service flow. They allow services to widely distribute resources to engage many people simultaneously, and they can utilize robots that are designed for a specific job. However, the decision to include person transfers is often a result of trade-offs among different factors. For instance, in a human guidance scenario, using the same robot for both kiosk-type services and guidance services will be faster as there is no time spent on a transfer. Furthermore, transfers would require additional logic and interaction complexity that must be explained and understood by people, which increases the odds of technological mistakes and failures. On the other hand, splitting the interaction might be more cost-effective because the service could use simpler, specialized robots for different parts of the task and serve multiple people. Recent work by Yedidsion et al. [92] has explored this trade off in the domain of robot guidance by choosing the transfer strategy (guide or instruct) based on the complexity of the instructions and traversability of the region. Designers and service providers should consider whether the efficiency and service flow that person transfers between multiple robots create are important enough to outweigh the problems of person transfers: development cost, user misunderstandings, and potential technological error.

4.7.2. Complexity in Transfer Interaction Design

If a service decides to undertake person transfers, several aspects that the taxonomy points out need to be considered: How far one robot is from the other (spatial context), how much waiting time might arise between the first robot and the second robot (temporal context), how does each robot look (form), and how should robots keep the person formed (information context)

A temporal gap between the two interactions with robots can reduce the quality of the experience. Designers need to consider how to engage or support a person during the gap. Similarly, large spatial gaps require ways to convey directions and additional support for lost users. The larger the gap, the more difficult it will be to create a seamless flow from one interaction to the other, and the less likely it will be for the interaction to be perceived as a single, holistic one.

The form of the robots can also significantly impact the interaction and the potential success of a transfer. Any transfer between two robots with different shapes or exterior designs will require some visual cue to communicate that they are part of the same interaction. The less similar the two robots are, the more a transfer entity will need to clearly communicate the role of each robot and information shared among them. Similar forms can also be achieved through robot re-embodiment, where the robot's "personality" or "identity" transitions from one robot body to another and indicates to the user that a transition has occurred [14, 93].

Lastly, depending on the user's mental model, they might be unsure about what the second robots knows. They might believe all the robots were connected and all information has been shared. They also could have envision them to be separate and required some form of information transmission. If they do need to transfer information, how should it be done and how to convey the information transfer to the user? We explored these questions further in Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer).

4.8. Limitations

Our design workshops primarily focused on service contexts where people are the customers. While we attempted to cover a large variety of contexts, such as healthcare, transportation, etc., other contexts might present unique characteristics and alternative needs for person transfers. Furthermore, our work did not touch on non-service environments, such as the home and workplace. Those spaces present additional challenges and considerations, such as robots' dynamics with different members of the family.

The workshops also only explored scenarios where a second entity was available. This underlying assumption may not always be valid. There will be scenarios where a second robot might not always be available. Failed transfers should also be further explored.

One of the challenges of this work was that the phrase "person transfers" is somewhat ambiguous, and could mean a variety of things to participants based on their personal interpretation. Participants mostly understood "person transfers" as a sequential action where someone moves from one entity to another, but they occasionally included ideas outside of this scope. In some scenarios, participants provided scenarios where the word "person" was one of the entities involved and the word "transfers" was the verb (e.g. 'Sushi delivery line transfers sushi to customers."). This ambiguity caused us to ignore multiple scenarios in our analysis as they fall under one of the other two types of sequential interactions. However, some transfers of information and products can be seen as the precursor or post-cursor events during a person sequential transfer.

"Apps tell user about what they should see at museum (or Siri/Alexa, etc)" - P7

. While the app does not actively transfer the user, it encompass the interaction where a system is telling a user where to go next and being transferred.

Lastly, we chose to be flexible with the interpretation of our scenarios as the categorization of each scenario was loosely defined. For example, in the scenario *"iPad on restaurant table for ordering food and calling staff to deliver check and pay"*, if we added "elsewhere" at the end, the transfer could change from a co-located spatial transfer to a remote transfer. Thus, we extracted insightful categories, even if they were clearly represented in only one or two scenarios in our collected scenarios. This was done in order to obtain a broad understanding of the factors that might impact this design space.

4.9. Conclusion & Contributions

The work discussed in this chapter led to (1) a categorization of multi-robot systems in a service environment and (2) a taxonomy and vocabulary to describe sequential *person transfers*. Our multi-robot categorization provides a framework to understand the roles of the robots in a service context. We discuss the three main types of sequential interaction and highlight how *person transfer between robots* is different. Our taxonomy provides a blueprint for HRI researchers and designers to create and evaluate different kinds of person transfers and explore potential opportunities in the area. We discuss opportunities and trade-offs of including person transfers across the four primary dimensions of the taxonomy – Rationale, Type, Design, and Information Shared. Through a better understanding of the dimensions of person transfers, it is possible to design better experiences that integrate diverse technological devices into a seamless service experience.

Inter-Robot Communication & Information Transmission In Person Transfer

5.

5.1 Overview 45
5.2 Method 46
5.3 Results 53
5.4 Discussion 57
5.5 Limitations 58
5.6 Conclusion & Contributions 59

5.1. Overview

Research in this chapter was conducted in collaboration with Samantha Reig, Elizabeth Carter, and Aaron Steinfeld.

Parts of this chapter are adapted from the following publications: *Xiang Zhi Tan, Samantha Reig, Elizabeth J Carter, and Aaron Steinfeld. 'From One to Another: How Robot-Robot Interaction Affects Users' Perceptions Following a Transition Between Robots'. In:* 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2019*

As explained in Section 3.2.2, robot-to-robot social communication has been an active area of research. Most of the prior work has focused on group interactions where multiple robots were present from the beginning of the experience. In co-located person transfers where the user first interacts with one robot before the other robot, it is unclear how the robots should interact with each other and how the first robot should convey information transfer between the robots. We identified this as one of the dimensions in our taxonomy (Subsection 4.5.5 (Information Shared)).

The arrival of a second robot raises questions about what the second robot knows about the previous interaction. As robots are not limited by human capabilities, they can transmit large amounts of information between each other without the user's knowledge. Prior work by Williams et al. [12] has explored a similar topic and found covert communication to be undesirable. Our work expanded on this by (1) focusing on person transfer rather than only fostering ongoing collaboration, (2) occurring in a different setting where robots work together to provide a service, and (3) within in a context where the robots have different social capabilities.

If covert communication is undesirable, the robots need to perform[1] a communicative act to convey the transfer of information. One method would be to equip robots with social capabilities, which has been shown to increase user engagement and enjoyment [95, 96]. However, the second robot, especially those that are task-driven, may not be designed to have human-like or animal-like features to leverage during social interactions. In a world of multiple robots and intelligent systems, it is likely that users will interact with a variety of robots and observe interactions between highly social robots and robots with only minimal social capabilities. (For readability, we will call these robots "nonsocial"[2].) In observing social behavior between a social robot and a nonsocial robot, users may attribute social intelligence, trust, and other properties normally associated with social robots to the nonsocial robot, even in the absence of its own social expressions. By being deliberate in the design of the way one robot treats

1: In our scenario, the act of the two robots communicating is entirely performative and no actual information is transferred in the act.

2: In the original publication [94], we also called the mobile robot "functional" in some parts of the paper, which implied that social robots are not functional and confused some readers. We intend to convey that the robot was designed primarily for a specific task and not social interaction with a human and has very little social capability another, we may be able to increase these attributions toward a robot that does not exhibit social qualities.

5.1.1. Chapter Summary

In this chapter, we examine how information transfer and social interactions between two different robots affect user perceptions during person transfer. We first describe a user study that tested three types of information transfer and three degrees of sociality in an interaction among two robots and a human participant. We then present our result that showed that the way a stationary social robot treats a nonsocial mobile robot with minimal social capabilities changes how humans perceive the mobile robot. Lastly, we discuss the implications of our study on how designers should design verbal interaction in not only person transfer scenario, but also multi-robot human interaction.

5.2. Method

We designed a laboratory study about how Information Transfer and Stationary Robot Behavior influenced participants' preferences and perceptions. As discussed in Chapter 4 (Design Space for Multiple Robots And Person Transfer), a person transfer between robots can happen due to the first robot unable to complete the task due to incompatible capabilities or physical constraints. The study followed this rationale and had users interacting with two robots with different capabilities. We created a navigation scenario in which a person requested assistance from a social, stationary robot that then summoned a nonsocial, mobile robot to lead the person to a destination. The study was approved by our university's Institutional Review Board.

5.2.1. Study Design

The study was a 3 x 3 mixed-design experiment with Information Transfer as the between-subjects manipulation and Stationary Robot Behavior as the within-subjects manipulation[3]. Information Transfer explored different ways for robot-robot information transfer to be signalled to the user. This manipulation was not to determine the best way for information to be transferred (electrical signaling is often the best option due to low noise, high reliability, and high bandwidth); instead, it was to learn how robots should indicate to their users that certain information has been shared between two robots. Our Information Transfer conditions were as follows:

3: We choose to have participants experience all three Stationary Robot Behavior conditions so that participants could compare and rank them.

- *Silent* – The stationary robot did not repeat the user's request and did not explicitly acknowledge that the request had been transferred to the mobile robot;
- *Explicit* – The stationary robot did not repeat the user's request but did explicitly acknowledge that the information had been sent to the mobile robot;

- *Reciting* – The stationary robot recited the user's request out loud to the mobile robot.

The Stationary Robot Behavior conditions described how a stationary robot might interact with a mobile robot that lacks speech capabilities. The conditions were as follows:

- *Representative* – The stationary robot did not speak directly to the mobile robot, but instead spoke to the participant on behalf of the mobile robot;
- *Direct* – The stationary robot turned to and spoke directly to the mobile robot, delivering the participant's request in a complete sentence;
- *Social* – The stationary robot turned to and spoke directly to the mobile robot, supplementing the participant's request with social conversational behavior.

In our mixed-design study, each participant was assigned to one of the Information Transfer conditions and experienced each of the three Stationary Robot Behavior conditions. The order of the Stationary Robot Behavior conditions was counterbalanced and the Information Transfer conditions were distributed equally across the 6 unique permutations of Stationary Robot Behavior. An illustration and exact dialog of each condition is shown in Figure 5.1.

5.2.2. Hypotheses

Our hypotheses predicted that both Information Transfer and Stationary Robot Behavior would affect participants' preference and perception of both robots.

H1. Participants will perceive the stationary robot to be more social, competent, and likable in the *social* condition than in the *direct* and *representative* conditions.

H2. Participants will perceive the mobile robot to be more social, competent, and likable in the *social* condition than in the *direct* and *representative* conditions.

H3. Information Transfer will have an effect on participants' perception of both robots.

(a) Participants will perceive the robots to be more competent in the *reciting* condition.

(b) Participants will perceive lower competence in the mobile robot and be more wary of and disturbed by the stationary robot's behavior in the *silent* condition.

H4. Participants will be more likely to see the robots as each other's equals in the *social* condition.

H5. Participants will have a higher preference to work with the mobile robot they encounter in the *social* condition.

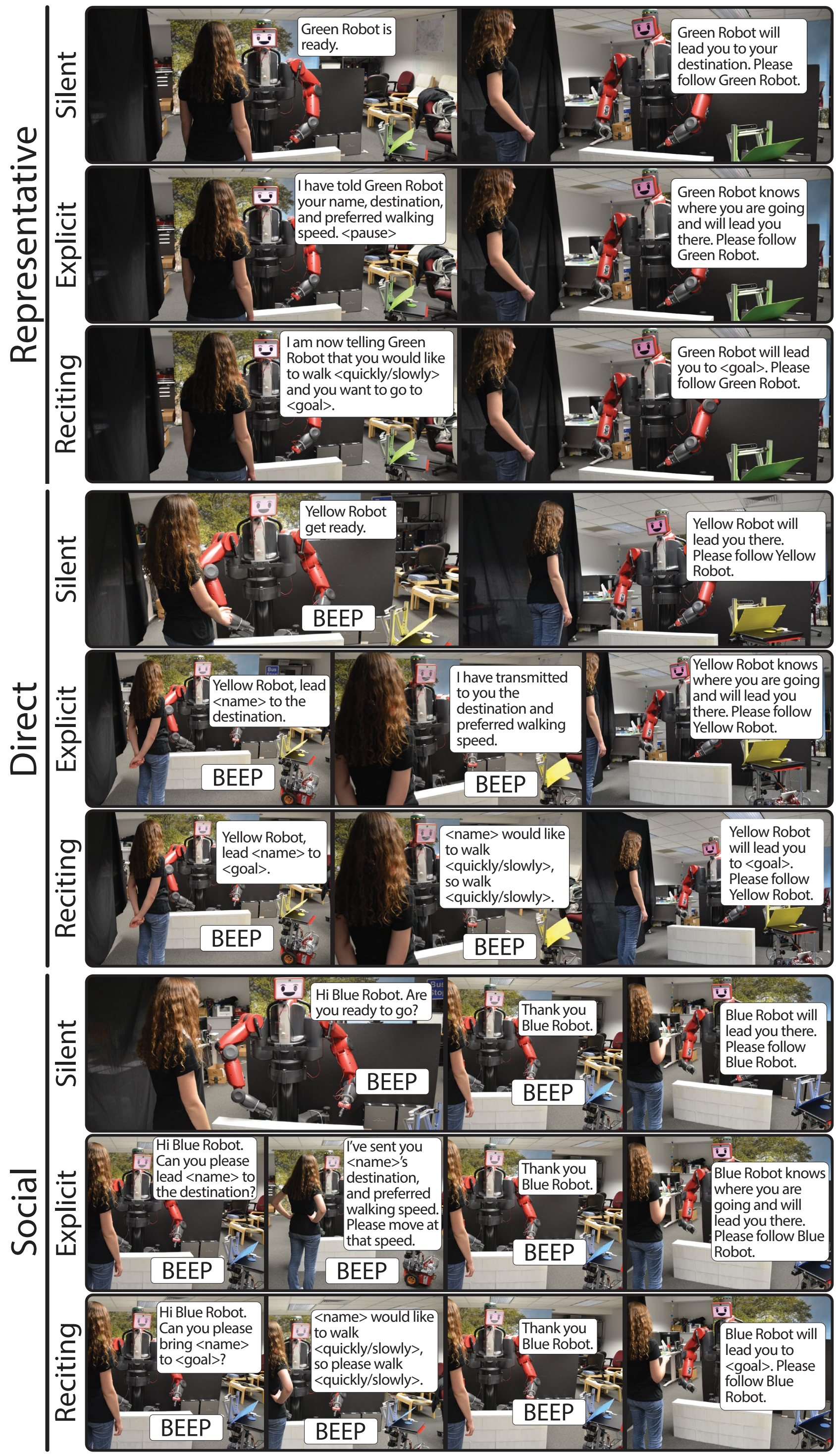


Figure 5.1.: Stationary and mobile robots' dialogue for each condition.

5.2.3. System and Study Setup

The study was conducted in an enclosed lab space at Carnegie Mellon University's Pittsburgh campus. The layout of the study space is shown in Figure 5.2. The stationary robot was a humanoid Baxter robot by Rethink Robotics with speakers added for projecting a clear robot voice and a camera mounted on the chest to record participant reactions. The Baxter robot wore a different name tag in each of the Behavior conditions to add to the perception that the context had been switched between each of the 3 interactions that constituted the within-subjects manipulation. The robot was controlled through its ROS-based SDK. Realistic speech was generated using the Amazon AWS Polly SDK, and natural human responses were recognized through a pipeline consisting of Google Cloud Speech and SNIPS Natural Language Understanding Engine [97].

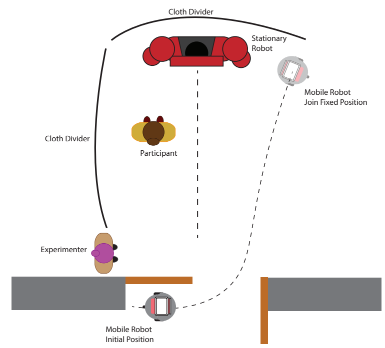


Figure 5.2.: Illustration of top down view of the study location.

The mobile robot was a custom-modified Mobile Robots P3DX. The robot was controlled with ROS and could autonomously navigate to multiple fixed waypoints throughout the study area using the ROS Navigation Stack [98]. Additional motion corrections could be made using a joystick controller if the robot moved erroneously. The robot was decorated with colorful accessories to allow participants to easily distinguish between trials and enhance the illusion that each trial had a different mobile robot. All participants saw a "blue robot", a "green robot", and a "yellow robot". There were two mappings of colors to Stationary Robot Behavior conditions.

Because both robots were independent units that operated separately by default, we wrote a ROS package, inspired by [58], to facilitate communication between the two robots through the ROS Bridge protocol [99]. This package allowed us to send signals between the robots so that one robot could tell the other robot to execute certain actions or that an action had been completed. In our study, the stationary robot wirelessly coordinated the experience and told the mobile robot when certain behaviors were needed. To move the mobile robot, the stationary robot sent a signal with a waypoint (next to the participant, at the door to the lab, etc.) to the mobile robot. The mobile robot then autonomously planned and executed a path to the given waypoint. When the path had been completed, the mobile robot sent back a "done" signal to the stationary robot, which waited for that signal before continuing the interaction with the participant.

Both robots operated autonomously during the study. Occasionally, the experimenter adjusted the mobile robot's motion if it moved too slowly or too close to an obstacle. We have released the code behind the study online[4].

4: The code can be found at `https://github.com/CMU-TBD/HRI19-MultiRobot-Transition-Study`.

5.2.4. Procedure

After obtaining informed consent, the experimenter explained to the participant that they would take part in a scenario in which they required guidance to navigate to a room in an unfamiliar building. The experimenter told the participant that they would engage in the interaction four times—with the first being a practice run—and that they would verbally interact with a differently programmed robot during each trial.

To mitigate the impact of language processing failures, we told participants to be aware that the system was not perfect and might fail to recognize some commands, and we provided tips on what to do when failures occurred.

In the practice run, the participant interacted with the stationary robot, which summoned the mobile robot to lead the participant from our lab to a door in the hallway. Once the mobile robot arrived at its destination, it beeped once to signal task completion. The experimenter then led the participant back to the room. The experimenter explained that the robot would drive itself back to the "charging station" outside of the lab. We chose to have the mobile robot park itself outside the lab (and out of sight of the participant) between trials for two reasons: (1) to prevent participants from thinking that the mobile robot gathered information by overhearing their conversation with the stationary robot rather than obtaining it directly from the stationary robot, and (2) to contribute to the illusion that there were three different mobile robot behavior "programs" by having the robot leave and return with a different appearance. When the participant re-entered the lab, they completed a demographic survey and the first of three questionnaires (Q1)[5].

5: The first questionnaire was about the participant's perception of the robots in the practice run. The questionnaire served to familiarize participant's with the questionnaire format.

Following Q1, the experimenter explained that the participant would need to follow the robot only to the door of the lab to complete each trial rather than all the way to a destination. We chose not to have the mobile robot navigate to a destination to ensure that the participant's experience and opinion of the mobile robot were based on the person transfer rather than its success in guiding the participant. The participant was also told that they would be prompted to select a walking speed ("quickly" or "slowly"), but that the speed selection would not determine the mobile robot's actual speed; this was to ensure that perceptions were not affected by the mobile robot's success or failure to match expectations of speed. This clarification stems from pilot studies where participants rated the robot's competence not on the interaction but failure to see a difference when they asked for different speeds.

When the participant was ready, the experimenter pressed a button to start the next interaction. The stationary robot's face appeared on its screen. The stationary robot waited for the participant to initiate the interaction and began when it heard a prompt. The robot then greeted the participant and asked for the participant's name (which was entered in advance to ensure correctness). The robot then asked what kind of assistance the user required. Our system used the natural language pipeline to parse the request. Once the system correctly extracted the destination, it asked the participant to confirm, and then asked them for their preferred walking speed. Then, it summoned the mobile robot to the fixed waypoint to the right of the participant. When the mobile robot arrived, it turned to the stationary robot and beeped to signal its arrival. In the *explicit* and *reciting* conditions, the stationary robot turned its arms and head towards the mobile robot. Dialogue for the condition was then executed (Figure 5.1), ending with the stationary robot instructing the participant to follow the mobile robot. The mobile robot moved outside of the lab, at which point the experimenter stopped the trial and administered the next questionnaire (Q2). The experimenter then exited the room, stating that they needed to switch the programs of both of the robots while the participant completed Q2. In reality, the experimenter only switched the program of the stationary robot, the name tag of the stationary robot, and the color accessories on the mobile robot. This interaction was then repeated two more times, once for each of the other within-subjects conditions. Each condition was associated with a fictional setting (Goliath National Bank, Echo Credit Union, Hugo National Bank) and a color (blue, yellow, green). To ensure that robot color was not a confound, we rotated the color order after 18 sessions (6 per between-subjects condition).

After completing Q2 for the last trial, the participant completed another questionnaire (Q3) comparing all three trials. We then conducted a semi-structured interview to gain further insight into participants' impressions. The study took about 30 minutes and participants were compensated 8 USD.

5.2.5. Measures

Because we were interested in the participants' perceptions of the interactions, we relied primarily on subjective measures. Our measures included several Likert scales drawn from prior work, experiment-specific forced-choice questions, yes-or-no questions about the robots' knowledge, and open-ended questions. When drawing from validated scales, we selectively omitted less relevant items to prevent survey fatigue. The complete set of questionnaires used in the study can be found in Appendix A (Questionnaires for Chapter 5).

Perception of Social Properties

To assess both robots, we combined sections of the Robotic Social Attributes Scale (RoSAS) [100] and the Godspeed questionnaire [101]. Participants were asked to rate the robot(s) with respect to 12 words from the *warmth* and *competence* RoSAS factors. We also asked them to rate 3 words from Godspeed (Likable, Mean, and Friendly) to measure perceived *likability*.

Trust in Guide Robot

Participants' trust in the system was measured by 6 questions, 4 of which were modified from Jian's trust scale [102] and two that were specific to the task.

Other Likert Questions

We included another 9 task specific Likert questions that are designed to access the relationship between robots and participants empathy towards the robots (modified from the Empathy Concern factor of the interpersonal reactivity index [103]).

Open Ended Questions

At the end of each trial, we asked participants to describe the relationship between the robots and what they liked or disliked about the interaction. After the third trial, we asked participants how they believed information transfer had occurred and which pieces of information had been transferred between robots.

Preferred Robot

The final questionnaire also included forced-choice questions asking participants which of the three mobile robots they would most want to use, which one they felt most connected to, which one they preferred the least, which one they found most likable, and which one they believed to be the most knowledgeable. We also asked in which trial the stationary robot was most likable and least preferred.

5.2.6. Other Data

We recorded each session and logged how often the robot repeated a question due to a natural language pipeline failure.

5.2.7. Participants

We recruited 44 participants from the Pittsburgh metropolitan area using an online recruitment tool. Participants were between 18 and 61 years old. Eight participants were excluded due to logistical or technical issues, resulting in 36 valid sessions (12 per between-subject condition; Table 5.1). Participants reported using computers on a near-daily basis, $M = 6.89$, $SD = 0.32$, on a 7-point Likert scale that ranged from Never (1) to Daily (7). Participants also reported some familiarity with robots, $M = 3.17$, $SD = 1.28$ on a 7-point scale.

	Female	Male	Other
Silent	8	4	0
Explicit	9	3	0
Reciting	8	4	0
	Age (Std. Dev.)		
Silent	29.5 (16.2)		
Explicit	27.1 (12.5)		
Reciting	26.3 (8.5)		

Table 5.1.: Participant demographics (36 valid sessions)

5.3. Results

Unless otherwise noted, we analyzed the results by fitting a multi-level linear model using REstricted Maximum Likelihood (REML) [104, 105] for all measures with Information Transfer and Behavior as fixed effects and participant as a random effect nested within Information Transfer[6]. The number of system mistakes (times the robot repeated a question due to language pipeline errors) was treated as a covariate and included as a fixed effect to ensure that differences in ratings were due to our manipulation and not due to system usability. All post-hoc analyses used Tukey's Honest Significant Difference (Tukey HSD). We report significant differences ($p < 0.05$) and important trends ($p < 0.1$).

6: We used JMP13 for all calculations.

5.3.1. Measure Reliability and Confounds

The RoSAS *warmth* factor was reliable for both the stationary robot (Cronbach's $\alpha = 0.92$) and the mobile robot (Cronbach's $\alpha = 0.94$). The *competence* factor was also reliable for both robots (Cronbach's $\alpha = 0.96$ and 0.92, respectively). The Likability measure was calculated by averaging participants' responses to items pertaining to the robot's *Likability* and *Friendliness*, which had item reliability of Cronbach's $\alpha = 0.90$ for the stationary robot and Cronbach's $\alpha = 0.83$ for the mobile robot. We did not include the reverse coding of *meanness* in this index because it had a low correlation with the other items. Instead, we analyzed *meanness* individually for both robots. Five of the six items that assessed trust in the mobile robot were highly correlated ($r > 0.85$). The exception was the item "I am wary of the guide robot", which was weakly correlated with the other items ($r < 0.34$). We combined the five highly correlated items (Cronbach's $\alpha = 0.98$) as a measure of trust in the mobile robot. To assess the perceived relationship between robots, we constructed a *relationship* factor composed of responses to Likert items pertaining to beliefs that the robots *knew each other well*, *ignored each other* (reverse coded), and *liked each other* (Cronbach's $\alpha = 0.76$).

To evaluate the possible effect of the mobile robot's color, we included color in a similar multi-level linear model. We did not find significant effects of color on our dependent measures.

5.3.2. Perception of Mobile Robot

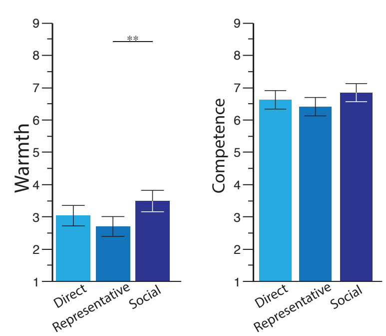


Figure 5.3.: Effect of Stationary Robot Behavior on Participants' Perception of Mobile Robot Warmth & Competence.

We measured participants' perceptions of the mobile robot through RoSAS and other measures. For the *warmth* measure, we found that Behavior had a significant effect, $F(2, 58.22) = 5.70$, $p = 0.006$. Pairwise analysis showed that participants felt that the *warmth* of the mobile robot was significantly higher in the *social* condition ($M = 3.45, SE = 0.33$) than in the *representative* condition ($M = 2.67, SE = 0.31$), $p = 0.004$. No other pairwise difference was found. We also found a trend wherein Stationary Robot Behavior impacted perceived competence of the mobile guide robot, $p = 0.092$. On the mobile robot's likability, we found a significant effect of

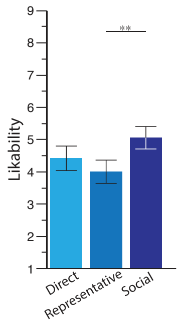


Figure 5.4.: Effect of Stationary Robot Behavior on Participants' Perception of Mobile Robot Likability

the stationary robot's Behavior, $F(2, 58.09) = 6.62$, $p = 0.003$. A pairwise comparison showed that when the stationary robot was *social* toward the mobile robot ($M = 5.04, SE = 0.38$), the mobile robot was more likable than when the stationary robot was *representative* ($M = 3.99, SE = 0.37$), $p = 0.002$. No other pairwise difference was found.

We also found a trend where Information Transfer influenced the participant's wariness of the mobile robot, $p = 0.057$. In particular, participants were more wary in the *silent* condition ($M = 3.12$) than in *reciting* condition ($M = 1.73$).

We found no significant difference among factors for participants' perceived *meanness* or trust of the guide robot.

5.3.3. Perception of Stationary Robot

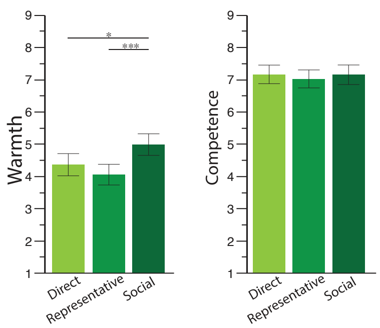


Figure 5.5.: Effect of Stationary Robot Behavior on Participants' Perception of Stationary Robot Warmth & Competence.

The Stationary Robot Behavior significantly affected perceived *warmth*, $F(2, 58.19) = 8.53$, $p < 0.001$. Pairwise comparisons showed that participants rated the stationary robot higher on *warmth* in the *social* ($M = 5.01, SE = 0.34$) condition than in the *direct* ($M = 4.37, SE = 0.33$) and *representative* ($M = 4.03, SE = 0.33$) conditions, $p = 0.028$ and $p < 0.001$, respectively. Though participants were treated equally by the stationary robot across conditions, their perceptions of its warmth changed when it treated the mobile robot in a nonsocial manner.

We found a trend wherein Information Transfer influenced participants' *competence* ratings of the stationary robot, $p = 0.064$, such that participants viewed the stationary robot as more competent in the *reciting* condition ($M = 7.93$) than in the *silent* condition ($M = 6.36$).

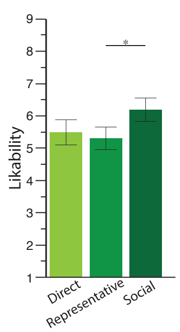


Figure 5.6.: Effect of Stationary Robot Behavior on Participants' Perception of Stationary Robot Likability

Participants' ratings of the stationary robot's likability were significantly affected by Information Transfer, $F(2, 31.84) = 3.90$, $p = 0.031$, and the Stationary Robot Behavior, $F(2, 58.22) = 3.81$, $p = 0.028$. Pairwise comparisons revealed that participants perceived the stationary robot as more likable when it used *reciting* ($M = 6.60, SE = 0.53$) than when it was *silent* ($M = 4.55, SE = 0.52$), $p = 0.025$. Participants also perceived the robot to be more likable in the *social* condition ($M = 6.20, SE = 0.37$) than in the *representative* condition ($M = 5.30, SE = 0.35$), $p = 0.025$.

Both Information Transfer and Behavior significantly affected *meanness* ratings for the stationary robot, $F(2, 34.47) = 3.31$, $p = 0.049$ and $F(2, 62.82) = 4.91$, $p = 0.011$. Again, pairwise comparisons showed that participants perceived that the robot was less mean in the *reciting* condition ($M = 1.31, SE = 0.37$) than in the *silent* condition ($M = 2.61, SE = 0.35$), $p = 0.038$. Participants rated the *social* ($M = 1.33, SE = 0.31$) stationary robot as significantly less mean than the *direct* ($M = 2.45, SE = 0.29$) stationary robot, $p = 0.011$. There was also a trend in which the stationary robot was perceived as less mean in the *social* condition than in the *representative* condition, $p = 0.053$.

5.3.4. Robot-Robot Relationship

We explored the perceived relationship between the robots via (1) a *relationship* factor and (2) an analysis of participants' responses to a free response question asking them to describe the robots' relationship after each trial. We found Behavior to have a significant effect on how participants perceived the relationship, $F(2, 59.77) = 13.43$, $p < 0.001$.

However, we also found there was a significant interaction effect of Behavior and Information Transfer, $F(4, 59.46) = 2.56$, $p = 0.048$. The effect of Behavior differed depending on Information Transfer: participants perceived the mobile robot to have a better relationship with the *social* stationary robot than with the *representative* stationary robot in the *explicit*, $d = 1.377$, $p = 0.031$, and *silent*, $d = 2.01$, $p < 0.001$, Information Transfer conditions. Participants perceived the robots in the *social* condition to have a better relationship than the *direct* condition when the Transfer condition was *silent*, $d = 1.51$, $p = 0.013$. This interaction effect showed that the difference in robot Behavior was mainly in the *silent* condition.

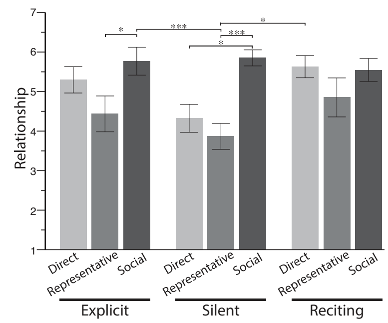


Figure 5.7.: Effect of Stationary Robot Behavior and Information Transfer on Participants' Perception of Robot Relationship.

For the qualitative responses about the robots' relationship, we annotated answers to the open-ended questions: "How would you describe the relationship between the receptionist robot and the guide robot?" and "What did you like and/or dislike about the interaction with the robots?" for information about the *type* and *nature* of the perceived relationship between the robots. The author and another research team member inspected responses and identified 3 categories of relationship type: *equal*, *unequal*, and *no relationship*. Within the *equal* relationship type, there were 2 categories of relationship *nature*: *prescribed*, e.g., commanded by the programmer to act as equals; and *independent*, e.g., friends or coworkers. Within the *unequal* type, there were 3 categories of relationship *nature*: *positive*, e.g., teacher and student; *neutral*, e.g., boss and employee; and *negative*, e.g., master and slave. Some participants did not address relationships in their answers, and in this case, a code of "N/A" was assigned.

	Equal	Unequal	Other
Social	20	8	8
Direct	8	19	9
Rep.	11	18	7

Table 5.2.: Coded perceived robots' relationship by Stationary Robot Behavior.

Two coders coded 25% of the data to calculate inter-rater reliability. For relationship type, Cohen's κ was 0.84, and for relationship nature, Cohen's κ was 0.79. One coder coded the rest of the data. The three Behavior conditions were analyzed individually within each Transfer condition.

In the *explicit* Transfer condition, a Fisher's Exact test revealed an association between Behavior conditions and relationship *nature*, $p = 0.033$. The *representative* and *direct* conditions were more likely to merit perceptions of *unequal negative* and *unequal neutral* relationships than the *social* condition, but pairwise comparisons using a Bonferroni corrected α of 0.0166 ($\frac{0.05}{3}$) did not reveal significant differences.

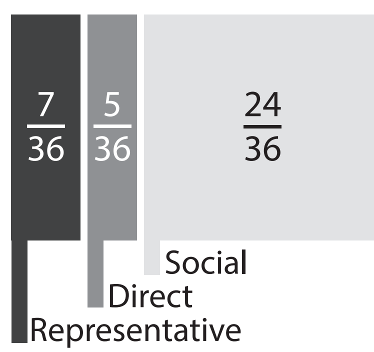


Figure 5.8.: Participant preferences for Stationary Robot Behavior.

5.3.5. User Preference

At the end of the study, participants chose which robot they preferred and matched certain descriptive words to one of the three mobile robots. When asked which mobile robot they preferred to lead them to their destination, participants generally preferred the robot from the *social* condition (24/36), followed by the *representative* condition (7/36) and finally the *direct* condition (5/36). A Fisher's Exact test found a significant association between choice of robots and Information Transfer condition, $p = 0.008$. Individual pairwise comparisons with a Bonferroni-corrected α of 0.0166 revealed a significant difference in robot preference between the *silent* and *reciting* conditions, $p = 0.005$: participants in the *silent* condition were more likely to prefer the mobile robot in the *social* condition. When participants reported which robot they preferred the least, 20 mentioned the robot in the *representative* condition, 12 picked the robot in the *direct* condition, and 4 chose the robot in the *social* condition. Fisher's Exact tests on all other forced-choice questions including least preferred showed no significant differences across conditions.

At the end of the study in Q3, we also asked participants to rate the competence of each mobile robot on a 7-point scale. We combined the responses and analyzed them using the same multi-level linear model but without mistakes as a covariate. We found a main effect of Information Transfer on perceived competence of the mobile robot, $F(2, 87.92) = 5.26$, $p = 0.007$. Pairwise comparisons found that participants rated the mobile robot as more competent in the *reciting* ($M = 6.28, SE = 0.26$) and *explicit* ($M = 6.25, SE = 0.23$) conditions than in the *silent* condition ($M = 5.32, SE = 0.23$), $p = 0.020$ and $p = 0.016$.

5.3.6. Other Findings

What Information Was Transferred

In the final survey, we also asked whether participants believed the mobile robot knew certain information. Nearly all of the participants believed the robot knew their desired walking speed (34/36), their destination (34/36), and that they had requested help navigating (33/36). However, only 17 out of the 36 reported believing that the robot knew their name. In the open-ended questionnaire responses and interviews, multiple participants mentioned that since the mobile robot could not speak, there was no reason for it to know their name and/or confirm knowing their name.

How Information Was Transferred

Participants also explained how they believed the information was transferred between robots. We grouped responses by similarity and found that participants believed that information was transferred between the robots in several ways. A few participants thought that the robots communicated out loud, and that the mobile robot interpreted the stationary robot's speech and beeped back in response. For example, P301 thought that information was transferred when *"the guide robot beeped loudly... signaling to the receptionist that it was ready to begin"*. Other participants believed (accurately) that the robots were exchanging information through wireless signals. Several participants thought that communication was occurring via a combination of verbal and electronic signals. The conversation between the participant and the receptionist robot in fact depended on real-time events (though it was heavily structured), while the interaction between the receptionist robot and the guide robot was entirely predetermined. Some participants picked up on this and said that information was transferred via a combination of verbal and electronic signals. P217 said that the guide robot relied on *"wireless data transfer, speech recognition"*, P305 said *"I think the beep signal from the guide robots were some sort of signal otherwise I think through internal communication"*, and P314 thought that the information transfer happened when utterances were *"listened to and compared to a list of things the guide robot is programmed to know"*. Three participants suspected that the entire interaction had been pre-programmed (e.g., *"It could be programmed to just seem like the information was transmitted"*, P310).

5.4. Discussion

5.4.1. Hypotheses Support

We found evidence partially supporting **H1** whereby Stationary Robot Behavior changed participants' overall perceptions of the stationary robot. Participants felt that the stationary robot displayed more warmth in the *social* condition than in other conditions. The stationary robot was also more likable in the *social* condition than in the *representative* condition. Participants also rated the stationary robot as meaner in the *direct* condition than the *social* condition. While we found the stationary robot to be more social and likable, the stationary robot's Behavior did not change the perceived competence of the robot. Furthermore, the likability of the robot was only significantly different between the *social* and *representative* conditions (and not the *direct* condition).

H1 – Participants will perceive the stationary robot to be more social, competent, and likable in the social condition than in the direct and representative conditions.

H2 was partially supported as Behavior affected perceived socialness and likability of the mobile robot (*social* led to higher warmth and likability compared to *representative*), but we were unable to find strong evidence that it influenced perceptions of the robot's competence and trust of the robot. We believe our inability to find an effect for both competence and trust is attributable to a ceiling effect: participants reported high confidence in the mobile robot. They reported an average trust of 6.0 ($SD = 0.98$) on a 7-point Likert scale and rated competence at 6.6 ($SD = 1.69$) on a 9-point scale. Their experience of successful guidance in the practice run might have also biased towards a belief that the mobile robot was capable of completing the task. There was a trend in which the mobile robot in the *social* condition was considered more competent than in the *representative* condition. In this study, we intended to evoke social perceptions about a nonsocial robot by changing the way it interacted with a social robot. Collectively, our results suggest that this was achieved.

H2 – Participants will perceive the mobile robot to be more social, competent, and likable in the social condition than in the direct and representative conditions.

We found some support for **H3**. The type of Information Transfer changed the effect of Stationary Robot Behavior on the perceived relationship between the robots. Ratings of stationary robot competence trended higher

H3 – Information Transfer will have an effect on participants' perception of both robots.

in the *reciting* condition than in the *silent* condition, but this is not sufficient evidence to support **H3(a)**. Participants also felt that the stationary robot was less mean and more likable in the *reciting* condition than in the *silent* condition. There was also a trend where participants reported being more wary of the mobile robot in the *silent* than the *reciting* conditions. Together, these results support **H3(b)** and prior research [12] stating that people do not like covert and silent communication between robots. While we found an effect of *explicit* Information Transfer on the perceived robot relationship, how it affected the exact nature of the relationship was unclear.

H3(a) – Participants will perceive the robots to be more competent in the reciting *condition.*

H3(b) – Participants will perceive lower competence in the mobile robot and be more wary of and disturbed by the stationary robot's behavior in the silent *condition.*

While **H4** was supported in that participants felt the robots in the *social* condition had the best relationship, the effect of Stationary Robot Behavior on perceived relationship was affected by the form of the information transfer. The difference between the *social* and other Behavior conditions was more evident in the *silent* condition. We hypothesize that the absence of similar effects in the other Information Transfer conditions may be due to participants inferring a long-term relationship between the robots after seeing them hold a longer interaction. The open-ended responses also suggest that more participants perceived an equal relationship in the *social* Behavior condition.

H4 – Participants will be more likely to see the robots as each other's equals in the social *condition.*

When asked to choose a preferred mobile robot, most participants picked the robot in the *social* condition and found the mobile robot in the *representative* condition to be the least preferable. This provides support for **H5**.

H5 – Participants will have a higher preference to work with the mobile robot they encounter in the social *condition.*

5.4.2. Social Interaction Between Robots

We found on average participants perceive the robots as more likable and social when they observed a social interaction between the two robots. Multiple participants wished that they had seen the robots treat each other the way humans do—that is, even more cordially—but appreciated that the interaction was social. We found a preference for socialness, even for functional robots that lack social capabilities. However, a few participants also commented that the social interaction was unnecessarily long and they might prefer a streamlined interaction.

> *"The dialogue between the robots went on a bit longer. It didn't bother me, but I could see people get irritated with it."* - P210 - explicit/social

We believe the interaction can be optimized for brevity and socialness.

5.5. Limitations

In multiple sessions, the system had difficulty parsing the participant's speech and required them to repeat their request. Often, participants simply raised their voice in response to a delay, and that solved the issue. However, in a few cases, the experimenter had to step in and ask the participant to speaker louder or rephrase their answers. While we controlled for the potential effect of these mistakes through the inclusion of *system mistakes* as a covariate, it is likely that it still added noise to our study. Future work can explore how failures of one robot influence participants' trust and confidence in another robotSince this work was published, Reig et al. [57] investigated effects of person transfer (*call* condition in the paper) using one of the recovery strategies when encountering a failure. They found participants preferred the robot fixing itself over involving a second robot.. In designing the experiment, we knew that technical problems might emerge, and we considered having a human-in-the-loop or using a Wizard-of-Oz design to avoid such issues. However, we believe it is crucial for HRI studies to utilize real, operational technologies and systems: they better reflect how HRI practitioners may use findings in the field, and they have the potential to expose findings and insights that cannot be captured by other methods [106]. Lastly, we did not include gender as a variable in our analysis because we had fewer male participants. We do not believe that gender influenced the results, but future work should test for gender effects.

Although we counterbalanced the order of the within-subjects conditions and attempted to control for novelty effects and comfort level with a practice run, there was still potential for learning effects during the study. Some participants may have picked up on the phrasing that worked well with the natural language pipeline. We used this design in spite of inherent learning effects because it allowed us to account for individual differences in people's perceptions of social robots.

5.6. Conclusion & Contributions

This work demonstrated that the interaction between robots when they transfer a user is not a trivial design problem, but an important aspect of a smooth person transfer. Through the expression of information exchange and designing social interactions between robots, we can instill confidence in the robots and change how users perceive their abilities. Having robots treat each other in ways that are consistent with human social expectations was the most preferred form of robot-robot interaction. However, it is clear there are subtleties that require careful designs.

Our work has design implications for facilitating transfers between robots. Besides following human social expectations, users may appreciate acknowledgment that certain information has passed between robots. Our results suggest that reciting the information aloud or simply acknowledging that a transfer has occurred are better than no confirmation in terms of establishing perceptions of robot competence. This also validated prior work by Williams et al. [12] which found that silent communication between robots was perceived as creepy. This emphasizes that when robots are transferring a user request from one robot to another, it is important for them to at least acknowledge that the transfer has occurred.

Part 3: Realizing Person Transfer

Interactive System 6.

6.1 Overview 63
6.2 Related Work 64
6.3 Overall Design 65
6.4 Sensory Layer 68
6.5 Execution Layer 69
6.6 Behavior Layer 70
6.7 Storage Layer & Analysis 72
6.8 Discussion 73
6.9 Conclusion & Contributions 73

6.1. Overview

Research in this chapter was conducted in collaboration with Prithu Pareek and Aaron Steinfeld.

In the following three chapters, we report on our efforts to realize person transfer in not only a controlled laboratory setting, but in the field. While we used a complete system in Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer) to investigate effects of verbal communication, the system was specifically made for the study with multiple limitations. In that study, we asked participants to stay in one side of the room that was sectioned off by tape on the floor[1]. The mobile robot's goal position was preprogrammed and the robots turned and faced each other by replaying saved robot joint keypoints. While suitable for the study, this system would not have worked in a real world environment with high variability in people's positions, actions, and goals.

1: These sorts of constraints are common in robotics settings where the robots are "fenced" off from people (e.g., factories).

To address this, we revamped our system and further extended its capabilities. Our new system's design and implementation drew on our own experience working with similar systems and other prior work in this area [25, 107, 108]. Interactive human-robot systems often use custom-designed control systems unique to the robotic platform and task, as we did in Chapter 5. While HRI researchers have slowly converged on using ROS for their underlying message passing framework, the design and implementation of a full stack HRI application remains in flux.

Our system was built with the following capabilities in mind:

Scene Understanding The system needs an understanding of the situated scene to make an informed decision about its actions and goals. For our use case, it needed to track people and robots in the scene to understand who are interacting with our robots and their relationships.

Multi-Modal The system needs to handle inputs and outputs in multiple modalities. For our use case, it needed to handle not only user speech and movement, but also robot speech generation, movement, and gaze[2].

2: We used the head orientation as a proxy for gaze in our study.

Interruptibility & Robustness The system needs to be robust to both unplanned human behavior and system failures. For our use case, it had to handle system failures and changes in interaction flow.

Platform Invariant The system needs to work across platforms and locations with minimal changes to the shared components and interaction scripts. For our use case, it needed to handle two different environments and different robots.

6.1.1. Chapter Summary

In this chapter, we will deep dive into the architectural choices and implementation details of the system. First, we will provide a brief overview of some interactive systems used in HRI and background on the underlying frameworks used in our system. We will then present the overall design of the system and its three main layers: “sensory”, “execution”, and “behavior”. We review each layer and their internal components. We also describe additional tools and systems used to analyze our data. Lastly, we discuss lessons learned and limitations of our system.

6.2. Related Work

6.2.1. Building Interactive Systems in HRI

Early examples of interactive robotic systems include museum robots [109, 110], custom robots built for grand challenges [111], and robot receptionists [112]. Often these systems served as limited, proofs-of-concept that can engage, avoid obstacles and move in complex spaces, and complete certain tasks. Since then, multiple papers have been published describing different ways to design and create interactive HRI systems [107, 113, 114].

Another recent trend in this area is creating tools to enable end-user programming of social robots. These works often integrate efforts from the Learning from Demonstration (LfD) literature, in which people teach robots new skills through physically moving the robot or teleoperation. An example of this is Code3 where the authors integrated keypoint LfD with Scratch to rapidly develop new applications on a PR2 robot [115]. More recent work has explored using physical objects such as arrows [116] and figurines [117] or design methods like bodystorming [118] to support the development of interactive systems with non-developer human users.

6.2.2. Our underlying frameworks

We used three different frameworks in our system: ROS, \Psi , and IPC & TDL.

ROS

Robot Operating System (ROS) [119] is software development kit that consists of a robust peer-to-peer message passing framework and a collection of libraries and tools to support development of robotic applications. The large ecosystem of applications, wrappers, and libraries in ROS has enabled it to become the de-facto robotic framework used by research labs and companies.

\Psi

Platform for Situated Intelligence (\psi) [120] consists of a time- and latency-aware message passing runtime, collections of components that support various human-AI interactions, and a visualization platform for debugging and annotation. \Psi has been used primarily in the multi-modal community to handle the processing of data [121].

IPC & TCM: Roboceptionist Implementation

As part of our study, we utilized the existing infrastructure for Roboceptionist. Roboceptionist was one of the first long-term deployed social robotic platforms [112]. Roboceptionist consisted of multiple programs that communicated with each other through Inter Process Communication (IPC) [122]. The two primary programs were *Expression* and *Robocept*. *Expression* was used to display the expressions of the robot, and *Robocept* was the program controlling the interaction flow of the whole experience. Both were written with Task Descriptive Language (TDL) [123], an extension of C++ that adds additional syntax to support asynchronous interactions. TDL adds syntax to support periodic events (e.g., making fake phone calls every few minutes) and to control flow callbacks (e.g., calling certain functions when a prior function is completed). TDL has also been used to write components in other HRI systems such as task manager in HRI/OS [124].

Here, we provide a brief overview of how Roboceptionist works internally. Please refer to Gockley et al. [112] for a more complete description. Using a laser scanner[3], the system prompts users for inputs if it detects the person is in a predefined range and might be interested in interacting with the robot. Users conversed with the robot through the keyboard and screen in front of Roboceptionist. The text inputs were then parsed using VAINE, a modified version of AINE—a pattern matching language parser. VAINE attempts to match the input with a predefined list of sentences[4]. The results are either immediate behavior outputs (asking about "bathroom" will lead to "segue_bathroom") or structured output that are further processed (e.g., "~D Zhi Tan ~D" will be returned if VAINE believes the user was asking for directions to my office). These outputs are then processed by the Roboceptionist program that decides what kind of behavior it should do depending on past interaction and results from external service requests (e.g., current weather, directions to office). The *Robocept* program then sends the requests to the *Expression* program to execute its responses. These requests can be either predefined expressions (speech + facial movements that were predefined) or direct commands (e.g. "speak("Hello")"). The expression program then parses the expression and runs them. The current Roboceptionist character implementation is named Tank [95].

3: In our final system, we used depth cameras in the scene instead of the laser scanner.

4: For example: "* NASA *" will match with any sentence that contains the word "NASA". When this match is found, it returns "~C (self : NASA)".

6.3. Overall Design

We took a modularized and layered [108] approach to our overall system design. This approach allowed us to test each component individually and easily change the robot platform based on the study needs.

On a conceptual level, our system can be separated into four parts. The first three parts are the three layers that sense, control, and operate the robots in real time. They consisted of:

Sensory Layer Handling of perception of the world and robot states.
Execution Layer Handling of the decision making process of our system.
Behavioral Layer Handling of how behaviors are executed on physical robot platforms.

The last part focused on the post-study processing, debugging, and analysis of the data.

We used the aforementioned frameworks for different parts of the pipeline. \Psi was used primarily at the sensory layer, for data storage and analysis, and for camera extrinsic calibration[5]. ROS handled the interaction merging, message passing between the frameworks, decision making, and the execution of robot actions. Lastly, as we used the Roboceptionist platform as the agent in our study, we communicated with the Roboceptionist components using IPC and reused multiple existing Roboceptionist components in our study. While having multiple frameworks communicating with ROS increased the complexity of the system, we relied on the strengths of each framework as they allowed us to use existing system (IPC), take advantage of open-sourced robot system components (ROS), and visualization and analysis tools built primarily for human behavior analysis (\Psi).

5: More details about the calibration procedures are available in Appendix B (Multi Depth Camera Calibration)

An abstracted view of our real time system and the associated frameworks is shown in Figure 6.1

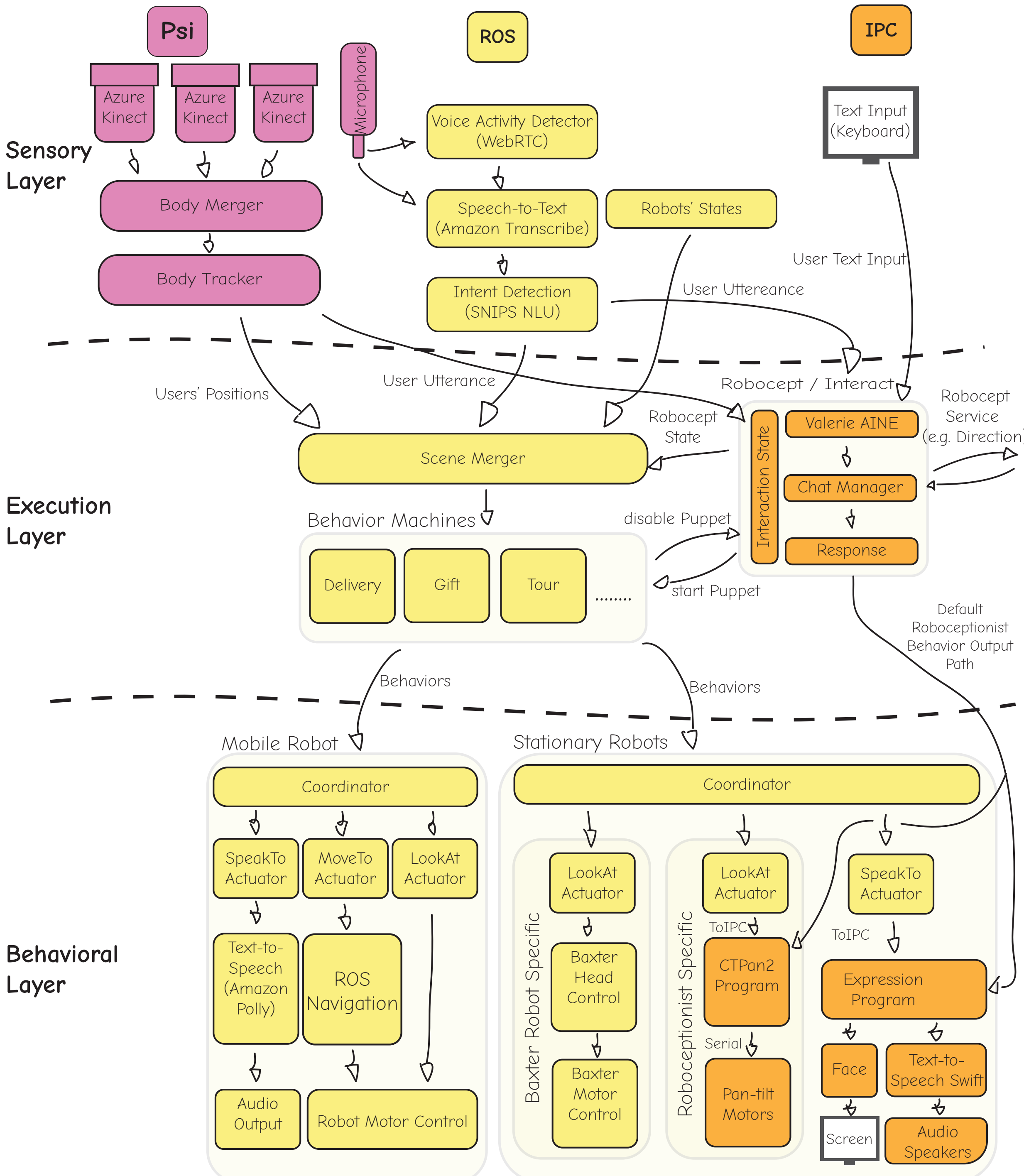


Figure 6.1.: Overview of our system.

6.4. Sensory Layer

The sensory layer handles the inputs from the world, immediate processing of the data, and extraction of features in the scene. In both of our study environments, the environment was augmented with multiple depth cameras and human user input devices. The depth cameras (3 Azure Kinect cameras and 1 Kinect V2 camera) not only provided RGB-D images of the scene, but also tracked body skeletons in the scene. The body skeletons were generated by body tracker modules provided by Microsoft[6]

6: Both cameras detected the skeleton through their depth images. Kinect V2 used a random forest approach, whereas Azure Kinect used a deep neural network approach.

All body skeletons detected by the cameras were fused and tracked by the system. We used multiple heuristics (e.g., euclidean distance difference, confidence of joint detection) to combine skeletons with missing joints, filter out invalid skeleton bodies[7], and track people when they were occluded in certain views of the cameras. Our system also kept a persistent track of the body skeletons in the scene and assigned a permanent ID to them. The combined tracked body skeletons were then sent to ROS and shared with ROS Nodes that required knowledge of user position in the scene. The tracked skeletons were used in:

7: For example, in the lab, we excluded body skeletons that were less than 1m in height and those detected in the area where the experimenter was standing. The height restriction was included to avoid cases where the mobile robot was misidentified as a person. This was an intentional bias we introduced into the system.

Decision Making The flow of the interaction depends on where people are in the scene.

Roboceptionist Instead of the original laser scanner module, we used the base position of the people in the scene to decide if Tank should interact with them.

Mobile Robot Navigation Our mobile robot navigation stack used the position of the people to update the costmap of our planner and intelligently avoid people and group formations. We included more detail in Appendix C (Social Navigation System).

In both the laboratory and the field, participants could converse with the robot through a microphone. The microphone was placed to the side of the stationary robot with the gain tuned to the environment. The audio was filtered by Windows proprietary audio filtering software to remove background noise and enhance voice for speech recognition. The audio was then sent to a Voice Activity Detector (VAD) that determined whether speech had occurred. The audio was sent through the natural language processing (NLP) pipeline only when the VAD believed the participant was speaking, none of the robots were speaking, and the user was not typing. While this made the robots unable to recognize cross-talk, we believe it was a reasonable trade-off to decrease false positive speech recognition. In the NLP pipeline, the audio data was first sent to Amazon's cloud-based speech-to-text recognizer[8]. The speech result was then sent to our local Natural Language Understanding component that extracted the speech's intent using SNIPS NLU [97].

8: https://aws.amazon.com/transcribe/. To ensure participants' privacy, we opted out of Amazon's AI Improvement Policy.

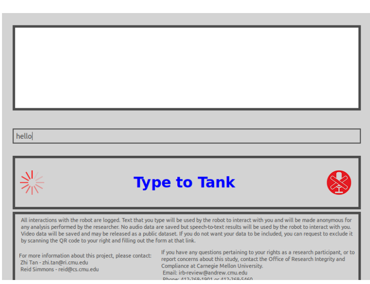


Figure 6.2.: Screenshot of the textinput screen.

In the field study context, participants could also interact with the robot through the existing text-based keyboard system (interface shown in Figure 6.2). The inputs from the keyboard system were sent directly to both the Roboceptionist component and the data store for post-study analysis. To reconcile the two types of input system, the speech system was disabled when keyboard inputs were detected and a mute icon was shown on the screen to cue users that the microphone had been disabled.

6.5. Execution Layer

The second part of the system is the Execution layer. It understands, decides, and controls the actions of both robots. It first creates a snapshot of the scene, decides the appropriate action based on the interaction, and sends the actions to the Behavior layer for execution.

6.5.1. Interaction Scene Generation

As a multi modal and complex interaction, our system needed to react to inputs from a variety of sources such as speech input, movement changes, and positions of the robots. Instead of individually subscribing and listening to each input source, our system contained a **scene merger** component that listened to all the different sources, processed the data, and combined it into a single scene object for decision making process. This scene object was updated as the merger received new data.

The merger receives low level information such as human skeletons (from the body tracker), robot positions (localization and joint sensing components), and participant inputs. This information was then combined with known information such as robot dimensions and shape to generate more useful information such as the robot's gaze direction and footprint (for navigation and position testing). Using this information, we also derived high-level interaction concepts such as engagement and group dynamics using a simplistic heuristic model. We designated one of the robots as the "main interaction agent" and used a combination of an individual's distance and body angle to determine their engagement and group. Each person was assigned a label of "observer", "interactor", or "bystander" based on their location and whether they were looking at the main robot. Afterwards, we iterated through the people in the scene to determine whether they were in a group with the main robot[9]. If a group was found, its members and center were stored to determine the mobile robot joining strategy and published for navigation. This prevented the robot from moving through the O-space of the group. More detail on this is in Appendix C (Social Navigation System).

9: In our study, we focused on groups with the main robot and ignored other human-only groups. Group detection in complex environments remains a challenge and is an area of active research

6.5.2. Behavior Machine

With the scene information, our system decides what kind of actions and behaviors the robot should do. The system does this through a Behavior Machine. Behavior Machine is a framework we developed that combines the properties of behavior trees [125, 126] and hierarchical state machines. Behavior Machine allows interaction designers to use patterns in behavior trees while also allowing them to use state machine-like transitions. It has built-in interruption handling and is inherently multi-threaded (each state runs on its own thread). The flexibility of the system allowed us to design interactions that are sequential (e.g., robots talking to each other according to script), parallel (multiple robots doing different tasks at the same time), and interruptible (the interaction flow can be changed by user input or experimenter control). Each scenario in our studies had its own Behavior Machine with multiple shared states to keep the interaction consistent. A complete explanation of Behavior Machine, its components, and design choices is included in Appendix D (Behavior Machine).

6.5.3. Behavior Machine in Practice

In all our scenarios, the Behavior Machine took in the scene and decided the actions the robots can take. The scene was updated at a frequency of 10Hz. The update retrieved the newest scene from the merger and saved it in the Behavior Machine's board. We wrote each machine to execute the desired service interactions and reused as many components as possible. In addition to the scene, the machine also reacted to information on other ROS topics through a dedicated ROS subscriber state. This state subscribed to each topic of interest and saved the latest message in the board.

To abstract away the complexity of robot controls and avoid limiting the actions to a specific robot platform, the Behavior Machine controlled the robot by sending high-level behavior requests to the behavior layers. The requests included both low- and high-level properties. For example, when commanding the robot to look at an object, the machine can specify either a ROS tf frame id or a fixed point in space. These requests were implemented with ROS actions which allowed us to cancel requests if the state was interrupted. The behaviors that we implemented were:

LookAt This is a high-level gaze action. The message consists of the type of gaze (Idle, Fixed, etc). The target can be provided as either a tf frame or a coordinates of a point in space.

SpeakTo This is a high-level utterance action. It consists of text that describes what the robot should say. There is also an optional target field (TF frames only).

MoveTo This is a high-level action that commands the robot to go to a specific position. While the code only allowed us to pass in fixed poses in the environment, a chaining of states would allow us to grab poses from a preprogrammed list of locations.

Besides these behavior actions, there are also pure ROS publisher states that enabled us to publish information on ROS Topics[10]. In the following section, we describe how these high-level behaviors are executed and created.

10: Examples were publishing debugging information or study state to ROS and subsequently passing that to the \psi storage components.

6.6. Behavior Layer

The behavior layer handles the coordination and execution of the high-level behaviors on each robot. As shown in Figure 6.1, each robot has its own behavior layer stack. First, a coordinator component listens to all the behavior requests and redirects them to actuators that create the movements or actions. By abstracting and using high-level behaviors, it allowed us to easily change the robot platform and environment.

6.6.1. Coordinators

Coordinators are components that reconcile the different and potentially competing behavior requests. For instance, a robot might nearly-simultaneously receive a **SpeakTo** request to look at a person and talk to them and a **LookAt** request to gaze at an object. The coordinator uses predefined rules to decide which behavior to execute and wait time between request. For example, an explicit command to gaze at an object is always prioritized over robot speech targets. The coordinator also encodes the timing, such as how long a gaze should last before speech starts. In our studies, we used simple heuristics to determine these behaviors, but future applications can expand upon this and have finer control over the timing and execution of interaction in different modalities.

6.6.2. Actuators

Actuators are robot and platform specific components that physically generate the high-level behaviors. While they could be shared between robots (e.g., speech generation), the majority of our components were unique to each of the robots because different robots used different motors and modalities to create high-level behaviors. For example, all three robot platforms perform gaze actions, but they do so in ways specific to their platforms. The stationary robot moved its head pan joint[11], whereas the mobile robot moved its whole body to face the person or change its eyes' position on the screen face. While the underlying systems for each actuator were drastically different, they used the same ROS action interface which allowed us to easily swap out a robot platform for another by simply changing the actuators. The use of the ROS action interface also allowed us to reason about whether an action had completed and stop behaviors midway.

11: The behavior of Baxter was further enhanced with the movement of its arms to give an impression of a larger movement.

In both the laboratory and field study, we used the same mobile robot and two different types of stationary robot (Baxter and Roboceptionist). Both stationary robot shared the same SpeakTo actuator. The details of the actuators used for each platform are shown in table 6.1.

Table 6.1.: List of actuators for robotic platforms used in our studies.

Platform	Actuator	Description
Stationary Robot (Baxter)	LookAt	We commanded Baxter's head pan angle such that the display faced the person. This was executed by sending a Pan ROS Msg to Baxter.
	SpeakTo	We sent the speech request to Roboceptionist's Expression program that executed the speech request.
Stationary Robot (Roboceptionist)	LookAt	We commanded the Roboceptionist pan joint such that its display faced the person. This was executed by sending a JointState Msg to Roboceptionist's CT Hybrid Module.
	SpeakTo	We sent the speech request to Roboceptionist's Expression program that executed the speech request.
Mobile Robot	LookAt	We first calculated the orientation such that the mobile robot would face the target. We then sent the move command with the current position but new orientation to the MoveTo component.
	SpeakTo	We generated the audio using Amazon Polly and sent the audio to speakers.
	MoveTo	We first checked if the orientation was the only difference. If only the orientation changed, we used a PD controller to rotate the Mobile Robot by directly sending motor commands. In cases where there was change in position, we sent the goal position to ROS Navigation which executed the move command. We modified the ROS Navigation costmaps to account for human positions and spatial formation. More detail in Appendix C (Social Navigation System).

6.7. Storage Layer & Analysis

As our research goals were explored through both qualitative observations (Chapter 8 (Person Transfers in the Field)) and quantitative comparisons (Chapter 7 (Spatial Formation in Person Transfers)), it was crucial to have easy ways to scroll through and analyze the interactions. Our study data was stored in \Psi and ROS 1 Bags. The features we were interested in, such as human positions, were stored in PsiStore[12] using their dataset architecture. The dataset feature allowed us to apply the same operations (extracting 2D position, etc.) across all participants and sessions.

12: More details about PsiStore can be found in 4.3.5 of [120]. It deserializes any C♯ object and stores them on disk.

We used PsiStudio[13] to visually inspect the interactions and help debug both latency and interaction issues. We also made a new 3rd party reader for PsiStudio that enabled it to read ROS 1 Bags.

13: More details about PsiStudio can be found in 5.1 of [120]. It allowed us to easily scroll through and replay data which is difficult to do in ROS.

6.7.1. Post-Processing

To analyze data in the lab study, we first trimmed each data session according to the flags set by the system that signified the beginning (when the experimenter informed the participant they could start interacting with the robot) and end (when the behavior machine finished running and the experimenter informed the participant it was the end of the study) of each session. We then calculated the position and orientation data for the participant relative to the main robot. These data were then paired with different states of the study to understand how human behavior changes at key points during the interaction.

To analyze data in the field study, we first identified the time stamps when the participant enter the scene and the end of the interaction[14]. We then extracted each session and saved them in a new data set for analysis.

14: Either the participants left the scene or they were being intercepted by the experimenter

6.8. Discussion

Our implementation and components are likely to be another robotic platform in the long line of robotic interactive systems. As shown in Section 6.2 (Related Work), our approach shared the layered design used in various past robotic systems. Our system contributes to ways others can design and implement multi-robot multi-platform human robot social interaction system.

As with many other implementations, our system relies on the engineer to craft the interaction, and to account for unexpected behaviors or failures. This balance allows for fine control over the high-level behaviors rather than risking uncontrolled emergent behaviors. However, other options are also available, such as cognitive robotic architectures that model cognitive processes and use them to decide behaviors (e.g. ACT-R/e [127], DIARC [128]).

The way our system was designed was highly influenced by the need to eventually transition the laboratory study into the field and switch between two stationary robot platform. Besides certain improvements and bug fixing, the major difference between the two study contexts was the usage of a different Behavior Machine. This was due to the need of a different scenario in the field study context. The abstraction of low-level control with high-level behavior allowed us to design scenarios without thinking about how the robot will create them. In both of our studies, the switch between robots simply required us to change which ROS program was running since both program listened for ROS actions on the ‘robocept/actuator/lookAt‘ topic.

6.9. Conclusion & Contributions

This chapter provided an overview of the interactive system used in the following two chapters. Together with the Appendices, it provides a detailed explanation of how our system works and justifications for the choices we made in designing and implementing it. We believe the contributions of this system design is both the individual system components and the novel integration of all of the components. We demonstrated and discussed a system to realize multi-robot human interaction with person transfers. Furthermore, the way our system was designed to respond to a need to switch between context and robotic platform may lead to a better understanding how we can create intelligent robot systems that are generalizable across domains and machines.

Another contribution of this chapter is the open source code for the system components. Excluding the Roboceptionist components, all code can be found at the following Github https://github.com/CMU-TBD/tan-dissertation-2022 .By sharing the system with the wider research community, we enable others to use all or parts of the system in their human-robot interaction system. This also allows our system to serve as an example for others, especially students, of how complex systems can work. We believe our codebase may help students better understand how such systems work. Building complex interactive systems like this is hard. Having both a blueprint (this document) and code will enable others to learn from our work and potentially find better methods to deploy this interaction experience.

7. Spatial Formation in Person Transfers

7.1 Overview 75
7.2 Robot Group-Joining Strategies 76
7.3 Method 78
7.4 Results 83
7.5 Exploratory Analysis . . . 90
7.6 Discussion 91
7.7 Conclusion & Contribution 93

7.1. Overview

Research in this chapter was conducted in collaboration with Elizabeth Carter, Prithu Pareek, and Aaron Steinfeld.

During immediate person transfers[1] a new interactant joins an existing one-on-one dyadic interaction and converts it into a group interaction. While the act of joining and exiting a social group may seem trivial, prior work has shown this is a complex act ([75, pg. 234] and Section 3.3 (Spatial Formations & Proxemics)). When a new party joins a group, the configuration of the group typically changes as members attempt to position themselves equidistantly from each other [75, pg. 220] and the group's center [129].

1: As discussed in Chapter 4, a transfer can happen immediately or with a gap.

Like a human, a robot summoned to join an existing dyadic interaction between a person and another robot needs to **(1) determine where it needs to position itself relative to the group** and **(2) understand how people may react to its choices**. As the mobile robot joins the interaction, the user might have to reorient to the new group. Similarly, as the robot departs, the user could either follow the new robot or rearrange themselves back to the original configuration.

We saw similar changes in position in Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer) (Figure 7.1). As the mobile robot arrived, some participants changed their orientation to face the robots.

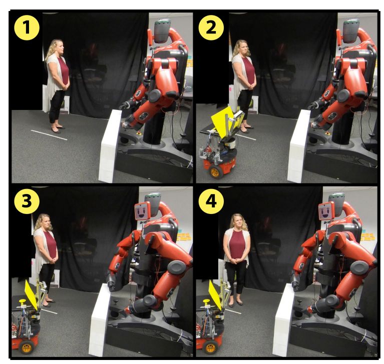


Figure 7.1.: Example F-formation in Chapter 5. Left to right: (1) Participant faces the stationary robot, (2) Mobile robot arrives, (3) Stationary robot turns to talk to the mobile robot, (4) Participant shifts her orientation to ensure a shared space.

To answer these questions, we investigated how a mobile robot should join an existing interaction between a user and a stationary robot through a laboratory study. We first identified two human group-inspired positioning strategies and designed an algorithm to realize them. We then compared them to one another, and also with an Improper strategy that violated human social expectations. The Improper strategy allowed us to explore cases where optimal positioning may be impossible or a system fails to find a reasonable position. As multi-robot interaction and person transfers can happen under a variety of scenarios, we also varied the services that the user received (e.g., obtaining information, picking up items) to explore how tasks might impact spatial behavior. While certain tasks such as pick up and drop off will require people to move, we are interested in how a person's position change overtime during the interaction. For example, after picking up an item from the second robot, would the person stay with the robot or move back to the first robot?

7.1.1. Chapter Summary

In this chapter, we first explain the different joining strategies and an algorithm to robustly pick the position to deploy the robot. We then describe the study we conducted to investigate the differences between mobile robot joining strategy and how the type of services influenced participants behavior and perceptions. Lastly, we present our findings, identify unresolved research questions in this area, and discuss how insights from this chapter can inform future research and robot interaction design.

7.2. Robot Group-Joining Strategies

We created three group-joining strategies. The first two strategies were based on human group spatial patterns described in prior work [75, pg. 213]. The last strategy simulated a failed joining where the mobile robot positioned itself far away from the existing interaction, violating human group norms. Specifically, the three strategies were:

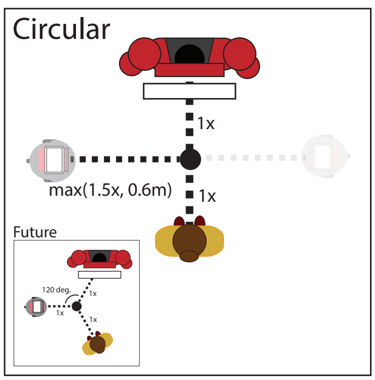


Figure 7.2.: Illustration of the Circular strategy.

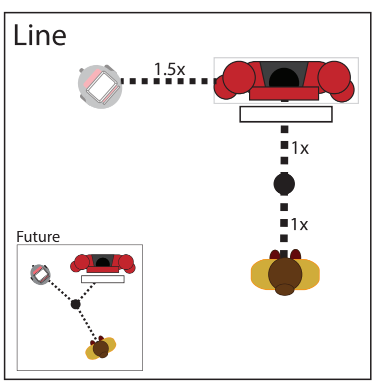


Figure 7.3.: Illustration of the Line strategy.

Circular – This strategy was based on the circular arrangement observed in human-human interaction [75, pg. 215]. After projecting a line between the person and the stationary robot, the system extrapolated a perpendicular line starting from the midpoint of the projected line. The mobile robot then picked the point along the perpendicular line that was 1.5x (multiplier) the distance between the person and the midpoint. We chose that multiplier to create a circle upon arrival with a radius equal to the person's original distance to the midpoint and the circle's center being the centroid of all three interactions (Figure 7.2 inset). We also forced the mobile robot to maintain at least 0.6m between itself and the midpoint to speed up collision checking as any distance smaller than 0.6m likely meant it was in the boundary box of the stationary robot.

Line – This strategy is based on the side-by-side arrangement described in prior work [130] where two individuals are positioned side-by-side while observing an event or a speaker. After projecting a line between the person and the stationary robot, the system extrapolated a perpendicular line starting from the stationary robot's endpoint. Because the stationary robot had a wider profile due to its arms, the line started at the tips of the downward facing arm, instead of the central base. The mobile robot then picked the point along the perpendicular line that was 1.5x (multiplier) the distance between the person and the midpoint (Figure 7.3). In this arrangement, the mobile robot and stationary robot form a line that faces the person.

Improper – This strategy is similar to **Circular** but with the distance between the second robot and the midpoint of the line segment formed by the human and the first robot either increased by a factor of 3 or by 2.5 meters, whichever was larger (Figure 7.4). This condition created a distance in which the second robot was too far away to form an O-space, even by the standards of the COVID-19 pandemic when people were accustomed to conversing at distances of 2 meters. The selected distance lied on the far phase of the social distance in proxemics which was further than the close phase of Social Distance (up to 2m), where gathering and work collaboration usually happens [131, pg. 115].

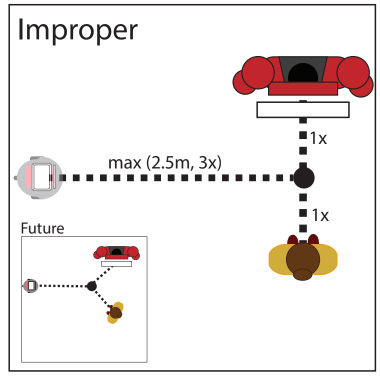


Figure 7.4.: Illustration of the Improper strategy.

7.2.1. Operationalizing Algorithm

In practice, a position chosen according to one of these three methods may not always be valid; it could overlap the position of another interactant or be outside the boundaries of the study zone. Furthermore, each method also outputted two valid positions for the second robot (the desired arrangement could be formed on either side of the human). To account for these two challenges, the strategies were operationalized through the 'PoseSelection' algorithm[2] described in Algorithm 1. The algorithm returned the best valid pose based on the initial scene information (location of users, robots, etc.). If no poses were valid, the algorithm adjusted the multiplier (e.g., 1.5 could become 1.6) and minimum required distance (e.g., 2.5m could become 2m) to search for alternative positions. In a nutshell, the algorithm performs a line search along the perpendicular line to find a valid position.

2: In the actual implementation of the code, the algorithm was realized through multiple states and consisted of additional checking and verification. The described algorithm is a abstracted version.

In the case of multiple valid candidates, the valid point closest to the second robot was chosen. Lastly, a backup position (1.4m to the right of and 0.2m in front of the stationary robot with it facing the robot and the person) was used if no valid position was found after multiple iterations. In actuality, the backup position was used once during the whole lab study (P12). The algorithm also generated the orientation to point the mobile robot towards the center of the group.

Algorithm 1 Mobile Robot Pose Selection Algorithm

1: **procedure** PoseSelection(scene, lookAt = "center")
2:     pos, attempt $\leftarrow$ None, 0
3:     param $\leftarrow$ initial parameters ▷ e.g. multiplier, min distance
4:     **while** attempt $\geq$ limit **do**
5:         c, d $\leftarrow$ center, distance (of dyad) from scene.
6:         param $\leftarrow$ UpdateParamByAttempt(param, attempt)
7:         choices $\leftarrow$ MobileRobotStrategy(c, d, param)
8:         $\text{pos} \leftarrow \underset{c \in choices}{\arg\max} \begin{cases} -1 & \text{not valid(c)} \\ e^{-1 \cdot \|m-c\|} & \text{otherwise} \end{cases}$
9:         **if** valid(pos) **then**
10:             ori $\leftarrow$ CalculateOrientation(c, d, pos, lookAt)
11:             **return** (pos, ori) ▷ Target pose
12:         **end if**
13:         attempt $\leftarrow$ attempt + 1
14:     **end while**
15:     **return** backup position
16: **end procedure**

In the algorithm, line 8 describes the scoring function that we extend further with other criteria in Chapter 8 (Person Transfers in the Field). $\|m - c\|$

calculates the distance between the position being tested, c, and the mobile robot, m.

7.2.2. Execution & Trajectory

Once a pose was selected, we generated the trajectory using a lattice-based global planner [132] in the ROS Navigation Stack [98]. The planner strongly penalized in-place rotation and generated trajectories that arrived at the target position with the correct orientation. Our planner was also fully aware of the position of the person and the interaction group. It penalized paths that passed through the O-space or were too close to the person. We describe our navigation stack in more detail in Appendix C (Social Navigation System).

7.3. Method

To evaluate the strategies in different contexts and understand how people position themselves during transfers between robots, we conducted an in-person[3] laboratory study where participants experienced transfers in various scenarios. The study was approved by our university's Institutional Review Board.

3: Due to the COVID19 pandemic, we considered other study methodologies such as simulations or online video study. We chose not to as we were concerned about whether it is possible to capture the motions that are manifested subconsciously. Our concern was supported by our results where we only found differences in the behavior data but not self-reported data.

7.3.1. Study Design

We created a 3x4 mixed design study where we manipulated the mobile robot's joining strategy (between-participants) and the service scenario (within-participants). Scenario order was counterbalanced using a balanced Latin square. The mobile robot's joining strategy was chosen as the between participant variable because if the participant observed multiple different positioning strategy, participants may interpret the difference in position as a system failure and act differently. In the study, the robot goal setting, navigation, interaction, and speech recognition were autonomous with the option for minor adjustments by the experimenter as needed (see Subsection 7.3.5 (Interaction System)).

7.3.2. Service Scenarios

We designed four simulated service scenarios and asked the participant to do a task that was appropriate for that setting. To avoid confounds, we kept the interaction dialogues as conceptually similar as possible and attempted to have the difference be primarily in the user action and service requirement. For instance, in all scenarios, when the mobile robot joined the interaction, the stationary robot always announced the mobile robot's arrival before they greet each other. Participants were told to use the name "Sam" for scenarios that asked for the participant's name.

Room **Guidance** – Participants were told to imagine that they were at a bank and needed help finding the conference room. They were instructed to begin the interaction by approaching the stationary robot. Then, the stationary robot asked the participant for their name and what help they required. It then summoned a mobile robot, which approached the dyad according to the group joining strategy. The stationary robot then restated the participant's request to the mobile robot. The mobile robot then asked the participant to follow it as it moved towards the door of the room. This scenario demonstrated a service transfer from a stationary robot to a mobile robot due a difference in capabilities.

Gift **Delivery** – Participants were told to deliver a small gift box to their friend who lived in an apartment building with a stationary robot doorman. After the participant approached and made their request, the stationary robot informed the participant that their friend was away but that the gift could be delivered to their doorstep. It then called the mobile robot, which joined the group and asked the participant to place the gift in the bin on top of its frame. The stationary robot restated the participant's request to the mobile robot, and the mobile robot departed with the gift. The stationary robot then told the participant that it would notify their friend about the delivery. This scenario demonstrated a service transfer from a stationary robot to a mobile robot due to a difference in capabilities and the participant physically interacting with mobile robot through dropping off the gift.

Instruction **Failure** – Participants were told to get help finding the office of a professor at a university campus from the stationary robot receptionist. The stationary robot then announced it had not updated its map and would summon a mobile robot who had the instructions. Once the mobile robot arrived, the stationary robot told the mobile robot that it had experienced a failure and asked it to provide the directions to the participant. After providing the instructions, the mobile robot left, and the stationary robot asked participants if they needed anything else. This scenario demonstrated a shift from 1-to-1 interaction to group interaction due to a failure.

Package **Pickup** – Participants were told to imagine that they lived in an apartment building with a stationary robot doorman. After asking the participant to provide their name for security purposes, the stationary robot informed them of a package delivery and summoned the mobile robot, which carried the package. The mobile robot arrived and asked the participant to retrieve the package from its bin. It then left while the stationary robot asked if the participant needed anything else. This scenario demonstrated a service transfer from a stationary robot to a mobile robot due a difference in capabilities and participant physically interacting with mobile robot by taking the package.

Prior to the four main service scenarios, participants also experienced two "Familiarization" scenarios to reduce novelty effects and help calibrate their expectations.

Guidance **Drop-Off** – Participants were told to imagine that they were at the airport and were following a mobile robot to their gate. Participants followed the mobile robot for a few minutes to a predefined location to the right of the stationary robot[4] The mobile robot then informed the

4: The position was 0.8m to the right and 0.8m to the front of the stationary robot.

participants that they had arrived at the gate and that the stationary robot would continue the service. The mobile robot left while the stationary robot inquired about participants' flight meal preferences.

Airport **Survey** – Participants were told to imagine they were at the airport and instructed to approach the stationary robot. The stationary robot greeted the participants and administered a survey that asked about their experience.

7.3.3. Hypotheses & Evaluation

Because there is sparse research on this topic, our analysis consisted of both hypothesis testing based on prior work in human behavior and an exploratory analysis grounded in research questions that arose during the study. We predicted:

H1 *Participants will move to rearrange themselves in the Improper condition more than in other joining strategy conditions (across all service scenarios).*
H2 *The effects of repositioning will be larger in scenarios where humans are required to move.*
H3 *Participants will rate the interactions in the Circular and Line as more comfortable and easy than in Improper.*
H4 *Within Improper, participants will rate scenarios with physical actions to be less comfortable and harder than other scenarios.*
H5 *Participants will rate the mobile robot as less competent in Improper.*
H6 *Participants will rate the mobile robot's goal position in Improper as less socially appropriate.*

In addition to our hypotheses, we explored how different aspects in the interaction could influence human spatial behaviors. Our research questions were:

RQ1 *How do social behavior and cues influence participants' spatial behavior?*
RQ2 *How does physical action during certain scenarios influence participants' spatial behavior?*

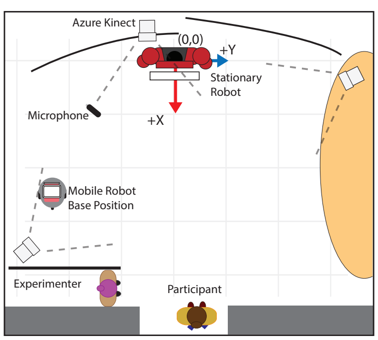


Figure 7.5.: Top down illustration of the study location.

7.3.4. Study Physical Environment

The study was conducted in an enclosed lab space at Carnegie Mellon University's Pittsburgh campus. The study space was separated from the lab's office space by black cloths. Furniture in the space was also cleared out or pushed to the corners in order to increase the sense of space. The main study area measured about 5.5m wide and 3m long.

In this study, the stationary robot was a Rethink Robotics Baxter robot (1.78 meters tall). The mobile robot was a Mobile Robot P3DX (1 meter tall) with a custom structure on top housing a tablet with a screen face[5] and an open transparent bin to carry items. The mobile robot's base location was in the same room, off to one side[6]. The positions of the participants were captured by 3 Microsoft Azure Kinect sensors that generated a skeleton

5: This was different from Chapter 5 where it did not have a face. The mobile robot also had speech capabilities.

6: We had the robot in the same room since we wanted to limit the study to a single enclosed space for COVID safety

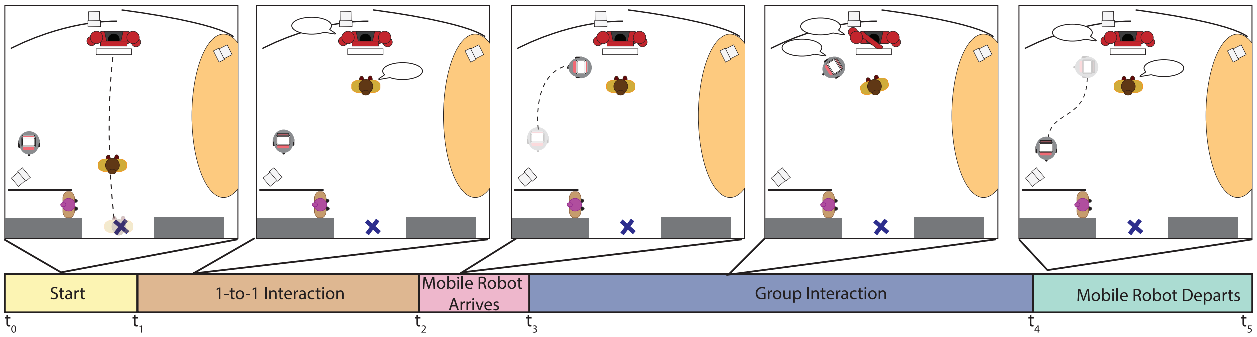


Figure 7.6.: Illustrations of the stages of the interaction.

view of them in the scene. A microphone hung from the ceiling to capture speech. The study area is illustrated in Figure 7.5.

7.3.5. Interaction System

The study used the system described in Chapter 6 (Interactive System) to manage the interaction flow and execute each scenario. Each scenario was written as its own Behavior Machine, but they shared similar components (e.g., experiment states and initialization sequences). The machine also included an experimenter bypass. The bypass allowed the experimenter to use a controller to provide additional motion controls for the mobile robot, progress the interaction when the system encountered speech-to-text failures and unplanned responses from the participants. While the machine was capable of automatically running the whole study, we chose to manually progress certain parts of the interaction, such as when the user placed an item in the mobile robot's bin or took an item from it, so that the interaction was maximally reliable, safe, and consistent across participants. The same experimenter was present in all study sessions and attempted to maintain the same demeanor and responses with all participants, for all conditions. Due to the complexity of the system, some participants experienced sessions with minor hiccups, such as delay in the interaction, a delay in robot gaze behavior, speech-to-text issues where the system completely misunderstood a response, or the mobile robot having trouble planning an exit trajectory due to participants standing too close.

7.3.6. Study Procedure

The experimenter first administered COVID-19 screening questions and obtained informed consent from the participant. Then, the experimenter introduced the participant to the study and explicitly told them that there was no restriction on how they could move and where they could stand. They then experienced the *Guidance* and *Survey* scenarios, followed by the four main scenarios. Participants also completed a short questionnaire about the experience at the end of each scenario.

In all four main scenarios, the experimenter gave the participant an explanation of the scenario, where they are, and their goal and objective.

Participants were asked to stand on a blue X-mark near the entrance to the room (Figure 7.5) and told to ignore the experimenter, who stood behind a divider. The experimenter then started the session.

In all scenarios, the system automatically segmented the interaction into stages (Figure 7.6) based on the interactions and movement of the robot(s). There were four stages in the transfer:

1-to-1 Interaction – This stage starts when the stationary robot begin interacting with the participant after recognizing that the participant was in-range and have stop moving. In this stage, participants were only interacting with the stationary robot.

Mobile Robot Arrival – This stage starts when the the stationary robot tells participants the mobile robot was summoned and can physically see the mobile robot moving.

Group Interaction – This stage starts when the mobile robot arrives at its goal position and joined the interaction. All interactors (robots + participant) then interacted together. Any physical actions (picking up and dropping off item) also happen in this stage.

Mobile Robot Departure – This stage starts when the mobile robot begins to leave the interaction. Except for *Guidance*, participants continued to interact with the stationary robot similar to 1-to-1.

In the study, the stationary robot was called "Red Robot" in all scenarios. To give an illusion of difference between scenarios, the mobile robot has a different color and name ("[COLOR] robot") in each scenario. The assignment of color was randomized for each participant. At the end of the study, the participant filled out a final questionnaire that included additional questions about the overall study and demographic information.

7.3.7. Measures

Participants' positions were recorded during each scenario and analyzed with the stationary robot as the point of origin with the positive X-axis pointing forward and the positive Y-axis pointing to the left (Figure 7.5). For subjective measures, we used a combination of custom questions and established scales. At the end of each scenario, we asked the participant to rate their comfort interacting with the robots, the task difficulty, and task confidence on custom 7-point scale items. Participants also described what they liked and suggestions for improvements. At the end of the study, participants completed a questionnaire about each robot using the validated RoSAS scale [100], seven custom rating items about spatial behavior of the mobile robot, and perceived distance to each robot. Participants used their preferred distance unit (feet or meters) when providing the estimate on a linear scale with 0.1 resolution. The complete set of questionnaires used in the study can be found in Appendix E (Questionnaires for Chapter 7).

7.3.8. State of Study Environment & Participants

The study was conducted during August 2021 with COVID-19 mitigation protocols still in place. Participants wore face coverings, and all touched surfaces were sanitized before and after the study. Participants were recruited from our community using flyers and our university's research participant recruitment system. 32 of 36 participants were local university students.

We recruited 58 participants, 22 of whom were excluded due to system failures and irregularities during the interaction (e.g., the robot gaze failed, the recording was not captured, time syncing errors). Our study had 12 participants per between-subject condition, for a total of 36 participants with ages ranging from 18 to 68 years. (Table 7.1). 20 out of 36 participants listed English as their first language. The study took about 45 minutes and participants were compensated USD $15.

	Female	Male	Other
Circular	4	8	0
Line	7	4	1
Improper	7	4	1
	Age (Std. Dev.)		
Circular	26.75 (14.2)		
Line	30 (10.71)		
Improper	23 (3.33)		

Table 7.1.: Participant demographics (36 valid sessions)

7.4. Results

Unless stated otherwise, results involving repeated data were analyzed by fitting of a mixed effect linear model using REstricted Maximum Likelihood (REML)[7]. Results involving singular responses from participants were analyzed through a linear regression model. All post-hoc analysis were conducted using Tukey's Honestly Significant Difference (HSD) test.

7: Our model consisted of scenario and group joining strategy as fixed effects with participants nested in scenario as a random effect. We used JMP 16 for our analysis.

7.4.1. Validation

To validate that participants experienced differences in the arrangements, we compared their distances to the mobile robot at the beginning of the group interaction stage for each strategy. Our comparison found that the distance to the mobile robot was significantly different between strategies ($F(2, 132) = 348.41$, $p < 0.0001$). Pairwise comparisons showed that the mobile robot was farthest in the Improper (M = 2.54m, SD = 0.35m), followed by Line (M = 1.35m, SD = 0.32m), and Circular (M = 1.01m, SD = 0.16m), all $p < 0.0001$. The difference in distance met our expectation. We also found significant differences in the self-reported perceived distance to the mobile robot between strategies ($F(2, 33) = 7.40$, $p = 0.0022$). Pairwise comparisons showed that participants perceived the mobile robot to be significantly farther away in Improper (M = 1.21m, SD = 0.58m) than in Circular (M = 0.70m, SD = 0.23m, $p = 0.0057$) and Line (M = 0.70m, SD = 0.18m, $p = 0.0061$).

7.4.2. Subjective Responses

Responses Per Scenario

We found significant difference in participants reported task difficulty ($F(3, 132) = 4.986, p = 0.0026$) and confidence ($F(3, 132) = 3.533, p =$

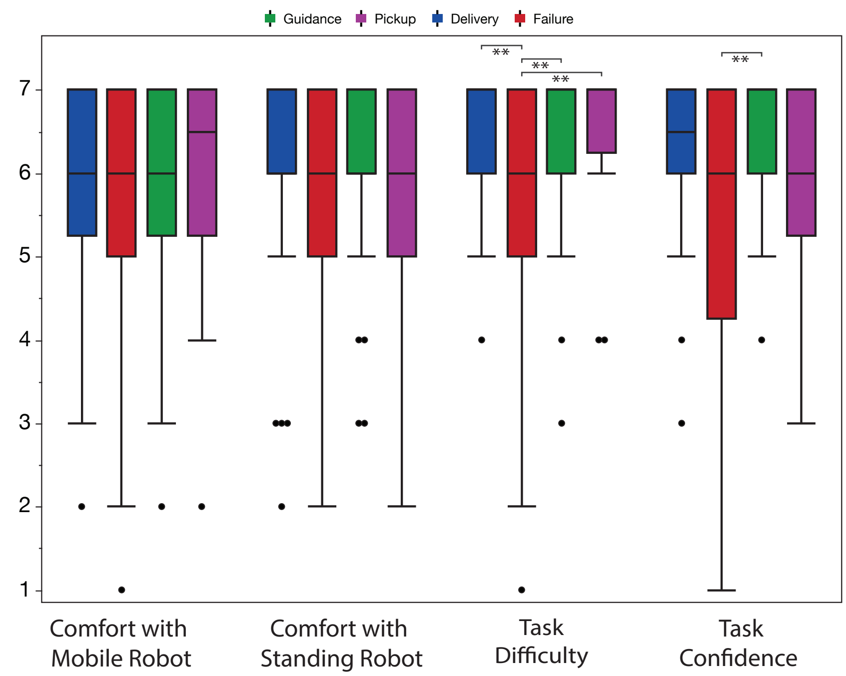


Figure 7.7.: Participants' responses to 4 subjective questions by scenario.

0.0167) by scenarios. Pairwise comparisons showed participants reported the Failure scenario was significantly harder (M = 6.52, SD = 0.73) than the other scenarios ($p < 0.0279$) and they were less confident on what to do next in the Failure scenario (M = 5.47, SD = 1.78) compared to the Guidance scenario (M = 6.36, SD = 1.21, $p = 0.0158$). We believe the differences stem from participants having to remember instructions that spanned two floors and we did not provide a chance for the participants to ask for any repeats. We found no significant differences in participants' reported comfort interacting with each robot.

Overall, participants were very comfortable interacting with the stationary robot (median = 6 , IQR = 1, on a 7-point scale with 7 = extremely comfortable) and the mobile robot (median = 6 , IQR = 2, on a 7-point scale with 7 = extremely comfortable). Even though there was significant difference across scenarios, participants still rated the tasks as extremely easy (median = 7, interquartile range (IQR) = 1, on a 7-point scale with 7 = extremely easy) even for the Failure scenario (median = 6, IQR = 2, on a 7-point scale with 7 = extremely comfortable).

Perception of Robots

To analyze the RoSAS, we first re-examined the three factors of Warmth, Competence, and Discomfort as described in the paper [100]. The items for each factor had high item reliability (Cronbach's α scores: Stationary Robot: Warmth - 0.884, Competence - 0.945, Discomfort - 0.783; Mobile Robot: Warmth - 0.800, Competence - 0.833, Discomfort - 0.816). No significant differences between conditions were found for any factor. An exploratory analysis on the individual items using the Wilcoxon test found no differences.

User Perception of Spatial Behavior

	Social Appropriateness	Predictability
The mobile robot's position made me uncomfortable. (R)	**-0.908**	0.180
The mobile robot stood really close to me. (R)	**-0.528**	-0.137
The mobile robot's position was socially appropriate.	**0.670**	-0.183
The mobile robot movement was predictable.	0.065	**0.998**
I was able to predict the position of the mobile robot.	0.037	**0.604**
The mobile robot position itself too far away from me.	0.015	0.156
I have trouble seeing the mobile robot.	-0.101	-0.079

Table 7.2.: Factor Loading. Items marked with (R) are reverse-coded.

For the questions about the mobile robot's spatial behavior, we conducted an exploratory factor analysis on our 7 items. Three factors explained 72.54% of the variance and had an eigenvalue above 1 [133]. We used Varimax rotation to determine the factor loading for each item and used a cutoff of 0.4 for clustering (Table 7.2). One factor was excluded because it only consisted of one item that did not load onto any factor, and another item was excluded for not loading onto any factor. We determined that the first factor (3 items) described the social appropriateness of the position and the second factor (2 items) described the predictability of the mobile robot's position.

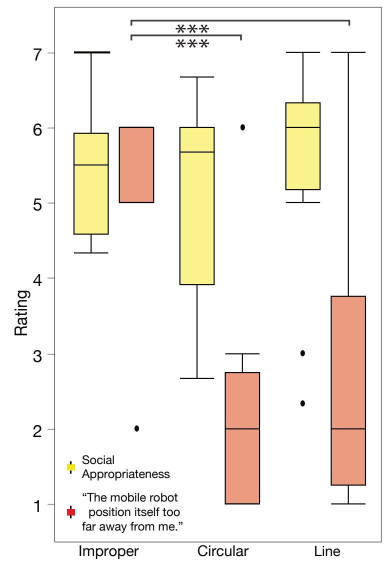


Figure 7.8.: Participant responses to social appropriateness and whether the mobile robot position was "too far" by strategy. *** means $p < 0.001$.

Participants perceived the mobile robot's position to be socially appropriate (M = 5.35, SD = 1.21) but were unsure about predictability (M = 3.66, SD = 1.36). We found a significant difference across strategies for the individual question *"The mobile robot position itself too far away from me"* ($F(2, 33) = 15.496$, $p < 0.0001$). Pairwise comparisons showed that participants reported stronger agreement with the statement in Improper (M = 5.33, SD = 1.15) than in Circular (M = 2.17, SD = 1.40, $p < 0.0001$) and Line (M = 2.75, SD = 1.81, $p = 0.0005$). No other significant differences were found.

7.4.3. Participant Spatial Behaviors

To test hypotheses H1 & H2 – that users would move to compensate for improper spatial arrangements – we compared participants' distances to the center, to both robots, and their ratio at the the beginning of each stage. Our linear mixed effect model consisted of the scenario, mobile robot strategy, the stage, and their combinations as fixed effects; and participant ID nested in both scenario and stage as a random effect. For each measure, we excluded some stages or scenarios where the values were manipulated (e.g., the mobile robot's distance changed as it joined the interaction). In our analysis, individual effects were not reported if interaction effects were found.

Distance to Mobile Robot

Our analysis focuses on time points when the mobile robot joined the group interaction or was just about to leave ($t_3 \rightarrow t_4$).We found significant interaction effects between the stage of the interaction and the strategy on participants' distance to the mobile robot ($F(6, 264) = 27.996$, $p < 0.0001$).

Figure 7.9.: Participants' aggregated distances to the stationary robot and mobile robot in the group interaction stage. The shaded region shows the range of values, the region inside the dotted line is the interquantile range, and the solid line is the mean. As the length of the interaction varied across participants, the shown aggregated value ends when half of the participants completed the stage.

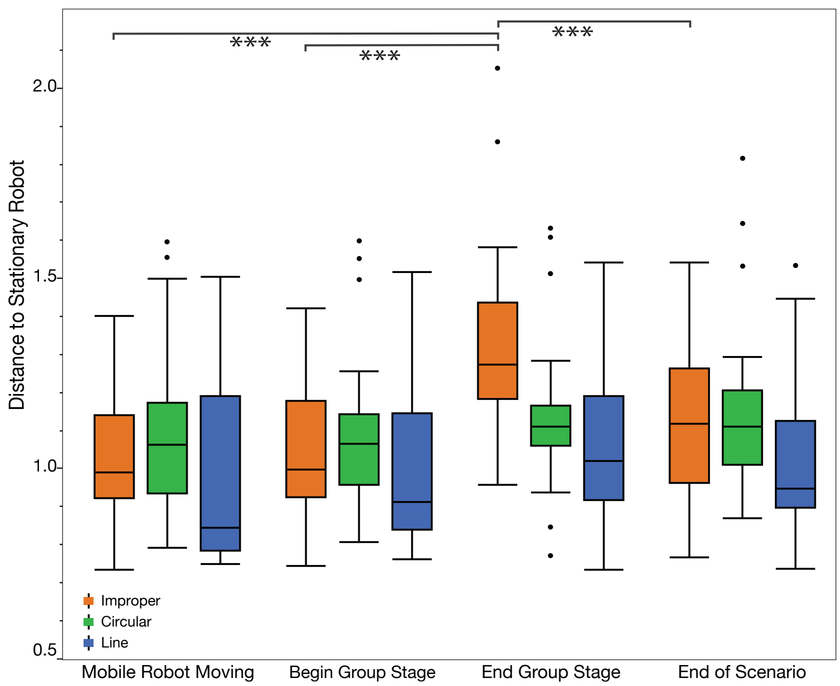


Figure 7.10.: Participant's distance to the stationary robot by stage for each robot joining strategy. *** means $p < 0.001$.

Post-hoc analyses showed that for Improper, the distance to the mobile robot at the beginning of the group interaction stage (M = 2.54m, SD = 0.35m) was significantly higher by an average of 0.7m compared to the end of that stage (M = 1.84m, SD = 0.57m, $p < 0.0001$). There was no evidence of a similar decrease in distance between stages in Line or Circular. The distances in Improper at both stages were significantly higher than distances in both Circular (begin: M = 1.01m, SD = 0.16m, $p < 0.0001$; end: M = 0.99m, SD = 0.13m, $p < 0.0001$) and Line (begin: M = 1.35m, SD = 0.32, $p < 0.0001$; end: M = 1.24m, SD = 0.27m, $p < 0.0001$). The differences we observed here matched our manipulation.

Distance to Stationary Robot

For the distance to the stationary robot, we included all stages from when the mobile robot was summoned to the end of the interaction ($t_2 \longrightarrow t_5$) and excluded the Guidance scenario as participants moved to follow the mobile robot.

We found an interaction effect between stage and strategy for participant distance to the stationary robot ($F(6, 396) = 3.408$, $p = 0.0027$). In the pairwise comparisons, we found that in Improper, participants' distance to the stationary robot at the end of the group interaction (M = 1.32m, SD = 0.23m) was significantly larger than either of the other two conditions at the same point ($p < 0.0016$ for both comparisons). Interestingly, the distance at that point was also significantly different from the distance at the start of the group interaction stage (M = 1.05m, SD = 0.17m) and at the end of the scenario (M = 1.12m, SD = 0.20m).

However, we did not find a difference between the start of the group interaction and end of scenario. This provides some evidence that after the mobile robot left the interaction, participants slowly reverted back to the original distance.

Difference in Distances to both Robots

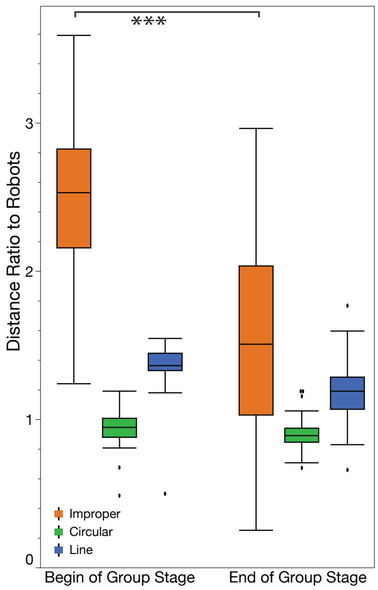


Figure 7.11.: The distance ratio by stage for each robot joining strategy. *** means $p < 0.001$.

One of the properties of an F-formation is that interactants jointly maintain a shared space in which they attempt to maintain constant distance between each parties, creating a balanced configuration [75]. As we hypothesized, participants moved to compensate for improper spatial formations. To quantify this, we calculated the distance ratio (participants' distance to the mobile robot over the distance to the stationary robot). The ratio encodes how imbalanced a configuration was. We compared the ratio at the start and the end of the group interaction stage ($t_3 \longrightarrow t_4$).

Similar to the distance results, we found significant interaction effects of the mobile robot joining strategy and stage of the interaction ($F(2, 264) = 46.357$, $p < 0.0001$). Pairwise comparisons showed that only Improper had a decrease in ratio from the beginning of the group interaction (M = 2.468, SD = 0.480) to the end of the group interaction (M = 1.506, SD = 0.658). For comparison, Line had an average ratio of 1.194 (SD = 0.190) and Circular had an average ratio of 0.894 (SD = 0.111) at the end of the group stage.

Distance to Group Centroid

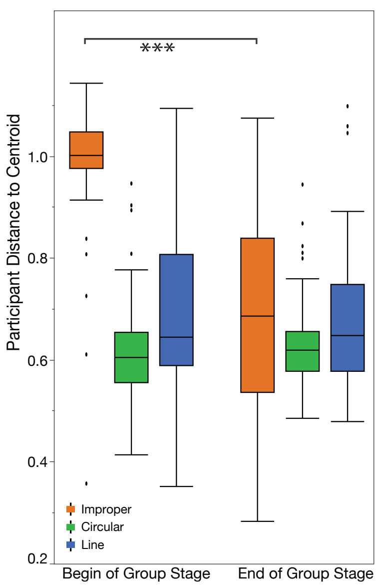


Figure 7.12.: The distance to the centroid by stage and mobile robot joining strategy. *** means $p < 0.001$.

Another property is the distance to the group centroid for all interactants. We calculated the participant's distance to the centroid and how it changed from the start to the end of the group interaction stage ($t_3 \longrightarrow t_4$). Similar to prior results, we found an interaction effect between the stages and strategy ($F(6, 264) = 31.197$, $p < 0.0001$). Pairwise comparisons showed the distance in Improper at the beginning of the group session (M = 0.98m, SD = 0.13m) to be significantly greater than other distances ($p < 0.0001$), including at the end (M = 0.68m SD = 0.20m). We found no significant differences in the distance to the centroid between strategies at the end of the group stage.

7.4.4. Qualitative Responses

At the end of each scenario, participants were asked to answer two qualitative questions about their experience: "*What did you like about the robots or the way they behaved?*" and "*What do you wish was different about the robots or the way they behaved?*". We reviewed the responses to identify patterns.

Difficulty of the Failure Scenario

Multiple participants talked about how the instructions in the Failure scenario were lengthy and hard to remember. Many participants suggested that the robot could guide them to their destination instead of giving directions, or wished that here were visual aids to help them understand the instructions. This was also reflected in participants rating this scenario to be significantly harder compared to other scenarios.

> *"I wish the directions were also shown visually on the robot or saved in some way, or sent to my phone. It would've been helpful in remembering the information to be referenced."* - P107 - Failure/Circular

> *"I also would prefer the purple robot to have taken me to the office instead of giving me directions."* - P123 - Failure/Line

As participants experienced the "Tour" scenario at the beginning, participants could have been primed with the idea that the mobile robot could have led them there.

Responses to Improper in Delivery and Pickup

As we were interested in participant's perception of the mobile robot's joining position, we labelled the instances where participants suggested the mobile robot should get closer. Two coders independently coded all 144 responses for whether the participant's response indicated a desire for the mobile robot to get closer to the them. The two coders had an agreement of 97.2 percent[8]. We then took the union of "Yes" labels by both coders.

8: We choose to report the agreement instead of Cohen's Kappa because the full set of labels was heavily biased towards "No".

We found 11 instances where participants expressed the mobile robot should get closer. Except for one instance in the Circular condition, all other responses were in the Improper condition. The single instance in the Circular condition reflected participant uncertainty.

> *"Was a little confused when package came that could just grab it, I think the robot could either come closer or there could be a signal on the robot that signals to take the package"* - P107 - Pickup/Circular

In the Improper condition, we found none of these suggestions in the Failure, one in the Guidance, two in the Delivery, and the majority (8 out of 12) in the Pickup scenario.

> *"The moving robot could have come closer to me or gave a quicker indication that I was supposed to grab the package because it was a bit further away and I couldn't tell whether it would move anymore or if I was supposed to move."* - P115 - Pickup/Improper

> *"I was confused if the mobile robot would actually come right beside me or not to deliver the package."* - P118 - Pickup/Improper

This gives some support to the idea that the action itself leads people to be more aware of the Improper configuration. The difference between Delivery and Pickup could have been that the participants were more unsure when they could pick up the object or if the robot had finished moving. However, we asked participants what they wished to change, so it is possible that other issues in Delivery might have overshadowed their responses. We looked at the responses in Delivery and were not able to find any common reason. Some participants mentioned the interaction felt long while others reported interaction failures, such as cross talk.

> *"Interaction would benefit from slightly faster response times."* - P126 - Delivery/Improper

> *"I wish the standing robot could stop talking when I was talking."* - P103 - Delivery/Improper

7.5. Exploratory Spatial Analysis

7.5.1. Spatial Formation and Social Cues (RQ1)

While our results demonstrated that there is a measurable change in participants' position at the beginning and end of the group interaction, we did not evaluate when exactly the shift happened during the group interaction. As we can visually observe in Figure 7.9, participants did not immediately shift their positions when the group interaction started; instead, they moved later. We suspected that their movements might correlate with social cues displayed by the stationary robot. When the mobile robot first joined the group, the stationary robot looked at the participant and announced the arrival of the mobile robot ("Here's [COLOR] Robot."). It then looked at the mobile robot and said, "Hi [COLOR] Robot". The mobile robot then turned to look at the stationary robot (except for the Failure scenario due to a bug) and verbally returned the greeting ("Hi Red Robot"). We extracted participants' distance to the robots and their ratio at the beginning and end of each social cue and compared their differences. We found that strategy ($F(2, 924) = 1134.195$, $p < 0.0001$), scenario ($F(3, 924) = 3.820$, $p = 0.0098$), and time points ($F(6, 924) = 4.964$, $p < 0.0001$) significantly affected the distance ratios. Pairwise comparisons on the effect of scenario showed the average ratio was lower in Pickup (M = 1.46, SD = 0.71) than in Failure (M = 1.57, SD = 0.68, $p = 0.0066$). No other significant differences were found.

Pairwise comparisons on the time points showed that the average ratio after the greeting by the mobile robot (M = 1.43, SD = 0.68) was significantly lower than the ratio at and before the point when the stationary robot announced the arrival of the mobile robot (M = 1.57, SD = 0.71). The p-values for the comparisons ranged from 0.0344 to 0.0054.

For the distance to the mobile robot, we again found significant effects of strategy ($F(2, 924) = 1550.76$, $p < 0.0001$), scenario ($F(3, 924) = 3.696$, $p = 0.0116$), and time points ($F(6, 924) = 5.571$, $p < 0.0001$). Pairwise comparisons on the effect of scenario found the average distance to the mobile robot was significantly lower in Pickup (M = 1.51m, SD = 0.71m) than in Failure (M = 1.60m, SD = 0.69m, $p = 0.0142$) and Guidance (M = 1.59m, SD = 0.65m, $p = 0.0499$). Similar to the ratio, pairwise comparisons showed the distance to the mobile robot was smaller at the time point when the mobile robot started greeting the stationary robot (M = 1.48m, SD = 0.69m) than the point when the stationary robot announced the arrival of the mobile robot (M = 1.61m, SD = 0.71m).

These results suggest that human movement might coincide with the robots starting to greet each other. However, this sequence of action between robots only lasted on average 9.02 seconds (SD = 0.73s), so it is possible that the movements were a function of time and not social behavior. The measured effect by scenario suggested that we may have lacked correct tests for H2 as effects were only detectable before the action.

7.5.2. Effects of Physical Action (RQ2)

Our second research question looked at exactly how physical action changed the spatial behavior of the participants. We compared the distance ratio at four time points in the group interaction: the beginning, the point when the robot started asking the person to take/place the item, the point after person completed the action and stopped moving, and the end.

We found a significant interaction between strategy and time point on the distance ratio ($F(6, 264) = 14.520$, $p < 0.0001$). Pairwise comparisons showed that in Improper, the ratio at the time after the person completed the action (M = 1.24, SD = 0.48) was significantly lower than all time points (M = 2.41; M = 2.08) before it (p-values were < 0.0001). There were no significant differences between that point and the end of the stage (M = 1.38, SD = 0.51m). The point at the end of the stage was also significantly different from the times before the user action ($p < 0.0001$). These changes were not seen in the other strategies. This demonstrated that physical action amplified the change in position and arrangement.

7.6. Discussion

7.6.1. Hypotheses Support

We found that when faced with an improper spatial arrangement, participants moved to balance the arrangement. We measured a significant decrease in distance ratio and distance to the mobile robot as the group interaction progressed. This provided **support for H1**.

H1 – Participants will move to rearrange themselves in the Improper condition more than in other joining strategy conditions (across all service scenarios).

We found **no support for H2**. We did not find any measurable differences in the distance ratios or distances to each robot between scenarios.

H2 – The effects of repositioning will be larger in scenarios where humans are required to move.

We found **no support for H3**. There was no measurable difference between the mobile robot joining strategies. All strategies were rated on average as being socially appropriate.

H3 – Participants will rate the interactions in the Circular and Line as more comfortable and easy than in Improper.

H4 – Within Improper, participants will rate scenarios with physical actions to be less comfortable and harder than other scenarios.

We found **no support for H4**. Besides the Failure scenario being deemed harder than other scenarios, participants did not rate scenarios with physical actions to be less comfortable or more difficult. This might be due to a ceiling effect; most participants rated the tasks moderately or extremely easy (123/144) and were moderately or extremely comfortable interacting with the stationary (109/144) and mobile robots (107/144). While we did not observe an effect on comfort and difficulty, majority of the participants in the Improper scenarios wished the mobile robot had moved closer in the Pickup scenario. This pattern was not observed in other scenarios.

H5 – Participants will rate the mobile robot as less competent in Improper.

We found **no support for H5**. We found no measurable difference in the mobile robot's perceived competence between mobile robot joining strategies.

H6 – Participants will rate the mobile robot's goal position in Improper as less socially appropriate.

We found **no support for H6**. While participants in the Improper condition reported that the mobile robot positioned itself "too far away" compared to the other conditions, the difference did not cause participants to rate the position as less socially appropriate.

7.6.2. Strategy Effectiveness

We believe that our results demonstrated the effectiveness of our two human-inspired group joining strategies (Line & Circular) and showed that both strategies had similar effects. Participants in the Circular and Line conditions felt little need to correct issues in the spatial configuration by getting closer to or farther from the robots. Also, both strategies were rated as socially appropriate ($M = 5.55$, $SD = 1.46$ for Line; $M = 5.08$, $SD = 1.36$ for Circular). While future work should explore if there are meaningful differences between them, we believe that both strategies can be used in practical HRI applications.

7.6.3. Movement & Tolerance

Our results showed that participants in the Improper condition moved more and attempted to compensate for the bad formation after the mobile robot joined the interaction (H1). However, we did not find any evidence that this made the task more difficult for the participants or that it was socially inappropriate (H3 & H6). This could be due to our measures not being sensitive enough to capture differences, or it could be that the corrective action was so trivial or subconscious that participants did not perceive it as an inconvenience (or at all).

7.6.4. Effects of the Environment, Embodiment, & Context

Prior work on F-formations [75] suggested that they can be affected by external factors, such as environment and cultural context. Putting additional large objects in our scene, placing the stationary robot elsewhere in the room, or conducting the experiment in a busy hallway could have led to different participant movements and responses. Regardless, we believe our algorithm would be robust to such changes through multiplier adjustments.

Our strategies provide a general framework for thinking about the spatial relationships among robots without considering differences in morphologies and sizes. Morphology of a robot changes where the endpoint for distance measurement should be, and potentially needs to compensate for large robot parts like the stationary robot's arm through additional offset. This made it difficult to determine endpoints for the distance measurement. For each robot, we defined a fictional center of presence that we believed users would face and focus their attention on when interacting with the robots. We used this point to calculate distance and added an additional offset, like in the Line strategy, to compensate for large robot parts like the stationary robot's arm. In practice, interaction designers need to consider robot appearances when specifying how they should approach human-robot groups.

Societal events impacted our study. Our experiment was conducted during the COVID-19 pandemic, during which people were advised to socially distance (2m or 6ft). It is unclear what influence this may have had on participants' spatial behaviors during our study and therefore on our results. In Improper, participants (now accustomed to social distancing) agreed with the statement that the robot was "too far away", but still felt that the position was "socially appropriate". There is a need for further research after the pandemic subsides.

7.6.5. Variability in Complex Interaction System

Our study used the complex interaction system described in Chapter 6 (Interactive System) that autonomously detected user position and reasoned about mobile robot goal positions. This was a research system and utilized four different computers to manage the two robots and multiple sensors, so there were unavoidable issues with latency as large amount of information was communicated through the system in real-time. The robot's gaze length and responsiveness also depends on where people are standing and what other events are happening. Some participants still experienced sessions with minor hiccups, such as a few seconds of delay in behavior and speech-to-text issues where the system completely misunderstood a response. We believe these trade-offs were worthwhile, as the system overall demonstrated that the group joining strategies and algorithm could be used in real-world settings.

7.7. Conclusion & Contribution

In this chapter, we developed two human group-inspired joining strategies and used them to investigate how a second mobile robot should join an existing human-robot interaction. These were compared to an improperly designed strategy in which the mobile robot positioned itself further away than "normal". We also investigated the effects of service scenarios on human spatial behaviors. We found that group configurations appear to be robust to different positioning strategies, and that people corrected the spatial arrangement imbalance in an “improper” condition and decreased the distance to both robots and centroid. We also found evidence that task involving people initiating physical actions may make improper spatial behavior more noticeable. Our exploratory analysis found that tasks involving physical action encouraged participants to get close to the robot once it joined, but those differences in distance did not persist into later stages of the interaction. We also found that changes in spatial behavior happened only after the first robot interacted with the second robot. Our results suggest that (1) people are willing to tolerate improper spatial arrangements, and (2) there is a need to better understand effects of social cues by robots on human spatial behavior in multi-robot group settings.

Person Transfers in the Field | 8.

8.1 Overview 95
8.2 Method 95
8.3 Findings & Discussion . . 99
8.4 Limitations & Future Work 103
8.5 Conclusion & Contributions 104

8.1. Overview

Research in this chapter was conducted in collaboration with Elizabeth Carter and Aaron Steinfeld.

As our last step, we conducted a short field study to better understand how findings from the controlled user studies might translate to the real world. Other prior work from our team has suggested that laboratory environments may lower perceptions of risk [29] and heighten awareness of certain details of robot behaviors [26]. Moreover, participants in one of our prior studies mentioned that their preferences for the amount of social interaction between the two robots might change based on their current sense of urgency [94]. Furthermore, we believe that real-world conditions, like having numerous people traversing the interaction environment and not having a scheduled appointment with the robots, can affect the way people interact with them and provide insights not available in a laboratory environment.

8.1.1. Chapter Summary

In this chapter, we describe the system and present the findings from our field study. First, we describe the changes made to the system from the study in Chapter 7 (Spatial Formation in Person Transfers). We then discuss the differences in the scenarios and environments. Lastly, we report the instances of person transfers that we observed and insights from the data that they generated.

8.2. Method

This study took advantage of an existing deployed robot on our university campus, Roboceptionist. This is a social robot system started in 2003 as a long-term robotic platform, and it has been involved in multiple prior studies in HRI [112, 134, 135]. Roboceptionist has undergone various changes throughout its deployment, most notably in its character and backstory. The latest version of Roboceptionist is an agent named "Tank".

Figure 8.1.: The scene of the study.

To capture the nuances missed in laboratory settings and observe unique one-off situations from the combination of various factors in the environment, we used a qualitative observational study approach. We recorded instances where people interacted with our robots and reviewed them for interesting factors of interactions and pitfalls in our system. After some interactions, we approached the participants and interviewed them about their experience.

As participants were likely to only interact with Tank once due to novelty, we chose to have a single scenario that every participant experienced. The new scenario involved the user(s) receiving stickers from the robots. We structured this scenario to further explore the differences between the *Circular* and *Improper* conditions in Chapter 7 (Spatial Formation in Person Transfers). The study was approved by our university's Institutional Review Board.

8.2.1. Study Environment & System

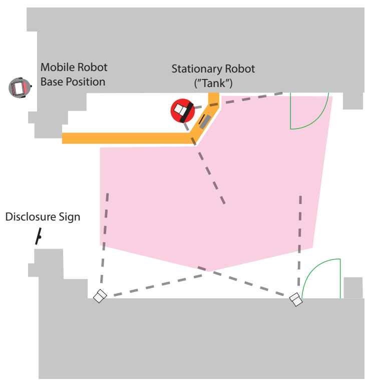


Figure 8.2.: An illustration of the study scene layout

This study took place in the 3rd floor entrance hallway of Newell-Simon Hall on Carnegie Mellon University's Pittsburgh campus. A bird's-eye view illustration of the layout is shown in Figure 8.2. The Roboceptionist system sits in a wooden booth in the entrance hallway. A partial wall in front of the booth creates a physical barrier between humans and the robot. A screen and keyboard for user input are placed on a wooden ledge directly in front of the robot. The robot is visible to visitors as soon as they enter the building from the main entrance. The Roboceptionist system consists of a stationary robot body (iRobot B21R) and a screen mounted on a pan-tilt unit that acts as the head.[1] Unlike the Baxter robot in the lab, Roboceptionist can only pan its head, which we used to convey gaze direction. The screen shows the animated face of "Tank", a muscular face that wears a headset that it uses to simulate taking phone calls. As part of our study, we made some modifications to the system design and behavior of Tank. We audited the existing codebase and fixed or replaced any failed components.[2] We also added a microphone and a speech-to-text software stack to the system, which allowed users to communicate with Tank through spoken speech. The speech-to-text was disabled when the keyboard was in use. Instead of using a laser scanner, Roboceptionist used the people-tracking component in Section 6.4 (Sensory Layer) to perceive the positions of people interacting with it. This extended Roboceptionist's capabilities to be aware of people in a larger area.

1: Gockley et al. [112] provides an introduction to the Roboceptionist project and the platform behind it. That paper describes the previous iteration of the Roboceptionist, *Valerie*.

2: The majority of the issues were missing dependencies or depreciated external services.

Similar to the previous experiment, we augmented the environment with three Azure Kinect cameras that provided information about people's poses and locations. The cameras were connected by wires to the three computers in the booth that controlled all aspects of the interaction. Our mobile robot hid around the corner out of view and only appeared when summoned.

As we could not account for all potential software issues through pilot testing, we addressed these issues as they arose during the study. These problems included invalid inputs and an overly aggressive interaction model (e.g., Tank greeted everyone even if they were not close by).

8.2.2. Robot Joining Strategy

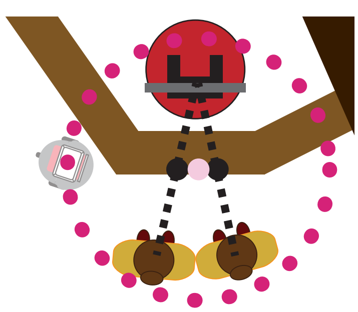


Figure 8.3.: Illustration of the group circular strategy.

One of the immediate differences from our previous study was that the existing *Circular* strategy is unsuitable for interactions involving more than 2 interactants because the perpendicular point can be occupied by another interactors. When there were multiple human interactors, we used a modified *Circular* strategy. As in the lab study, we first calculated the midpoint between an interactor and the main agent (here, Tank). We repeated the process for all interactors and obtained an average midpoint and distance. Afterwards, we obtained the desired distance from the the average midpoint by multiplying the distance with a multiplier. We then generated 24 points with the desired distance around the average midpoint with an interval of 15 degrees to approximate a circle around the average midpoint of the space in which the people and agent were interacting (Figure 8.3). These points were then evaluated by our "pick best" function that finds the best point for the mobile robot to join the interaction. If no valid points were found, the process was repeated with a higher distance multiplier. When evaluating a point, we considered three factors: (1) whether it was valid, (2) its distance to mobile robot, and (3) the smallest range in min-max distance to all members. The first two factors were the same as in Chapter 7 (Spatial Formation in Person Transfers).

The third factor was added to find the point that minimized the differences in the distances from the position to each group member. For example (as shown in Figure 8.3), when two people are facing Tank, the mobile robot will pick the leftmost point among the 24 candidate points as it is both valid (not colliding with anything) and has the smallest range in its distance to each member (i.e., the distances to Tank and the two people are similar).

The scoring function, f can also be expressed as the following equation:

$$f(p) = \begin{cases} -1 & \text{if } \neg\,\text{valid(p)} \\ W_d \exp(-\|p - m\|) + W_a \exp(\min_{e \in E} \|p - e\| - \max_{e \in E} \|p - e\|) & \text{otherwise} \end{cases} \tag{8.1}$$

where E is the set of all entities (robot and human) in the current group, and $\|p - m\|$ calculates Euclidean distance between the tested point p and the mobile robot m. The weights were fine-tuned for our scenario. In our study, we used $W_a = 1$ and $W_d = 8$.

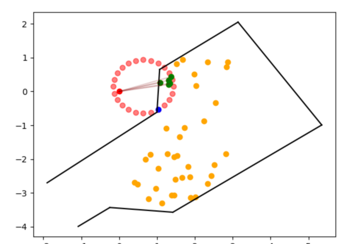


Figure 8.4.: A screenshot from simulated human data. In the figure, three people (green) stood in front of Tank and were recognized to be in a group. The system first calculated the average midpoint (black) and generated a circle of candidate points around it (red). The system then picked the best point (blue) based on the criteria. The area inside the black outline is the valid region which accounted for the mobile robot footprint.

To use this new strategy in our existing algorithm 1 from Chapter 7 (Spatial Formation in Person Transfers), we updated the algorithm (1) to choose the two variations based on the number of interactors and (2) to use the new scoring function.

8.2.3. Scenario

Instead of using the scenarios from the prior study, we created a new scenario similar to the "Pickup" scenario. When the study was active, Tank told the participants interacting with it that it was giving out stickers as part of its reopening. If the participants indicated they wanted the stickers, Tank would inform them that it had run out of stickers and would summon another robot ("green mobile robot") who had more stickers. The mobile robot would then drive around the corner and join the interaction. After exchanging greetings with Tank, the mobile robot prompted the participants to take a sticker from a top-mounted tray. The experimenter, who was standing nearby and tried to be out of the way, could command the robots to skip the prompt if participants had already taken the stickers. Afterwards, the mobile robot told Tank that someone would be coming with more stickers. We added this brief conversation to observe how people would react and observe people's movements while a group interaction with the two robots was in progress. After the conversation, the mobile robot informed the participant that it had to leave, and it departed. Tank then looked at the participant and told them about the study, mentioning that they could approach the experimenter if they had any questions.

8.2.4. Participants & Recruitment

Our study included four types of participants:

Passerby – These were people who passed by our robots without interacting with the robots or observing any human-robot interactions.

Observers – These were people who passed by and observed the robots interacting, e.g., by slowing down or stopping to watch someone else interact with it. However, they did not directly interact with the robots.

Participants Group A – These were people who took part in some or all of the multi-robot interaction[3].

3: Some participants left the interaction halfway.

Participants Group AA – These were people who were in group A and also answered a few questions that the experimenter asked. This interaction took less than 5 minutes. Participants in this condition were not compensated.

Participants Group AB – These were people who were in group A but also participated in a 15 minute interview and completed a questionnaire after the interaction. Participants were compensated $10 USD for their time.

After the participants completed the interaction, the experimenter approached the participants and asked if they have any questions and inform them about the interview. We were not able to intercept all participants as some left the building while the mobile robot was driving back[4]. Furthermore, not all participants agreed to be interviewed due to other commitments. We also did not interview participants who knew or recognized the experimenter.

4: We had to supervise the mobile robot going back and only approached the participants after the robot stopped.

There were two groups of approached participants (AA and AB) as we wanted to give them the option to answer a few questions instead of a 15 minute interview. Group AA also acted as a pilot group as we tested the questions to ask participants.

5: This was the duration with recorded interaction data (e.g., video data, robot state, etc).

We recorded 7 hours and 2 minutes of study data[5] over 4 work days. The recordings all took place between 11:22am and 6:30pm. We only enabled the microphone input system for some sessions. We observed during our pilot testing that vast majority of people who walked through the hallway ignored Tank. Therefore, we added an additional sign in front of Roboceptionist to advertise that it was giving away free stickers. In our post-study analysis, we observed 18 person transfer interactions involving at least 40 people[6]. We interviewed 11 people (3 in group AA and 8 in group AB).

6: We obtained this number by counting the number of people in front of Tank when the mobile robot was summoned. We believe this is a lower bound as there were multiple instances where people approached or interacted with the robots after the mobile robot showed up.

8.2.5. Technical Issues & Implications

Unfortunately, we were not aware that our system was not functioning as planned. When there was a high volume of people, our internal interaction model ran at a slower cadence than our input stream. This led to a backlog of data and the model lagging behind real-world data as time passed[7]. The system was run for hours at a time and led to our model being significantly behind. This caused the mobile robot joining strategy to use outdated data and led to 11 out of 18 of our sessions to use the backup position. Furthermore, as we did not save the necessary data, we were unable to verify the extent of the delay on the remaining 7 sessions.

7: We failed to enforce a maximum queue size for the ROS topic, which led the queue to grow very large.

While this was detrimental to our understanding of the effects of the joining strategies, the primary goal for this study was to understand how people responded to our robot in the field, and that was not affected by this error. In all cases, participants experienced a person transfer. Their behavioral responses and interview data can provide some insights into how these systems might work outside of a laboratory setting.

8.3. Findings & Discussion

8.3.1. Overall Impressions

Overall, participants told us that they thought the interactions were “cool”, “fun”, and “really neat”. For most participants, the highlight of the interaction was the arrival of the mobile robot. This was also reflected in our observations where participants expressed excitement when they saw the mobile robot approaching them.

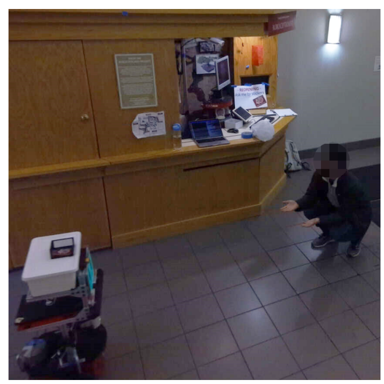


Figure 8.5.: A participant reacting to the arrival of the mobile robot.

8.3.2. Trust in Person Transfers

Figure 8.6.: A comic strip showing participants leaving after Tank informed them it had run out of stickers and coming back after the mobile robot showed up. (1) A group of participants interacting with Tank. (2) The group walked away after Tank told them it ran out of stickers. (3) One of the participants (in blue) saw the mobile robot, pointed at it, and called for the group. (4 & 5) The group came back and surrounded the mobile robot to get the stickers.

One of the common themes bought up by the interviewed participants was that they were surprised that the mobile robot showed up with more stickers. Participants told us that they thought the sticker was a lie and that Tank was joking in saying that another robot would come to give out stickers. Their rationales were that the joke fitted Tank's personality and that they had never seen another robot in this area before. Participants stated they did not believe it was true until they saw the mobile robot turning the corner. This reinforces the current novelty of interacting with multiple robots and the need for more research on this topic.

Participants' behaviors confirmed this disbelief in the promise of another robot: in several interactions, after Tank informed a participant that it had run out of stickers, the participant left the interaction. As the mobile robot moved towards Tank, the participants turned around and reengaged with the robots (Figure 8.6). There were also participants who intercepted the mobile robot and took the stickers from it as they left.

8.3.3. Mobile Robot Handover

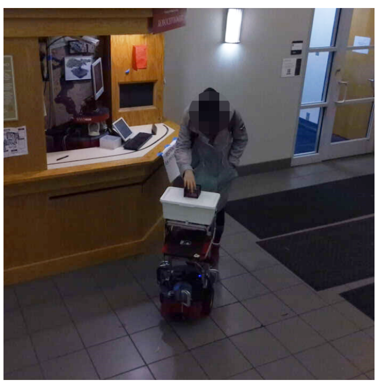


Figure 8.7.: A participant taking the stickers from the mobile robot as the mobile robot finished its movement.

The action of picking up the sticker from our mobile robot can be viewed as a "handover" of an item from our mobile robot to participants. One of the interesting findings in Chapter 7 (Spatial Formation in Person Transfers) was that the majority of the participants in the Improper condition talked about how far away the mobile robot was in the Pickup condition. We believed this might have been caused by participants being unsure about when they should approach the robot and take the package.

In 13 out of 18 person transfer instances, participants took the stickers while the mobile robot was still trying to reach its final position (10 "Backup" position, and 3 "Circular"). In the remaining 5 scenarios, participants only took the stickers after being prompted by the mobile robot. Besides one unique case where the participants were recording the interaction, the mobile robot was further away from the participants in the remaining 4 sessions (1 "Circular" and 3 "Improper").

These results provided insights into how people perceive receiving items from a robot. The act of receiving an item requires the person to decide when to approach and take the item. In most cases, as the robot got close to the participants, participants took that as permission to take the items. In cases where it was further away, participants waited for the signal from the robot that it had completed its movement before taking the item.

8.3.4. Change in Group Membership

Among the 18 sessions, we observed a few instances where someone joined the interaction halfway. In one case, one participant (P4) joined two others who were already interacting with Tank. The first two participants requested the stickers and the mobile robot approached the group, moving to its backup position. As the mobile robot got closer, P4 stepped back and moved away from the interaction. From the recording, we observed that

Figure 8.8.: A comic strip showing a participant (P4, blue backpack) joining halfway. From left to right: (1) P4 joined two other participants who were interacting with the robot. (2) The mobile robot was summoned and moved towards them. (3) P4 stepped back as the robot approached them. (4) P4 observed the interaction between the mobile robot and the other two participants. (5) P4 stepped further away as the interaction progressed.

P4 stood further away and even stepped back as the two robots interacted with each other.

In the post-interaction interview, P4 mentioned that it was unclear to them if they were part of the interaction because it was their friends who were initially interacting with the robot. As the mobile robot moved towards them, P4 was unsure if the mobile robot knew they were part of the group and moved out of the way. This sequence of interaction shows that the joining behavior has the potential to influence people's perception of group membership. A better and socially appropriate position could have made P4 feel confident they were part of the group and have been less likely to prompt them to move away.

In one of the sessions, we observed a participant who stood to the side observing the interaction between the robot and another group of participants. Once the participant observed the group taking the stickers, they stepped in, took a sticker, and left. It was unclear if the person knew anyone who was initially interacting with the robots.

8.3.5. Mobile Robot Joining Position

While the mobile robot moved to the backup position in the majority of the interactions, we still collected valuable feedback on the position choice. Participants generally found the chosen position to be appropriate. One of the participants who experienced the backup position stated that they wished the robot was closer. They talked about while the position of the robot was where they expect a person to be, they believe the robot needed to come closer because it lacked the manipulation capabilities to hand over the object like a human would. Because the participant had to lean forward and take the object, they talked about how the robot should be only "one hand" (arm's length of) distance away compared to the "two hands" of distance that it was[8] distance that they experienced.

8: "Two hand" as in one arm's length by the robot and one arm's length by the participants, similar to how a handover is done by humans.

These findings, together with our observation of handovers and changes in groups, demonstrated the importance of a task-aware and human-aware joining strategy. A static position may lead to the system accidentally excluding others in the group, lead to people misinterpreting the robot's action, and could likely be a poor position for certain tasks.

8.3.6. Effects of Keyboard Inputs & Failures of Speech-to-Text

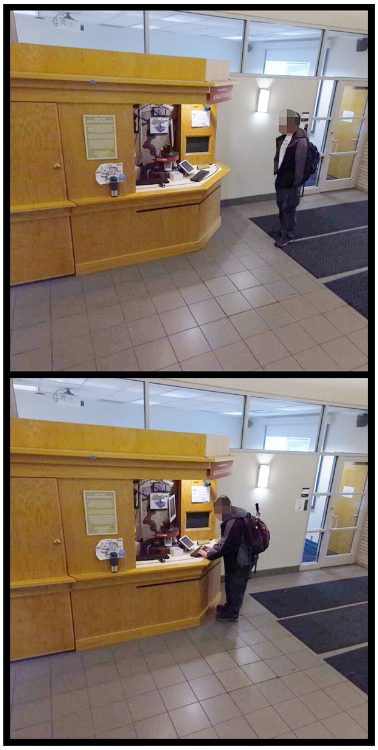


Figure 8.9.: An example showing a participant first attempting to use speech to interact with Tank (top) and eventually using the keyboard (bottom).

When we first designed this study, one concern we had was that the keyboard input would limit the movement of participants because it would require them to move back to the starting position (where the keyboard was) to provide input. We ran this study with both keyboard input only and the combination of keyboard input and microphone.

Due to the combination of ambient sounds, hallway acoustics, and COVID-19 masking guidelines, our speech-to-text system was unreliable. We observed multiple participants who first attempted to use the speech system before stepping forward and interacting with the robot through the keyboard. In the interviews, participants also expressed similar sentiment and talked about how the microphone was unreliable and they ended up using the keyboard.

We also found some evidence supporting our concern that participants would move to the starting position. We observed participants moving back to the keyboard to type responses such as "Thanks for the stickers", "got the sticker", ''thanks". This supports our intuition that the keyboard anchors the human position during the interaction and changes the spatial dynamics in the interaction. Future work should explore how these anchors[9] influence user positions and movements during person transfers and multi-robot interaction.

9: In addition to keyboards, task requirements such as a fixed tool could also change how people move during these group interactions.

8.3.7. Robot-Robot Communication

After the participant picked up the sticker, the robots had a quick conversation. Participants had mixed reactions to the exchange. In a few cases, participants left after taking the stickers and did not wait for the interaction to finish. This was understandable as the conversation did not add any value to the service. The majority of the participants did wait for the robots to finish their conversation. When asked about the conversation, one of the participants mentioned that it was a good addition as it showed they could communicate and were on the same team. Some participants also stated that it was artificial and a performative act. One participant also mentioned they were surprised that the mobile robot could talk at all.

8.3.8. Other People in the Scene

We were also interested in how bystanders and others in our study area interact with our robots. In the majority of the sessions, when the mobile robot was close to Tank, we observed people moving around the robots and the participants interacting with our robot. When there was a big gap between the participant and the mobile robot, we observed that most people in the passersby category simply walked through the gap, violating the human-robot group space (Figure 8.10). However, we did observe one instance where a person consciously walked around the mobile robot even when it was far away from the participant it was interacting with.

Figure 8.10.: A comic strip of a group of people walking through the group interactions. From left to right: The participant (red shirt) observed the greeting between robots. A group of people simply walked between the robot and the participants. After the group passed, the participant approached the mobile robot to get the sticker.

Figure 8.11.: The top down annotated point cloud view of the scene. We annotated the position of the participants (circle), the walls (white edges), the mobile robot (square), and the bystanders moving through (triangle).

It was unclear how much the layout of the hallway affected whether bystanders decided to walk through the gap. In the scene shown in Figure 8.10, the mobile robot moved to the middle of the hallway and slightly blocked the default route through the hallway (as shown in Figure 8.11). While there was sufficient space behind the mobile robot for people to move through, it required large trajectory changes, and the space behind the robot was unlikely to fit a group of people. The function of the hallway as a means of moving between spaces may have also led bystanders to believe it was socially acceptable to violate group O-spaces. Having said this, it was also unclear if the bystanders perceived the robot as being in a group with the two participants.

We also encountered issues where other people in the scene purposely blocked the movement of the robot[10].

10: A lot of the people walking through the hallway were students of our university and some tried to test the limits of the robot.

8.4. Limitations & Future Work

Our main limitation was the failure to validate the effectiveness of the new group circular joining strategy. While our simulated human data show that the position seemed reasonable, we were unable to verify that our new group joining strategy was socially appropriate in the real world. We hope we can address this in the future.

As pointed out above, group memberships and participants in the interactions were not static and constantly changed as people joined and left the interaction. We observed cases where up to 7 people concurrently interacted with our robots and cases where people left and rejoined the group throughout the interaction. We used a simple model to determine the group membership and only used it once when the mobile robot calculated its goal position. As the mobile robot approached a group, the group membership could change and the chosen position could be outdated. A more dynamic model may be more useful in the future.

8.5. Conclusion & Contributions

Our field study expanded our understanding of how person transfer works and is perceived in the field. Our findings demonstrated why a context-aware socially appropriate mobile robot joining strategy is needed. A fixed position strategy, similar to our backup position, will likely be less preferable for some tasks, alienate certain group members, and not react to the changes in group membership. Similarly, a bad, improper position can also lead to interruption by others as they walk through the interaction O-space. We also found it is important for the first robot to communicate clearly how and when a transfer is going to happen. We believe that the dialog by Tank ("I ran out of stickers. Let me call green robot with more stickers,") likely did not instill confidence as it did not convey when and how the second robot would arrive.

Part 4: Final Words

Future Work & Real World Considerations | 9.

The opportunities enabled by multi-robot interaction are rich and generate research questions that could span multiple PhDs. This dissertation focuses on person transfers, a very specific phase of multi-robot human interaction. In this section, we first briefly summarize and revisit the unanswered questions related to person transfer that this dissertation has exposed. We also discuss other areas of possible future exploration related to multi-robot interactions. We end by discussing research and ethical implications of our findings.

9.1. Unanswered Questions from this Dissertation

9.1.1. Composition of Robot Teams

One of the main limitations of this work is the focus on person transfer between two robots, in which humans only interact with at most two robots. An interesting unanswered question is how the results would scale and evolve when there are more than two robots. The robots' capabilities could also influence how people perceive the necessity of person transfers. We highlighted multiple reasons a service might utilize person transfers, but these choices might not be obvious to the users[1]. Therefore, future work should investigate how such interactions could be expanded beyond two robots.

1: For example, one of the study participants in Chapter 5 asked why the first robot did not have wheels so it could lead them to their destination. Likewise, a user might not be aware of a service requirement for a robot to stay on station or that the first robot had a drive failure.

Additionally, further exploration is needed to examine how the morphologies of the robots could impact people's perceptions of multi-robot interactions and other variables that our work has measured. In our studies, multiple participants commented that the mobile robot's eyes were cute and that Tank's face was robotic and creepy[2]. Different morphologies not only affect the appearances but also enable different functionalities: at least one participant expressed the opinion that the mobile robot should move closer to make up for the fact that it lacked the manipulation capability to hand over a sticker as a person would. It is worth considering how morphology can change perceptions of person transfer and how it can directly affect transfer. In Chapter 4 (Design Space for Multiple Robots And Person Transfer), robot form was one of the factors of our taxonomy. For scope reasons, none of our follow-up work explored this avenue. While morphological differences between robots in different scenarios are unavoidable[3], we can also explore ways of quantifying the differences and explore how they correlate with human acceptance and behavior[4].

2: It was designed that way by its creators to fit Tank's personality and storyline.

3: Even if we are able to standardize robot platforms across the globe, it is likely that different generations of robots will have different capabilities and subtle design changes that could influence human behavior.

4: An example of this type of work is Phillips et al. [136] where the authors quantified how the features of a robot face influenced how human-like a robot is.

9.1.2. Domain of Transfer

In this dissertation, we investigated our research topic through the lens of service interaction where users received services from robots (e.g., leading to another location, experiencing a system failure when presenting information). These tasks were based on current human-human person transfers. Deployed robotic systems that complete some of these tasks with a single robot also exist. However, we only sampled a portion of potential interactions and a different domain could influence human acceptance and design insights. For instance, the way robots should interact with each other in Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer) might change in high-stress scenarios where users prioritize completion over interaction. Users might also care less about where the mobile robot is and might even prefer joining positions that minimize task time over social normative positions. We also did not explore collaborative settings where the same robot repetitively leaves and rejoin an interaction.

9.1.3. Scheduling Person Transfers

While there are many cases where transfers between service robots are unavoidable (e.g., when a robot cannot physically complete the task, when a robot's path is blocked by physical obstacles, when a robot fails, etc.), there are also cases where transferring between service robots is preferred, although not required. In these scenarios, there are sometimes other constraints, such as an owner's desire for a robot to remain in a location or concerns about resource allocation. Prior works such as [16, 92, 137] have explored the reassignment cost—the cost of having the robots guide the users instead of completing their background tasks. However, whether a transfer is suitable or even allowed might also depend on the person's preference. For example, a person with low trust in automation might prefer single-robot interactions. Future work should further explore the role of user preferences in scheduling person transfers.

9.2. Other Interesting Areas in Multi-Robot Interaction

As described in Subsection 4.2.4 (Findings), multi-robot interaction exists in various other contexts. Person transfers could be the first phase or part of various types of multi-robot human interaction.

1. **Multi-Robot Collaborations** – While transfer can be a part of multi-robot collaborations, there are many cases where the collaboration between robots could be continuous. For instance, how should we model the continuous information transfer between robots when they are actively collaborating in a group with both humans and robots?

2. **Passive Multi-Robot Interactions** – In a world full of robots, it is highly likely there will be other robots living around us but not actively engaged in our service. For example, while we receive guidance from a mobile robot, there could be a Roomba nearby that is vacuuming the floor. Should there be interactions between the two robots? Should they have some coordination of their actions? If not, how should the other robot know the first robot is incorporating that knowledge into planning?
3. **Longitudinal Interactions** – Interacting with a robot is likely not a one time affair. Our examples of person transfer were often immediate or with some delay. However, we can also see repeated interaction between a person and a robot as transfers where the person leaves and reengages with the robot everyday. How might our findings change over repeated exposure?

9.2.1. Person Transfers in a Multi Intelligent System Future

Another interesting aspect of this problem that is worth investigating further is transferring users to other intelligent systems or smart environments. An example of such transfer is prior work by [47] in which a physical social robot redirects user attention to a virtual agent or my other work on a social robot managing user attention between itself and a smart screen [25]. How would the agency and embodiment of the other device change the findings in our work? Could the same transfer strategy work with a screen interface?

9.3. Ethical Considerations

Beyond the limitations and open questions in our study, we also reflect on the ethical considerations for our results and contributions in this dissertation.

All of our studies were conducted with approval by our university's Institutional Review Board (IRB). Among our studies, the field study described in Chapter 8 (Person Transfers in the Field) carried the most ethical risk. In this study, passersby were only informed about the recording through disclosure signs and participants were only informed they were in a study after they interacted with the robots[5].

5: All our other studies had recruited participants with informed consent at the start of the study. We also described in detail the optional permissions when presenting the consent form.

We placed multiple RGB-D cameras in the environment to capture the reaction and movement of both people interacting with Tank and other individuals who were moving through the hallway. While the hallway was deemed a public space with no expectation of privacy, we still placed signs around the area and gave people ways to anonymously contact us with requests to have their data deleted. We also stated in our IRB that any video data released would have human faces pixelated, unless agreed by the participants afterwards on a consent form. We believe it is our responsibility as researchers to weigh the possible benefits of sharing the data against the cost of the participants' privacy. While complete natural

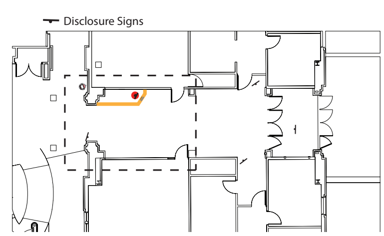


Figure 9.1.: Location of the disclosure signs in the study area.

data of people interacting with multiple robots would be an interesting dataset for future research, we believe large gains still can be made from a cleaned, anonymized interaction dataset.

The system described in Chapter 6 (Interactive System) relied on complete knowledge of people's position in the world. We implemented this approach by placing multiple RGB-D cameras in the environment to record people's movements and positions. We choose this approach to overcome the issue of occlusion, which is common in group settings. While this made it easier for the study, systems like this may not be the best way forward for robotic technology as it encourages a future with complete surveillance of the environment. We believe it is possible to gather the same information with some trade-offs from a single camera system.

Our participants were also from a population that does not reflect the general population. Our participants were younger, educated, and were conveniently sampled from those close to our university. The participants were also likely to be more familiar with technology or robots either through their education or past experiment experiences[6]. This bias could have an impact on our findings as it might not be generalizable, especially with older or novice users. Our small sample size likely further exaggerates this concern.

6: Some participants have disclosed to us that they had taken part in multiple human-robot experiments.

We believe the potential for misuse of our system and findings is minimal. It is possible that the social communication and the appearance of intelligence by our system could convey to people that the system is more intelligent and create over-trust [138]. While we believe it is important for researchers to understand the boundaries of future novel interactions through studies like this, we need create better ways to convey such limitations to novice users in the future. Furthermore, as we found support that explicit transfer of information was preferred by our participants, we need to consider how to determine the type information that should be transferred and whether it should be done discretely or explicitly. Improper disclosure or sharing of information can lead to a breach of privacy. Systems will also need to be transparent about what is being transferred and have ways to demonstrate that to unfamiliar users.

One of the primary concerns about robotic technologies is whether they are a net positive for our society or exaggerate inequality [139]. In our work, we see the role of the robots as being supplementary to those of humans, enabling humans to focus on tasks that are difficult for robots or rewarding for humans, and extending service reach during times when humans prefer to not work (e.g., in the middle of the night). For example, a transit station agent might deploy a mobile robot to distant entrances or supervise a kiosk robot at a remote location. However, this work has advanced our knowledge of how certain types of tasks could be automated, leading to fewer future job opportunities. We do not pretend to have an answer to this complex social-economical question about automation, but as a society, we need to better prepare for the addition of automation and robots in our workforce.

Conclusion 10.

10.1. Research Contribution

Our work contributed to to the field's understanding of an important phase in multi-robot interaction: the person transfer. We believe some of our most important contributions are demonstrating the nuances in this new research area, proposing new research questions, and inspiring more conversation in our research community. While simple, person transfers should be examined across more varied research lenses and contexts.

In our exploration, we investigated certain aspects of person transfers. We now reiterate the contributions of each chapter, categorized by the type of contribution to the field in the following sections.

10.1.1. Design Findings

The following are the contributions of this dissertation that may be useful to HRI designers:

Ch. 4 When designing person transfers, designers can use our taxonomy as a framework to guide their choices.

Ch. 4 When discussing person transfers with other stakeholders, our taxonomy provides a vocabulary and template to communicate and understand the problem space and the required criteria for the service.

Ch. 5 When designing interactions between multiple robots, it might be preferable to design verbal interaction to follow human social norms, even when robots lack human features.

Ch. 5 When designing the presentation of information exchange between robots in front of humans, it is inadvisable to have the information transfer be hidden. However, we have no evidence regarding what type of transfer (acknowledging vs. reciting) is preferable.

Ch. 7 When designing the arrival of the second robot, the designer can employ one of the two different robot joining strategies that our studies found to be perceived as socially appropriate.

Ch. 8 When designing the person transfer experience, designers can review our observations for inspiration and identify potential pitfalls.

10.1.2. Behavioral Findings

This dissertation contributes the scientific knowledge on people's behaviors when interacting with multiple robots during person transfers:

Ch. 4 Presents different factors that are potentially important for users in person transfers.

Ch. 5 Participants reacted more positively to the speaking robot when the verbal communication between robots followed social norms, as compared with when the verbal communication did not follow social norms or when there was no verbal communication during a transfer.

Ch. 5 Participants reacted more positively towards a minimally social robot when there was social interaction between it and a social robot, as compared to when the social robot spoke on behalf of the minimally social robot during a transfer.

Ch. 5 Participants found a robot to be less mean and more likable when it recited information aloud to another robot, as compared to when it silently transferred the information to the other robot while transferring a user between robots. This also validates prior work [12].

Ch. 7 When faced with "Improper" configurations, where one robot was further away than another, participants moved to rebalance the group's spatial configuration.

Ch. 7 Validates the existence of F-formations in multi-robot interaction by showing people corrected formations.

Ch. 7 Participants noticed bad spatial formations more during tasks with physical actions that required participants to take initiative.

Ch. 8 Insights on how people interact with multiple robots in a real-world person transfer scenario.

10.1.3. Technical Findings

This dissertation contributes the following technical advances:

Ch. 5 An autonomous proof-of-concept human transfer system demonstrating coordination and verbal exchange between robots. This demonstrated the feasibility of implementing such a system for future deployments.

Ch. 6 An open-source system[1] that allows other teams to advance research in multi-robot interaction or deploy commercial robot applications.

1: https://github.com/CMU-TBD/tan-dissertation-2022

Ch. 6 Multiple components that connected various frameworks used in our system (\Psi , IPC, and ROS). We also created various components to assist in the visualization of ROS data in PsiStudio.

Ch. 6 Open-sourcing of our components will enable other researchers to use all or parts of our components. It will also provide examples to interested parties of how we designed and implemented our interactions.

Ch. 7 & Ch. 8 Methods for a mobile robot to select socially appropriate goal positions when joining a group.

10.2. Closing Remarks

This thesis only explored a small part of the space of multi-robot human interactions, the person transfer. Even in this research area, there are still multiple unanswered questions (Chapter 9). One of the goals of this thesis is to inspire more research in this area, especially on how dyadic human-robot interactions are enhanced by other robots and intelligent systems. By using a wider lens and looking not only at individual robots in isolation but also at robots joining and leaving existing human-robot interactions, we can find new research questions that were present but concealed by the veil of "complexity". Our research has provided evidence that in multi-robot person transfer, the whole is greater than the sum of the parts. The addition of the second robot not only revealed questions pertaining to the inclusion of the second robot (e.g., joining strategy, rationale for inclusion) but also surfaced new questions (e.g., robot-to-robot communication, trust in person transfers) from the synergy of both robots. We hope this thesis provides a convincing argument to explore this vast new space and a good foundation for examining the crucial phase where a dyadic interaction transforms into a group interaction, and for discovering the richness of this space.

Appendices

Questionnaires for Chapter 5 | A.

The following pages are the questionnaires used in Chapter 5 (Inter-Robot Communication & Information Transmission In Person Transfer). The pages are arranged with the three questionnaires (Q1, Q2, and Q3) followed by the demographic questionnaire.

Participant ID (Number): Robot Name: Guide Robot

Part 1(A). Please answer the questions below on a scale from 1 to 7:

- I trust the guide robot will be able to guide me to my destination.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- The guide robot is dependable.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- The guide robot is reliable.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- I can count on the guide robot to guide me to my destination.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- I am wary of the guide robot.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- I am confident in the guide robot's ability to complete the task.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely

Part 1(B). Please answer the questions below on a scale from 1 to 7:

- The receptionist robot is a reliable partner.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- I feel kind of protective towards the guide robot.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The guide robot's problems do not disturb me a great deal.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots like each other.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots know each other well.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots compliment each other well.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The guide robot is a reliable partner.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots ignore each other.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- I was bothered by the receptionist robot's behavior.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree

GUIDE ROBOT EVALUATION:

Part 2. How closely are the words below associated with the **guide robot**?

Happy
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Feeling
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Social
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Organic
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Compassionate
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Emotional
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Capable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Responsive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Interactive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Reliable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Competent
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Knowledgeable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Likable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Mean
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Friendly
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

RECEPTIONIST ROBOT EVALUATION:

Part 3. How closely are the words below associated with the **receptionist robot**?

Happy
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Feeling
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Social
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Organic
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Compassionate
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Emotional
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Capable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Responsive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Interactive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Reliable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Competent
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Knowledgeable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Likable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Mean
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Friendly
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Part 4. Please answer the questions below in as much detail as you can provide.

1) How would you describe the relationship between the receptionist robot and the guide robot?

Participant ID (Number): Robot Name:

Part 1(A). Please answer the questions below on a scale from 1 to 7:

- I trust the guide robot will be able to guide me to my destination.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- The guide robot is dependable.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- The guide robot is reliable.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- I can count on the guide robot to guide me to my destination.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- I am wary of the guide robot.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely
- I am confident in the guide robot's ability to complete the task.
  Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Extremely

Part 1(B). Please answer the questions below on a scale from 1 to 7:

- The receptionist robot is a reliable partner.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- I feel kind of protective towards the guide robot.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The guide robot's problems do not disturb me a great deal.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots like each other.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots know each other well.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots complement each other well.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The guide robot is a reliable partner.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- The two robots ignore each other.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- I was bothered by the receptionist robot's behavior.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree

GUIDE ROBOT EVALUATION:

Part 2. How closely are the words below associated with the **guide robot**?

Happy
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Feeling
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Social
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Organic
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Compassionate
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Emotional
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Capable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Responsive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Interactive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Reliable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Competent
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Knowledgeable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Likable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Mean
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Friendly
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

RECEPTIONIST ROBOT EVALUATION:

Part 3. How closely are the words below associated with the **receptionist robot**?

Happy
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Feeling
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Social
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Organic
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Compassionate
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Emotional
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Capable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Responsive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Interactive
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Reliable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Competent
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Knowledgeable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Likable
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Mean
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Friendly
Definitely not associated 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 definitely associated

Part 4. Please answer the questions below in as much detail as you can provide.

1) How would you describe the relationship between the receptionist robot and ________________?

____

____

2) What did you like and/or dislike about the interaction with the robots?

____

____

____

Participant ID (Number):

Part 1. Check the box next to your choice.

Did the **receptionist robot** get your name right?

☐ Yes ☐ No

Do you think the **guide robots** (Blue Robot, Green Robot, Yellow Robot) knew the following information after the interactions were completed?

That you wanted help navigating to a location	☐ Yes	☐ No
Your destination	☐ Yes	☐ No
Your name	☐ Yes	☐ No
Your preferred walking speed	☐ Yes	☐ No

Please feel free to explain any of your answers:

____

____

____

Part 2. Please answer the questions below on a scale from 1 to 7:

- Yellow Robot is competent.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- Green Robot is competent.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree
- Blue Robot is competent.
  Strongly Disagree 1 | 2 | 3 | 4 | 5 | 6 | 7 Strongly Agree

Part 3. Please circle your choice for each of the questions below:

If you needed a robot to lead you somewhere, which one would you choose?

Yellow Robot | Blue Robot | Green Robot

I feel connected the most to ....

Yellow Robot | Blue Robot | Green Robot

Which robot do you prefer the least?

Yellow Robot | Blue Robot | Green Robot

Which is the most likable robot?

Yellow Robot | Blue Robot | Green Robot

Which is the most knowledgeable robot?

Yellow Robot | Blue Robot | Green Robot

In which round was the receptionist robot the most likeable? (Check the box next to your choice.)

- ❏ Round 1 - (Robot as guide)
- ❏ Round 2 - (Robot as guide)
- ❏ Round 3 - (Robot as guide)

Which interaction with the receptionist robot did you find the least preferable? (Check the box next to your choice.)

- ❏ Round 1 - (Robot as guide)
- ❏ Round 2 - (Robot as guide)
- ❏ Round 3 - (Robot as guide)

Part 4. Please answer the question below in as much detail as you can provide.

1) How do you think information is transferred between the two robots?

____

____

____

Participant ID (Number):

Part 1. Please answer the following questions:

1) What is your age? ____

2) What is your native language? _____________________________________

3) What other languages do you speak? _______________________________

4) What is your gender? ___________________

5) How often do you use a computer (answer below on a scale from 1 to 7) ?

Never 1 | 2 | 3 | 4 | 5 | 6 | 7 Often (daily)

6) How familiar are you with robots (answer below on a scale from 1 to 7) ?

Not at all 1 | 2 | 3 | 4 | 5 | 6 | 7 Very familiar

7) What is your occupation (field of study if student)? __________________

8) Do you own a pet? ▢ No ▢ Yes

If yes, What kind of pet do you own? _________________

9) Do you own a smart assistant (Amazon Echo, Google Home, etc)? ▢ No ▢ Yes

If yes, What kind of smart assistant do you own? _________________

9) Do you own a robot? ▢ No ▢ Yes

If yes, What kind of robot do you own? _________________

Multi Depth Camera Calibration B.

In Chapter 6 (Interactive System), our system used multiple depth cameras to track participants in the scene. To enable this tracking, we needed to know the position and orientation of each camera in the world[1]. This short chapter describes how we calculated the relative position of each camera. Our calibration process was similar to the one described in [140].

1: This process is also known as extrinsic calibration.

B.1. Procedure

Our calibration pipeline uses \Psi and OpenCV. To find the position of each camera, we first recorded a 30-second video of a 4x5 Charuco board (Figure B.1) that was visible to all the cameras. The color images from each camera were then processed by OpenCV's `aruco` contrib library that detected the Charuco board and returned an estimated pose of the board in each image if found. Because we were interested in the body tracker result that used the depth camera, all poses were first transformed from the color frame of reference to the depth camera frame of reference. This was done by using the factory-provided transformation matrix.

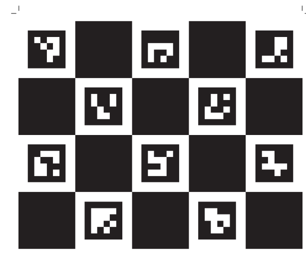


Figure B.1.: The Charuco board used for calibration.

We then synchronized the poses from all cameras and created pairs of poses from the same time point. Each pair was at most 10 milliseconds apart according to the capture times. Given four cameras, we had 6 pairs ($\binom{4}{2}$) of the poses. Not all pairs always had values as the board was not always visible at the same time for certain pairs. For example, in our laboratory study, the board was not visible to camera 1 and camera 3 at the same time.

Because both poses are describing the same object, just in different camera's frames of reference, we can use the following formula to find the transformation that links the cameras:

$$\hat{P}_{cam_A} T_{cam_A \rightarrow cam_B} = \hat{P}_{cam_B} \tag{B.1}$$

where P is the pose matrix (4×4) of the board in each camera frame, $\hat{P}$ describes the inverse of the pose matrix, and $T_{cam_A \rightarrow cam_B}$ (4×4 matrix) is the transformation matrix that describes the relationship between the two cameras and the matrix we are trying to solve for.

Assuming that we have n number of pairings for a pair of cameras, we can use the above equation to find the transformation matrix between the cameras. Because the last row of a transformation matrix is fixed $\begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$, we only have to solve for 12 elements in the transformation matrix. We can reconstruct B.1 with T as a single column vector:

$$\begin{bmatrix} P_{a11}, P_{a12}, P_{a13}, 0,0,0,0,0,0,0,0,0 \\ 0,0,0,P_{a11}, P_{a12}, P_{a13}, 0,0,0,0,0,0 \\ 0,0,0,0,0,0,P_{a11}, P_{a12}, P_{a13}, 0,0,0 \\ 0,0,0,0,0,0,0,0,0,P_{a11}, P_{a12}, P_{a13} \\ P_{a21}, P_{a22}, P_{a23}, 0,0,0,0,0,0,0,0,0 \\ 0,0,0,P_{a21}, P_{a22}, P_{a23}, 0,0,0,0,0,0 \\ 0,0,0,0,0,0,P_{a21}, P_{a22}, P_{a23}, 0,0,0 \\ 0,0,0,0,0,0,0,0,0,P_{a21}, P_{a22}, P_{a23} \\ P_{a31}, P_{a32}, P_{a33}, 0,0,0,0,0,0,0,0,0 \\ 0,0,0,P_{a31}, P_{a32}, P_{a33}, 0,0,0,0,0,0 \\ 0,0,0,0,0,0,P_{a31}, P_{a32}, P_{a33}, 0,0,0 \\ 0,0,0,0,0,0,0,0,0,P_{a31}, P_{a32}, P_{a33} \end{bmatrix} \begin{bmatrix} T_{11} \\ T_{21} \\ T_{31} \\ T_{12} \\ T_{22} \\ T_{32} \\ T_{13} \\ T_{23} \\ T_{33} \\ T_{14} \\ T_{24} \\ T_{34} \end{bmatrix} = \begin{bmatrix} P_{b11} \\ P_{b12} \\ P_{b13} \\ P_{b14} - P_{a14} \\ P_{b21} \\ P_{b22} \\ P_{b23} \\ P_{b24} - P_{a24} \\ P_{b31} \\ P_{b32} \\ P_{b33} \\ P_{b34} - P_{a34} \end{bmatrix} \tag{B.2}$$

As we have multiple pairings of P_{cam_A} and P_{cam_B}, we created an overdetermined system where the left-side "A" matrix has a size of ($N \times 12$) and the right-side "B" matrix has a size of ($12N \times 1$).

While a single pairing was sufficient to solve the equation, we chose to use multiple pairings to minimize systematic errors and noise. With N number of pairings, we used a least squares solver in NumPy's linear algebra package [141] to solve the equations above. In our application, we found that the variance could vary based on the position of the camera. For camera pairings that we cared about, we often found our residuals to be around 1.0 with 360 pairs (4320 rows). In practice, when validating whether the system was correct, we looked for an average residual [2] of less than 0.05.

2: Total residual divided by the number of pairs.

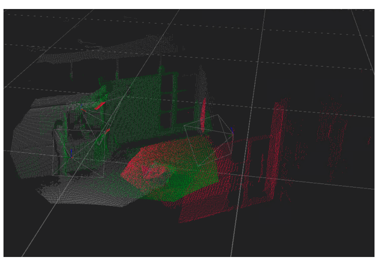


Figure B.2.: The combination of the pointclouds from all three Azure Kinect cameras in the hallway after calibration.

The least squares would give an estimate of the transformation matrix but it was likely improper as the rotation matrix was not orthogonal (special orthogonal group 3). To normalize the rotation matrix in our estimated transformation matrix, we used singular value decomposition:

$$R_{\text{init}} = U\Sigma V^T \tag{B.3}$$

$$R_{\text{final}} = UV^T \tag{B.4}$$

where R is the rotation matrix for the initial and final result of the transformation matrix. As U and V^T are orthogonal matrices, the resulting matrix will be orthogonal.

With a completed transformation matrix, we now have pairings between two cameras. This process was repeated for all of the cameras. We then choose one camera as the origin of the world frame[3]. In the lab study, it was the side camera. In the field study, it was the camera directly above Roboceptionist. We then found a chain of transformation from the origin camera to each camera in the scene. The combination of the chains gave us the transformation matrix that described the spatial relationship between each camera and the world frame.

3: Because we wanted the world frame to be at the base of the camera, we manually measured the height of the camera and included another transformation that corrected for its height.

We also conducted an additional validation step where we visually inspected the combined point clouds from all cameras and checked if they were aligned.

Social Navigation System C.

In Chapter 6 (Interactive System), the mobile robot needed to navigate around humans and avoid moving through the O-space of the interaction group. We used ROS Navigation [98] as our underlying navigation system. ROS Navigation is a modular framework that provides a common interface to connect different components (e.g., global planner, local planner, costmaps) commonly used in mobile robot navigation. It also provided out-of-the-box planners and costmaps.

In our studies, we used a lattice-based global planner [132] provided in the SBPL library as our global planner and a trajectory rollout local planner. In addition to fine-tuning both the motion primatives in the lattice global planner and parameters in the planners, we also added two additional costmaps to account for human position and group dynamics. The two costmaps were added to the default costmaps created from static map, the obstacles detected by the laser scanner, and the costmap inflation.

C.1. Human Costmap

This costmap added additional cost to the area around humans detected by the environment depth cameras. This layer is similar to ones described in prior work [142] where humans were treated differently from other dynamic obstacles. Our work differed from the prior work in that our costmap listened to our `/humans` ROS topic that provides a whole body skeleton and also integrated orientation into our model. Using the position and orientation of the person, we generated an ellipse with the major axis parallel to the direction in which the person was facing. We also moved the region forward such that the region was larger in front of the person. The ellipse also was elongated forward when the person was moving. We then used that region to update the costmap. The regions that the human ellipses occupied were deemed to be invalid for navigation purposes.

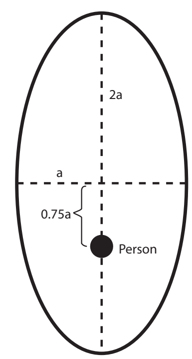


Figure C.1.: An illustration of the human ellipse model with our default parameter. In our studies, the value of a was set to 0.4. The ellipse ratio and ellipse offsets were configurable by RQT reconfig.

C.2. Interaction Space Costmap

In addition to avoiding people in the scene, the mobile robot also avoided infringing upon any existing interaction groups. This was especially important in our user studies because the robot should not go through the interaction space between the person and the other robot. For example in Chapter 7, if the robot is driving away, it should not go though the space between the participant and the stationary robot as they were still interacting.

To address this, we created an Interaction Space costmap that encoded the active interactions involving the robots in the scene. The costmap received a list of groups from the interaction framework. Each group consisted of a center, a list of points where the robots/agents were, and a list of points for each human member in the group. In the group scenario (> 2 members), we created a polygon with the points of all members. To avoid the robot getting stuck due to the costmap overlapping with the robot, we added some slack ($0.3m$) for all robot members. In dyadic cases (= 2 members), we chose to generate a rectangle with the ends being the members in the group [1]. The rectangle has a width of 0.5m. We then assigned a very high cost (225) to the region. While high, the region was not treated as invalid because we wanted the robot to go through the region if it was the only feasible path.

1: In testing, we found the polygon method to be unreliable because it generated a line that did not impose a large enough cost difference to change the selected path.

C.3. Limitations & Future Work

One of our design limitations was that our system that was designed well for the lab did not transfer well when we tested it in the field with a high volume of people. We were forced to turn off the human costmap layer in the field study as the number of people in the area often led to the mobile robot being unable to find a valid path. We believe our model was too conservative and resulted in too many invalid regions

Furthermore, we only consider interactions involving our robot and not other potential groups in the scene. Future work can explore taking advantage of other group detection models [129, 143] and encoding those group into our costmaps.

We also believe that there were multiple opportunities that we did not fully explore. These include but are not limited to:

Integration of verbal/nonverbal behavior Currently, after receiving a target, our robot simply drives towards the target. There are opportunities for robots to use different cues to not only indicate where it is going but also whether it has completed its path. This was one of the common points for improvement bought up by participants in Chapter 8.

Violating Social Norms While we try to avoid moving through the O-space when possible, our planner might choose that path if it is the only valid path. When such violation occurs, what should the robot do and how it can mitigate such violations?

Better Interaction Model Currently, we constructed a simple polygon model with a fixed cost. While easy to implement, it lacked the flexibility and complexity of real-life interactions. A simple change is to have a dynamic cost and change it based on distance. It is likely to be more acceptable to go through a group if they are farther apart. A better model can also tell us where the robot should go through if they have to violate a group's O-space.

Behavior Machine D.

In this supplementary chapter, we describe how the Behavior Machine works and how it was used.

D.1. State

The core of the behavior machine is the `State` class. All states and components are derived from the base `State` class which allows any state to be used in place of another or have unlimited nested hierarchical states. When the `State` is started (either by a Machine or manually), the execute function is run by itself in another thread.

To create a new state, one simply has to overwrite the base `execute` function.

```python
from behavior_machine.core import State, Board

class PrintState(State):

    def execute(self, board: Board) -> typing.Optional[StateStatus]:
        print("Hello World")
```

Listing D.1: State Example

As developers can also use patterns from Behavior Trees, each state can also return an optional `StateStatus` that describes whether or not the state was successful. The `StateStatus` enum has values of "success", "failed", "interrupted", "exception" (when an internal exception happened), or "NOT_SPECIFIED" (if the function did not return any status) at the end of execution [1].

1: If the status of a state is queried when it is running, it returns the "RUNNING" value instead.

D.2. Information Passing Between States

There are a few ways to pass information between states. The easiest way is through the `Board` object that is passed to the `execute` function. The instance of the board object is shared with every single state in an instance of the behavior machine. The object is similar to other memory objects, such as those in Behavior Trees. Any variable or objects can be placed or retrieved through the Board with a string identifier using the thread-safe 'get' or 'set' functions. In addition to the Board function, `State` also has optional `flow_in` and `flow_out` variables that allow a state to pass values to the immediate next state [2]. Lastly, because the program is implemented in Python, programmers can get around the state design and implement static information, class members or non-local variables that are set in the code. While these patterns are possible, they carry the risk of race conditions and undefined behavior when rerunning the state.

2: For the standard BT patterns, the flow mechanism works differently for each pattern. For example, a flowed-in value for a `Parallel` state is duplicated and passed to all internal states. Details about this can be found in the code base.

D.3. State Transitions & Machine

Similar to other state machine designs, state changes are done when certain conditions are satisfied. The checking of conditions is done through the `transition(evaluation_func, nextState)` member function in each state. The 'evaluation_func' is a function that takes in the current state and board and returns a boolean value that expresses whether the transition should be taken or not. Inside the state, each transition is evaluated in the order that it was added.

The frequency with which the transition functions are evaluated depends on the frequency of the `Machine`. `Machine` is a special type of `State` that takes in an initial state, a rate of transition evaluation, and an optional end-state IDs [3]; when started, it evaluates and transitions the internal state if the transition function is satisfied. As the `Machine` is just another `State`, it can be placed in any other position and nested if needed.

3: The machine accepts additional parameters that help with debugging and evaluation, but we omit them here for simplicity.

D.4. Interruption

One of our primary goals was to enable fast changing of states and allow sudden changes in the control flow (for example, if an interactor leaves in the middle of the interaction). The challenge with interruptions is not only that the interruption needs to be propagated to all states, but each running state also need to handle them gracefully. To enable interruption, the `State` class consists of `Interrupt()` and `isInterrupted()` methods as well as an interrupted event flag. When a state might take longer than 100ms or it involves waiting or polling, the state has to implement ways to check if it's being interrupted. To offload and speed up the polling, we used the `threading.event` flags in the Python library [4]. For example, in a lengthy `execute` operation, the state will periodically check if the state has been interrupted through the `isInterrupted()` method and it will quickly wrap up the operations and end its `execute` function.

4: Since Python3, the wait method is implemented in C instead of Python, which improved performance.

D.5. Behavior Trees Pattern

As previously mentioned, our codebase was also inspired by the pattern used in Behavior Trees. We borrowed similar ideas and implemented four common patterns that were used in behavior trees. Patterns allows us to quickly link up and separate states without manually creating the connections. As a pattern is a wrapper state, it can be nested or treated as its own state with transitions. The main patterns in our system are:

SequentialState This is also known as Sequence Node in the Behavior Trees literature. The state will run all internal states sequentially in the order that they are added as long as they are successful.

SelectorState This is also known as Selector Node in the Behavior Trees literature. The state will run all internal state sequentially until one of the states returns success.

ParallelState This is also known as Parallel Node in the Behavior Trees literature. All internal states are run simultaneously and this state only completes once all internal states have completed. The return state would be 'failure' if at least one state is 'failed'. This state has the property of continuing to run even when all but one state has ended.

AtLeastOneState This is similar to the ParallelState but will stop immediately (and interrupt the remaining state(s)) if one of the internal states completes [5].

5: We call it `AtLeastOneState` because we know at least one state has completed.

D.6. Illustrative Example

The following is a short illustrative example modified from our study code of how a behavior machine operates:

Listing D.2: Illustrative Example

```

r1 = RoboceptSpeakState("Hi, welcome to Goliath National Bank")
r2 = RoboceptSpeakState("For security purposes, can you provide me with your
      first name?")

robot_conversation = SequentialState("robot_conversation", children=[
    RoboceptSpeakConstructState("Hi {0}.", "person_name", target_frame_id="
     interactor"),
    PodiMoveToLocation("location_key"),
    RoboceptSpeakState("Hi Podi, please bring {0} to the conference room", "
     person_name", target_frame_id="podi")
    PodiSpeakState("Okay.", target_frame_id="robocept")
    ParallelState("lookAtInteractor", children=[
        LookAtPersonMachine("lookAt-r", "robocept", "interactor"),
        LookAtPersonMachine("lookAt-p", "podi", "interactor")
    ]),
    RoboceptSpeakConstructState("{0} will guide you to {1}. Please follow
     {0}", [
                                "podi_name", "fixed_location"]),
    PodiSpeakState("Please follow me.")
])
end = IdleState("end")

# add the connections and transitions

r1.add_transition_on_success(r2)

r2.add_transition(transition_when_utterances_contains(USER_NAME_LIST),
     robot_conversation)
r2.add_transition_after_elapsed(r2)
r2.add_transition(transition_when_utterances_contains_intent("repeat", r2)

machine = Machine(r1, end_state_ids=["end"], rate=30)
```

In the example above, the machine first commanded Roboceptionist to welcome the person and ask for their name. The state then transitioned based on whether the name matched a pre-programmed name, if they ask for the robot to repeat the question, or a certain amount of time has passed.

Once the system verified the person, Roboceptionist proceeded to have a structured dialog with the person and summoned the mobile robot, Podi. It used `ParallelState` to make the two robots look at the person at the same time before Podi asked to lead the participant away.

E. Questionnaires for Chapter 7

The following pages are the questionnaires used in Chapter 7 (Spatial Formation in Person Transfers). It was administered through Qualtrics Survey Software and participants completed the survey on a provided laptop.

Experimenter Setup

What is the condition?

Participant ID:

Welcome to TBD Lab Study -- Version 1.2

First Round

This is scenario A.

How comfortable were you when interacting with the **mobile robot**?

Extremely uncomfortable	Moderately uncomfortable	Slightly uncomfortable	Neither comfortable nor uncomfortable	Slightly comfortable	Moderately comfortable	Extremely comfortable

How comfortable were you when interacting with the **standing robot**?

Extremely uncomfortable	Moderately uncomfortable	Slightly uncomfortable	Neither comfortable nor uncomfortable	Slightly comfortable	Moderately comfortable	Extremely comfortable

How easy/difficult was the task?

Extremely difficult	Moderately difficult	Slightly difficult	Neither easy nor difficult	Slightly easy	Moderately easy	Extremely easy

How confident were you on what to do next during your interaction with the robot?

Extremely unconfident	Moderately unconfident	Slightly unconfident	Neither confident nor unconfident	Slightly confident	Moderately confident	Extremely confident

What did you like about the robots or the way they behaved?

What do you wish was different about the robots or the way they behaved?

Please wait for the experimenter in your seat. They will be with you shortly.

Second Round

This is scenario B.

How comfortable were you when interacting with the **standing robot**?

Extremely uncomfortable	Moderately uncomfortable	Slightly uncomfortable	Neither comfortable nor uncomfortable	Slightly comfortable	Moderately comfortable	Extremely comfortable

How easy/difficult was the task?

Extremely difficult	Moderately difficult	Slightly difficult	Neither easy nor difficult	Slightly easy	Moderately easy	Extremely easy

How confident were you on what to do next during your interaction with the robot?

Extremely unconfident	Moderately unconfident	Slightly unconfident	Neither confident nor unconfident	Slightly confident	Moderately confident	Extremely confident

What did you like about the robot or the way it behaved?

What do you wish was different about the robot or the way it behaved?

Please wait for the experimenter in your seat. They will be with you shortly.

In-Between

This is ${lm://Field/1}.

How comfortable were you when interacting with the **mobile robot**?

Extremely uncomfortable	Moderately uncomfortable	Slightly uncomfortable	Neither comfortable nor uncomfortable	Slightly comfortable	Moderately comfortable	Extremely comfortable

How comfortable were you when interacting with the **standing robot**?

Extremely uncomfortable	Moderately uncomfortable	Slightly uncomfortable	Neither comfortable nor uncomfortable	Slightly comfortable	Moderately comfortable	Extremely comfortable

How easy/difficult was the task?

Extremely difficult	Moderately difficult	Slightly difficult	Neither easy nor difficult	Slightly easy	Moderately easy	Extremely easy

How confident were you on what to do next during your interaction with the robots

Extremely unconfident	Moderately unconfident	Slightly unconfident	Neither confident nor unconfident	Slightly confident	Moderately confident	Extremely confident

What did you like about the robots or the way they behaved?

What do you wish was different about the robots or the way they behaved?

Please wait for the experimenter in your seat. They will be with you shortly.

Final Questions

Using the scale provided, how closely are the words below associated with the **mobile robot.**

1 = Definitely not associated 9 = Definitely associated

	1	2	3	4	5	6	7	8	9
Awful									
Dangerous									
Aggressive									
Competent									
Social									
Strange									
Reliable									
Compassionate									
Happy									
Awkward									
Capable									
Emotional									
Knowledgable									
Feeling									
Responsive									
Scary									
Interactive									
Organic									

Using the scale provided, how closely are the words below associated with the **standing robot.**

1 = Definitely not associated 9 = Definitely associated

	1	2	3	4	5	6	7	8	9
Capable									
Competent									
Reliable									
Feeling									
Scary									
Social									
Aggressive									
Responsive									
Strange									
Interactive									
Happy									
Emotional									
Organic									
Awful									
Knowledgable									
Compassionate									
Awkward									
Dangerous									

Block 8

Please rate your agreement with the following statements:

Other Choices

The mobile robot stood really close to me.

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

The mobile robot's position made me uncomfortable.

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

I have trouble seeing the mobile robot.

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

The mobile robot movement was predictable.

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

I was able to predict the position of the mobile robot.

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

The mobile robot position itself too far away from me.

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

The mobile robot's position was socially appropriate.

Strongly disagree	Disagree	Somewhat disagree	Neither agree nor disagree	Somewhat agree	Agree	Strongly agree

Distance Question

When gauging distances, are you more familiar with meters or feet?

Meters

Feet

Please estimate how far away each robot was relative to you in feet during the last 4 scenarios.

You (feet)

	0	1	2	3	4	5	6	7	8	9	10
Mobile Robot											
Standing Robot											

Please estimate how far away each robot was relative to you in meters during the last 4 scenarios.

You (meters)

	0	0.5	1	1.5	2	2.5	3	3.5	4	4.5	5
Mobile Robot											
Standing Robot											

Demographic Survey

What is your age?

What is your native language?

What other languages do you speak?

What is your gender?

[Optional] Are there other aspects of your identity that are important to you (racial, ethnic, or otherwise)?

How often do you use a computer?

Never	A little	A moderate amount	A lot	Often (daily)

How often do you use a smart voice assistant (Siri, Alexa, etc)?

Never	A little	A moderate amount	A lot	Often (daily)

How familiar are you with robots?

Not at all	Slightly familiar	Moderately familiar	Very familiar	Extremely familiar

What is your occupation (field of study if student)?

Do you own a pet?

No

Yes

What kind of pet do you have?

Do you own a smart assistant (Amazon Echo, Google Home, etc)

No

Yes

What kind of smart assistant do you own?

Do you own a robot?

No

Yes

What kind of robot do you own?

Have you interact with the Roboceptionist on Carnegie Mellon University before?

Yes

No

I don't know what Roboceptionist is.

Bibliography

Here are the references in citation order.

[1] Jenny Medeiros. *How robots in retail increase sales and customer engagement*. May 2018. URL: https://www.modev.com/blog/want-to-increase-customer-engagement-use-a-robot (cited on page 1).

[2] Marketta Niemelä, Päivi Heikkilä, and Hanna Lammi. 'A social service robot in a shopping mall: expectations of the management, retailers and consumers'. In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. 2017, pp. 227–228 (cited on page 1).

[3] Graeme Massie. *Meet the Army of robots that will help run the Tokyo Games*. July 2021. URL: https://www.independent.co.uk/sport/olympics/olympic-games-tokyo-robots-vehicles-b1884259.html (cited on page 1).

[4] Hamish Tennent, Solace Shen, and Malte Jung. 'Micbot: A Peripheral Robotic Object to Shape Conversational Dynamics and Team Performance'. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2019, pp. 133–142 (cited on page 1).

[5] Marynel Vázquez et al. 'Towards robot autonomy in group conversations: Understanding the effects of body orientation and gaze'. In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2017, pp. 42–52 (cited on pages 1, 12).

[6] Brennan Sellner et al. 'Coordinated multiagent teams and sliding autonomy for large-scale assembly'. In: *Proceedings of the IEEE* 94.7 (2006), pp. 1425–1444 (cited on pages 1, 2).

[7] Adam Rule and Jodi Forlizzi. 'Designing interfaces for multi-user, multi-robot systems'. In: *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*. ACM. 2012, pp. 97–104 (cited on pages 1, 23).

[8] Marlena R Fraune et al. 'Three's company, or a crowd?: The effects of robot number and behavior on HRI in Japan and the USA.' In: *Robotics: Science and Systems*. 2015 (cited on pages 1, 16).

[9] Nicole Salomons et al. 'Humans conform to robots: Disambiguating trust, truth, and conformity'. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2018, pp. 187–195 (cited on pages 1, 16).

[10] Shuhei Tsuchida, Tsutomu Terada, and Masahiko Tsukamoto. 'A System for Dancing with Multiple Robotic Balls'. In: *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology*. ACE '15. Iskandar, Malaysia: ACM, 2015, 40:1–40:3. DOI: 10.1145/2832932.2832943 (cited on page 1).

[11] Darren Guinness et al. 'RoboGraphics: Dynamic Tactile Graphics Powered by Mobile Robots'. In: *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. ACM. 2019, pp. 318–328 (cited on pages 1, 18).

[12] Tom Williams, Priscilla Briggs, and Matthias Scheutz. 'Covert robot-robot communication: Human perceptions and implications for human-robot interaction'. In: *Journal of Human-Robot Interaction* 4.2 (2015), pp. 24–49 (cited on pages 2, 11, 18, 45, 58, 59, 112).

[13] Brian R Duffy et al. 'Agent chameleons: Agent minds and bodies'. In: *Proceedings 11th IEEE International Workshop on Program Comprehension*. IEEE. 2003, pp. 118–125 (cited on page 2).

[14] Michal Luria et al. 'Re-Embodiment and Co-Embodiment: Exploration of social presence for robots and conversational agents'. In: *Proceedings of the 2019 on Designing Interactive Systems Conference*. ACM. 2019, pp. 633–644 (cited on pages 2, 9, 10, 42).

[15] Marlena R Fraune et al. 'threatening flocks and mindful snowflakes: How group entitativity affects perceptions of robots'. In: *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. IEEE. 2017, pp. 205–213 (cited on pages 2, 16).

[16] Piyush Khandelwal and Peter Stone. 'Multi-robot human guidance: Human experiments and multiple concurrent requests'. In: *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*. International Foundation for Autonomous Agents and Multiagent Systems. 2017, pp. 1369–1377 (cited on pages 2, 40, 108).

[17] Michal Luria et al. 'Social Boundaries for Personal Agents in the Interpersonal Space of the Home'. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 2020, pp. 1–12 (cited on page 3).

[18] Holly A Yanco and Jill L Drury. 'A taxonomy for human-robot interaction'. In: *Proceedings of the AAAI Fall Symposium on Human-Robot Interaction*. sn. 2002, pp. 111–119 (cited on page 9).

[19] Xiang Zhi Tan et al. 'Go That Way: Exploring Supplementary Physical Movements by a Stationary Robot When Providing Navigation Instructions'. In: *The 21st International ACM SIGACCESS Conference on Computers and Accessibility*. ASSETS '19. Pittsburgh, PA, USA: ACM, 2019, pp. 299–311. DOI: 10.1145/3308561.3353805 (cited on page 11).

[20] Shiri Azenkot, Catherine Feng, and Maya Cakmak. 'Enabling building service robots to guide blind people a participatory design approach'. In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2016, pp. 3–10 (cited on pages 11, 24).

[21] Xiang Zhi Tan and Aaron Steinfeld. 'Haptic Interaction for Human-Robot Communication Using a Spherical Robot'. In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2018, pp. 732–739 (cited on page 12).

[22] Amal Nanavati, Xiang Zhi Tan, and Aaron Steinfeld. 'Coupled Indoor Navigation for People Who Are Blind'. In: *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2018, pp. 201–202 (cited on page 12).

[23] Jirachaya "Fern" Limprayoon et al. 'Robot Trajectories When Approaching a User with a Visual Impairment'. In: *The 23rd International ACM SIGACCESS Conference on Computers and Accessibility*. ASSETS '21. Virtual Event, USA: Association for Computing Machinery, 2021. DOI: 10.1145/3441852.3476538 (cited on page 12).

[24] Allison Sauppé and Bilge Mutlu. 'Robot deictics: How gesture and context shape referential communication'. In: *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. ACM. 2014, pp. 342–349 (cited on page 12).

[25] Xiang Zhi Tan et al. 'Now, over here: Leveraging extended attentional capabilities in human-robot interaction'. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. 2020, pp. 468–470 (cited on pages 12, 63, 109).

[26] Xiang Zhi Tan et al. 'Inducing bystander interventions during robot abuse with social mechanisms'. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2018, pp. 169–177 (cited on pages 13, 95).

[27] Drazen Brscić et al. 'Escaping from children's abuse of social robots'. In: *Proceedings of the tenth annual acm/ieee international conference on human-robot interaction*. ACM. 2015, pp. 59–66 (cited on page 13).

[28] Todd Leopold. *HitchBOT, the hitchhiking robot, gets beheaded in Philadelphia*. 2015. URL: https://web.archive.org/web/20190617095558/https://www.cnn.com/2015/08/03/us/hitchbot-robot-beheaded-philadelphia-feat/index.html (visited on 11/02/2019) (cited on page 13).

[29] Cecilia G Morales et al. 'Interaction Needs and Opportunities for Failing Robots'. In: *Proceedings of the 2019 on Designing Interactive Systems Conference*. ACM. 2019, pp. 659–670 (cited on pages 13, 95).

[30] Mark Weiser. 'The Computer for the 21 st Century'. In: *Scientific american* 265.3 (1991), pp. 94–105 (cited on page 15).

[31] Frederik Brudy et al. 'Cross-Device Taxonomy: Survey, Opportunities and Challenges of Interactions Spanning Across Multiple Devices'. In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. ACM. 2019, p. 562 (cited on page 15).

[32] Giuseppe Ghiani, Fabio Paternò, and Carmen Santoro. 'On-demand cross-device interface components migration'. In: *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services*. ACM. 2010, pp. 299–308 (cited on page 15).

[33] David Dearman and Jeffery S Pierce. 'It's on my other computer!: computing with multiple devices'. In: *Proceedings of the SIGCHI Conference on Human factors in Computing Systems*. ACM. 2008, pp. 767–776 (cited on page 15).

[34] Stephanie Santosa and Daniel Wigdor. 'A field study of multi-device workflows in distributed workspaces'. In: *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM. 2013, pp. 63–72 (cited on page 15).

[35] Dimitrios Raptis, Jesper Kjeldskov, and Mikael B Skov. 'Continuity in multi-device interaction: an online study'. In: *Proceedings of the 9th Nordic Conference on Human-Computer Interaction*. ACM. 2016, p. 29 (cited on page 15).

[36] Peter H Kahn et al. 'The new ontological category hypothesis in human-robot interaction'. In: *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2011, pp. 159–160 (cited on page 15).

[37] Salvatore Andolina and Jodi Forlizzi. 'A multi-touch interface for multi-robot path planning and control'. In: *International Workshop on Symbiotic Interaction*. Springer. 2015, pp. 127–132 (cited on page 16).

[38] Terrence Fong, Charles Thorpe, and Charles Baur. 'Multi-robot remote driving with collaborative control'. In: *IEEE Transactions on Industrial Electronics* 50.4 (2003), pp. 699–704 (cited on page 16).

[39] Kexi Liu et al. 'Roboshop: multi-layered sketching interface for robot housework assignment and management'. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM. 2011, pp. 647–656 (cited on page 16).

[40] Meher T Shaikh and Michael A Goodrich. 'Design and evaluation of adverb palette: A gui for selecting tradeoffs in multi-objective optimization problems'. In: *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. IEEE. 2017, pp. 389–397 (cited on page 16).

[41] Pragathi Praveena et al. 'Characterizing Input Methods for Human-to-robot Demonstrations'. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2019, pp. 344–353 (cited on page 16).

[42] Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 'A motion retargeting method for effective mimicry-based teleoperation of robot arms'. In: *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2017, pp. 361–370 (cited on page 16).

[43] Oliver Liu et al. 'Understanding human-robot interaction in virtual reality'. In: *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2017, pp. 751–757 (cited on page 16).

[44] Nhan Tran, Josh Rands, and Tom Williams. 'A hands-free virtual-reality teleoperation interface for wizard-of-oz control'. In: *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*. 2018 (cited on page 16).

[45] David Whitney et al. 'Ros reality: A virtual reality framework using consumer-grade hardware for ros-enabled robots'. In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE. 2018, pp. 1–9 (cited on page 16).

[46] Michael Walker et al. 'Communicating robot motion intent with augmented reality'. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2018, pp. 316–324 (cited on page 16).

[47] Brian Scassellati et al. 'Teaching language to deaf infants with a robot and a virtual human'. In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM. 2018, p. 553 (cited on pages 16, 109).

[48] Marlena R Fraune et al. 'Rabble of robots effects: Number and type of robots modulates attitudes, emotions, and stereotypes'. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2015, pp. 109–116 (cited on page 16).

[49] Solomon Asch and Harold Guetzkow. 'Effects of group pressure upon the modification and distortion of judgments'. In: *Documents of gestalt psychology* (1951), pp. 222–236 (cited on page 16).

[50] Clay Beckner et al. 'Participants conform to humans but not to humanoid robots in an English past tense formation task'. In: *Journal of Language and Social Psychology* 35.2 (2016), pp. 158–179 (cited on page 16).

[51] Masahiro Shiomi and Norihiro Hagita. 'Do synchronized multiple robots exert peer pressure?' In: *Proceedings of the Fourth International Conference on Human Agent Interaction*. ACM. 2016, pp. 27–33 (cited on page 16).

[52] Masahiro Shiomi and Norihiro Hagita. 'Do the number of robots and the participant's gender influence conformity effect from multiple robots?' In: *Advanced Robotics* (2019), pp. 1–8 (cited on page 16).

[53] Joe Connolly et al. 'Prompting Prosocial Human Interventions in Response to Robot Mistreatment'. In: *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '20. Cambridge, United Kingdom: Association for Computing Machinery, 2020, pp. 211–220. DOI: 10.1145/3319502.3374781 (cited on page 17).

[54] Henny Admoni et al. 'Are you looking at me?: perception of robot attention is mediated by gaze type and group size'. In: *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*. IEEE Press. 2013, pp. 389–396 (cited on page 17).

[55] Raquel Oliveira et al. 'Friends or Foes?: Socioemotional Support and Gaze Behaviors in Mixed Groups of Humans and Robots'. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '18. Chicago, IL, USA: ACM, 2018, pp. 279–288. DOI: 10.1145/3171221.3171272 (cited on page 17).

[56] Marlena R. Fraune et al. 'Is Human-Robot Interaction More Competitive Between Groups Than Between Individuals?' In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Daegu, Korea (South): IEEE, Mar. 2019, pp. 104–113. DOI: 10.1109/HRI.2019.8673241 (cited on page 17).

[57] Samantha Reig et al. 'Flailing, Hailing, Prevailing: Perceptions of Multi-Robot Failure Recovery Strategies'. In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 2021, pp. 158–167 (cited on pages 17, 59).

[58] Takayuki Kanda et al. 'Multi-robot cooperation for human-robot communication'. In: *Proceedings. 11th IEEE International Workshop on Robot and Human Interactive Communication*. IEEE. 2002, pp. 271–276 (cited on pages 17, 49).

[59] Shujie Zhou and Leimin Tian. 'Would You Help a Sad Robot? Influence of Robots' Emotional Expressions on Human-Multi-Robot Collaboration'. In: *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. Naples, Italy: IEEE, Aug. 2020, pp. 1243–1250. DOI: 10.1109/RO-MAN47096.2020.9223524 (cited on page 17).

[60] K. Hayashi et al. 'Robot Manzai - robots' conversation as a passive social medium'. In: *5th IEEE-RAS International Conference on Humanoid Robots, 2005.* Dec. 2005, pp. 456–462. DOI: 10.1109/ICHR.2005.1573609 (cited on page 17).

[61] Masahiro Shiomi et al. 'Interactive humanoid robots for a science museum'. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM. 2006, pp. 305–312 (cited on page 17).

[62] Kotaro Hayashi et al. 'Humanoid robots as a passive-social medium-a field experiment at a train station'. In: *2007 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2007, pp. 137–144 (cited on page 17).

[63] Takamasa Iio, Yuichiro Yoshikawa, and Hiroshi Ishiguro. 'Pre-scheduled turn-taking between robots to make conversation coherent'. In: *Proceedings of the Fourth International Conference on Human Agent Interaction*. ACM. 2016, pp. 19–25 (cited on page 17).

[64] Marlena Fraune and Selma Šabanović. 'Robot gossip: Effects of mode of robot communication on human perceptions of robots'. In: *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2014, pp. 160–161 (cited on page 18).

[65] Magnus Söderlund. 'The Robot-to-Robot Service Encounter: An Examination of the Impact of Inter-Robot Warmth'. In: *Journal of Services Marketing* 35.9 (Dec. 2021), pp. 15–27. DOI: 10.1108/JSM-01-2021-0006 (cited on page 18).

[66] Masahiro Shiomi et al. 'Field trial of networked social robots in a shopping mall'. In: *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2009, pp. 2846–2853 (cited on pages 18, 40).

[67] Iolanda Leite et al. 'Emotional storytelling in the classroom: Individual versus group interaction between children and robots'. In: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2015, pp. 75–82 (cited on page 18).

[68] Marynel Vázquez et al. 'Spatial and Other Social Engagement Cues in a Child-robot Interaction: Effects of a Sidekick'. In: *Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction*. HRI '14. Bielefeld, Germany: ACM, 2014, pp. 391–398. DOI: 10.1145/2559636.2559684 (cited on pages 18–20).

[69] Edward T Hall and ET Hall. 'The hidden dimension, vol. 1990'. In: *NY: Anchor Books* (1969) (cited on page 19).

[70] Agnieszka Sorokowska et al. 'Preferred interpersonal distances: a global comparison'. In: *Journal of Cross-Cultural Psychology* 48.4 (2017), pp. 577–592 (cited on page 19).

[71] Michael L Walters et al. 'The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment'. In: *ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005.* IEEE. 2005, pp. 347–352 (cited on page 19).

[72] Michael L Walters et al. 'Human approach distances to a mechanical-looking robot with different robot voice styles'. In: *RO-MAN 2008-The 17th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE. 2008, pp. 707–712 (cited on page 19).

[73] Leila Takayama and Caroline Pantofaru. 'Influences on proxemic behaviors in human-robot interaction'. In: *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2009, pp. 5495–5502 (cited on page 19).

[74] Jonathan Mumm and Bilge Mutlu. 'Human-robot proxemics: physical and psychological distancing in human-robot interaction'. In: *Proceedings of the 6th international conference on Human-robot interaction*. ACM. 2011, pp. 331–338 (cited on page 19).

[75] Adam Kendon. *Conducting interaction: Patterns of behavior in focused encounters*. Vol. 7. CUP Archive, 1990 (cited on pages 19, 75, 76, 88, 92).

[76] Bokyung Lee et al. 'An Empirical Study of How Socio-Spatial Formations Are Influenced by Interior Elements and Displays in an Office Context'. In: *Proceedings of the ACM on Human-Computer Interaction* 3.CSCW (Nov. 2019), 58:1–58:26. DOI: 10.1145/3359160 (cited on page 19).

[77] Hideaki Kuzuoka et al. 'Reconfiguring spatial formation arrangement by robot body orientation'. In: *Proceedings of the 5th ACM/IEEE international conference on Human-robot interaction*. IEEE Press. 2010, pp. 285–292 (cited on page 19).

[78] Marynel Vázquez, Aaron Steinfeld, and Scott E Hudson. 'Maintaining awareness of the focus of attention of a conversation: A robot-centric reinforcement learning approach'. In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE. 2016, pp. 36–43 (cited on page 19).

[79] Dan Bohus, Sean Andrist, and Eric Horvitz. 'A study in scene shaping: Adjusting F-formations in the wild'. In: *Proceedings of the 2017 AAAI Fall Symposium: Natural Communication for Human-Robot Collaboration*. 2017 (cited on page 19).

[80] Takahiro Matsumoto et al. 'Where should robots talk?: Spatial arrangement study from a participant workload perspective'. In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. ACM. 2018, pp. 270–278 (cited on page 20).

[81] Xiang Zhi Tan et al. 'Charting Sequential Person Transfers Between Devices, Agents, and Robots'. In: *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. 2021, pp. 43–52 (cited on page 23).

[82] Xiang Zhi Tan, Michal Luria, and Aaron Steinfeld. 'Defining Transfers Between Multiple Service Robots'. In: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '20. Cambridge, United Kingdom: Association for Computing Machinery, 2020, pp. 465–467. doi: 10.1145/3371382.3378258 (cited on page 23).

[83] John Zimmerman, Jodi Forlizzi, and Shelley Evenson. 'Research through design as a method for interaction design research in HCI'. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2007, pp. 493–502 (cited on page 23).

[84] Hee Rin Lee and Laurel D Riek. 'Reframing assistive robots to promote successful aging'. In: *ACM Transactions on Human-Robot Interaction (THRI)* 7.1 (2018), pp. 1–23 (cited on page 24).

[85] Dan Lockton et al. 'New Metaphors: A Workshop Method for Generating Ideas and Reframing Problems in Design and Beyond'. In: *Proceedings of the 2019 on Creativity and Cognition*. C&C '19. San Diego, CA, USA: ACM, 2019, pp. 319–332. doi: 10.1145/3325480.3326570 (cited on page 25).

[86] Hugh Beyer and Karen Holtzblatt. *Contextual design: A customer-centered approach to systems designs*. 1997 (cited on pages 25, 32).

[87] Charles W Bytheway. *FAST creativity and innovation: Rapidly improving processes, product development and solving complex problems*. J. Ross Publishing, 2007 (cited on page 32).

[88] Leonard L Berry and Sandra K Lampo. 'Teaching an old service new tricks: The promise of service redesign'. In: *Journal of Service Research* 2.3 (2000), pp. 265–275 (cited on page 34).

[89] Jennifer Goetz, Sara Kiesler, and Aaron Powers. 'Matching robot appearance and behavior to tasks to improve human-robot cooperation'. In: *The 12th IEEE International Workshop on Robot and Human Interactive Communication, 2003. Proceedings. ROMAN 2003.* Ieee. 2003, pp. 55–60 (cited on page 38).

[90] Valerie K Sims et al. 'Anthropomorphism of robotic forms: A response to affordances?' In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 49. 3. SAGE Publications Sage CA: Los Angeles, CA. 2005, pp. 602–605 (cited on page 38).

[91] Kristin E Schaefer et al. 'Classification of robot form: Factors predicting perceived trustworthiness'. In: *Proceedings of the human factors and ergonomics society annual meeting*. Vol. 56. 1. SAGE Publications Sage CA: Los Angeles, CA. 2012, pp. 1548–1552 (cited on page 38).

[92] Harel Yedidsion et al. 'Optimal Use of Verbal Instructions for Multi-robot Human Navigation Guidance'. In: *International Conference on Social Robotics*. Springer. 2019, pp. 133–143 (cited on pages 41, 108).

[93] Michita Imai, Tetsuo Ono, and Tameyuki Etani. 'Agent migration: communications between a human and robot'. In: *IEEE SMC'99 Conference Proceedings. 1999 IEEE International Conference on Systems, Man, and Cybernetics (Cat. No. 99CH37028)*. Vol. 4. IEEE. 1999, pp. 1044–1048 (cited on page 42).

[94] Xiang Zhi Tan et al. 'From One to Another: How Robot-Robot Interaction Affects Users' Perceptions Following a Transition Between Robots'. In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2019, pp. 114–122 (cited on pages 45, 95).

[95] Rachel Kirby, Jodi Forlizzi, and Reid Simmons. 'Affective social robots'. In: *Robotics and Autonomous Systems* 58.3 (2010), pp. 322–332 (cited on pages 45, 65).

[96] Iolanda Leite, Carlos Martinho, and Ana Paiva. 'Social robots for long-term interaction: a survey'. In: *International Journal of Social Robotics* 5.2 (2013), pp. 291–308 (cited on page 45).

[97] Alice Coucke et al. 'Snips Voice Platform: an embedded Spoken Language Understanding system for private-by-design voice interfaces'. In: *arXiv preprint arXiv:1805.10190* (2018), pp. 12–16 (cited on pages 49, 68).

[98] Eitan Marder-Eppstein et al. 'The office marathon: Robust navigation in an indoor office environment'. In: *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE. 2010, pp. 300–307 (cited on pages 49, 78, 131).

[99] Russell Toris et al. 'Robot Web Tools: Efficient messaging for cloud robotics.' In: *IROS*. 2015, pp. 4530–4537 (cited on page 49).

[100] Colleen M Carpinella et al. 'The robotic social attributes scale (rosas): Development and validation'. In: *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*. ACM. 2017, pp. 254–262 (cited on pages 51, 82, 84).

[101] Christoph Bartneck et al. 'Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots'. In: *International journal of social robotics* 1.1 (2009), pp. 71–81 (cited on page 51).

[102] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury. 'Foundations for an empirically determined scale of trust in automated systems'. In: *International Journal of Cognitive Ergonomics* 4.1 (2000), pp. 53–71 (cited on page 51).

[103] Mark H Davis et al. 'A multidimensional approach to individual differences in empathy'. In: (1980) (cited on page 52).

[104] HD Patterson. 'Maximum likelihood estimation of components of variance'. In: *Proceeding Eight International Biometric Conference, 1975*. Biometric Soc. 1975 (cited on page 53).

[105] Walter W Stroup. *Generalized linear mixed models: modern concepts, methods and applications*. CRC Press, 2016 (cited on page 53).

[106] Aaron Steinfeld, Odest Chadwicke Jenkins, and Brian Scassellati. 'The oz of wizard: simulating the human for interaction research'. In: *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM. 2009, pp. 101–108 (cited on page 59).

[107] Dylan F Glas et al. 'The network robot system: enabling social human-robot interaction in public spaces'. In: *Journal of Human-Robot Interaction* 1.2 (2012), pp. 5–32 (cited on pages 63, 64).

[108] David Kortenkamp, Reid Simmons, and Davide Brugali. 'Robotic systems architectures and programming'. In: *Springer Handbook of Robotics*. Springer, 2016, pp. 283–306 (cited on pages 63, 65).

[109] Sebastian Thrun et al. 'MINERVA: A second-generation museum tour-guide robot'. In: *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*. Vol. 3. IEEE. 1999 (cited on page 64).

[110] Illah R Nourbakhsh, Clayton Kunz, and Thomas Willeke. 'The mobot museum robot installations: A five year experiment'. In: *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*. Vol. 4. IEEE. 2003, pp. 3636–3641 (cited on page 64).

[111] Reid Simmons et al. 'Grace: An autonomous robot for the AAAI robot challenge'. In: *AI magazine* 24.2 (2003), pp. 51–51 (cited on page 64).

[112] R. Gockley et al. 'Designing Robots for Long-Term Social Interaction'. In: *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Aug. 2005, pp. 1338–1343. DOI: 10.1109/IROS.2005.1545303 (cited on pages 64, 65, 95, 96).

[113] Dan Bohus, Chit W Saw, and Eric Horvitz. 'Directions robot: in-the-wild experiences and lessons learned'. In: *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*. Citeseer. 2014, pp. 637–644 (cited on page 64).

[114] Mary Ellen Foster et al. 'MuMMER: Socially Intelligent Human-Robot Interaction in Public Spaces'. In: *arXiv:1909.06749 [cs]* (Sept. 2019) (cited on page 64).

[115] Justin Huang and Maya Cakmak. 'Code3: A system for end-to-end programming of mobile manipulator robots for novices and experts'. In: *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI*. IEEE. 2017, pp. 453–462 (cited on page 64).

[116] Yasaman S. Sefidgar et al. 'RobotIST: Interactive Situated Tangible Robot Programming'. In: *Proceedings of the Symposium on Spatial User Interaction*. SUI '18. New York, NY, USA: Association for Computing Machinery, Oct. 2018, pp. 141–149. DOI: 10.1145/3267782.3267921 (cited on page 64).

[117] David J. Porfirio et al. 'Figaro: A Tabletop Authoring Environment for Human-Robot Interaction'. In: *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 414. New York, NY, USA: Association for Computing Machinery, May 2021, pp. 1–15 (cited on page 64).

[118] David Porfirio et al. 'Bodystorming Human-Robot Interactions'. In: *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology*. UIST '19. New York, NY, USA: Association for Computing Machinery, Oct. 2019, pp. 479–491. DOI: 10.1145/3332165.3347957 (cited on page 64).

[119] Morgan Quigley et al. 'ROS: an open-source Robot Operating System'. In: *ICRA workshop on open source software*. Vol. 3. 3.2. Kobe, Japan. 2009, p. 5 (cited on page 64).

[120] Dan Bohus et al. 'Platform for situated intelligence'. In: *arXiv preprint arXiv:2103.15975* (2021) (cited on pages 65, 72).

[121] Kalin Stefanov et al. 'OpenSense: A Platform for Multimodal Data Acquisition and Behavior Perception'. In: *Proceedings of the 2020 International Conference on Multimodal Interaction*. 2020, pp. 660–664 (cited on page 65).

[122] Reid Simmons. *A Brief History of Our Research in Task-Level Control Architectures*. Mar. 2001. URL: https://web.archive.org/web/20200219182304/http://www.cs.cmu.edu/afs/cs/project/TCA/www/TCA-history.html (cited on page 65).

[123] Reid Simmons and David Apfelbaum. 'A task description language for robot control'. In: *Proceedings. 1998 IEEE/RSJ International Conference on Intelligent Robots and Systems. Innovations in Theory, Practice and Applications (Cat. No. 98CH36190)*. Vol. 3. IEEE. 1998, pp. 1931–1937 (cited on page 65).

[124] Terrence Fong et al. 'The human-robot interaction operating system'. In: *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. 2006, pp. 41–48 (cited on page 65).

[125] Damian Isla. *Handling complexity in the Halo 2 AI*. Mar. 2005. URL: https://www.gamedeveloper.com/programming/gdc-2005-proceeding-handling-complexity-in-the-i-halo-2-i-ai (cited on page 69).

[126] Michele Colledanchise and Lorenzo Natale. 'On the Implementation of Behavior Trees in Robotics'. In: *IEEE Robotics and Automation Letters* (2021) (cited on page 69).

[127] J. Gregory Trafton et al. 'ACT-R/E: An Embodied Cognitive Architecture for Human-Robot Interaction'. In: *J. Hum.-Robot Interact.* 2.1 (2013), pp. 30–55. DOI: 10.5898/JHRI.2.1.Trafton (cited on page 73).

[128] Matthias Scheutz et al. 'Novel mechanisms for natural human-robot interactions in the diarc architecture'. In: *Workshops at the twenty-seventh aaai conference on artificial intelligence*. 2013 (cited on page 73).

[129] Hooman Hedayati et al. 'REFORM: Recognizing F-Formations for Social Robots'. en. In: *arXiv:2008.07668 [cs]* (Aug. 2020), pp. 11181–11188. DOI: 10.1109/IROS45743.2020.9340708 (cited on pages 75, 132).

[130] Adam Kendon. 'Spacing and Orientation in Co-Present Interaction'. en. In: *Development of Multimodal Interfaces: Active Listening and Synchrony*. Ed. by Anna Esposito et al. Vol. 5967. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010, pp. 1–15. DOI: 10.1007/978-3-642-12397-9_1 (cited on page 76).

[131] Edward Twitchell Hall. *The hidden dimension*. Vol. 609. Garden City, NY: Doubleday, 1966 (cited on page 77).

[132] Maxim Likhachev and Dave Ferguson. 'Planning long dynamically feasible maneuvers for autonomous vehicles'. In: *The International Journal of Robotics Research* 28.8 (2009), pp. 933–945 (cited on pages 78, 131).

[133] Andy Field, Jeremy Miles, and Zoë Field. *Discovering statistics using R*. Sage publications, 2012 (cited on page 85).

[134] Maxim Makatchev and Reid Simmons. 'Incorporating a User Model to Improve Detection of Unhelpful Robot Answers'. In: *RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication*. Sept. 2009, pp. 973–978. DOI: 10.1109/ROMAN.2009.5326140 (cited on page 95).

[135] S. Sabanovic, M.P. Michalowski, and R. Simmons. 'Robots in the Wild: Observing Human-Robot Social Interaction Outside the Lab'. In: *9th IEEE International Workshop on Advanced Motion Control, 2006.* Mar. 2006, pp. 596–601. DOI: 10.1109/AMC.2006.1631758 (cited on page 95).

[136] Elizabeth Phillips et al. 'What is human-like? Decomposing robots' human-like appearance using the anthropomorphic robot (abot) database'. In: *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*. 2018, pp. 105–113 (cited on page 107).

[137] Piyush Khandelwal, Samuel Barrett, and Peter Stone. 'Leading the way: An efficient multi-robot guidance system'. In: *Proceedings of the 2015 international conference on autonomous agents and multiagent systems*. International Foundation for Autonomous Agents and Multiagent Systems. 2015, pp. 1625–1633 (cited on page 108).

[138] Paul Robinette et al. 'Overtrust of robots in emergency evacuation scenarios'. In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE. 2016, pp. 101–108 (cited on page 110).

[139] Pew Research Center. *Looking to the Future, Public Sees an America in Decline on Many Fronts*. en-US. Tech. rep. Washington, D.C., 2019 (cited on page 110).

[140] Bo Li et al. 'A multiple-camera system calibration toolbox using a feature descriptor-based calibration pattern'. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2013, pp. 1301–1307 (cited on page 129).

[141] Charles R. Harris et al. 'Array programming with NumPy'. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2 (cited on page 130).

[142] David V Lu and William D Smart. 'Towards more efficient navigation for robots and humans'. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE. 2013, pp. 1707–1713 (cited on page 131).

[143] Mason Swofford et al. 'Improving social awareness through dante: Deep affinity network for clustering conversational interactants'. In: *Proceedings of the ACM on Human-Computer Interaction* 4.CSCW1 (2020), pp. 1–23 (cited on page 132).