Investigating Positioning and Gaze Behaviors of Social Robots

People's Preferences, Perceptions and Behaviors

Michiel Joosse

INVESTIGATING POSITIONING AND GAZE BEHAVIORS OF SOCIAL ROBOTS

People's Preferences, Perceptions, and Behaviors

Michiel Joosse

Graduation committee

Chairman and Secretary:

Prof. dr. P.M.G. Apers University of Twente

Supervisor:

Prof. dr. V. Evers University of Twente

Members:

Prof. dr. T. Belpaeme University of Plymouth

Prof. dr. F. Eyssel Bielefeld University

Prof. dr. M.A. Neerincx Delft University of Technology

Prof. dr. ir. P.P.C.C. Verbeek University of Twente

Prof. dr. D.K.J. Heylen University of Twente

Paranymphs:

Bas Joosse

Jered Vroon

CTIT

CTIT Ph.D. Thesis Series ISSN: 1381-3617, No. 17-440

Center for Telematics and Information Technology

P.O. Box 217, 7500 AE Enschede, The Netherlands

SIKS Dissertation Series No. 2017-27

The research reported in this thesis has been carried out under the auspices of SIKS, the Dutch Research School for Information and Knowledge Systems.

The research reported in this thesis received funding from the European Community's 7th Framework Programme under grant agreement FP7-ICT-600877 (SPENCER).

The research reported in this thesis was carried out at the Human Media Interaction group of the University of Twente.

©2017 Michiel Joosse, Enschede, the Netherlands.

Typeset with $\LaTeX$. Printed by Ipskamp Printing Enschede.

ISBN: 978-90-365-4376-7

ISSN: 1381-3617 (CTIT Ph.D. Thesis Series No. 17-440)

Available online at <https://doi.org/10.3990/1.9789036543767>

All rights reserved. No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without prior permission from the copyright owner.

INVESTIGATING POSITIONING AND GAZE
BEHAVIORS OF SOCIAL ROBOTS

PEOPLE'S PREFERENCES, PERCEPTIONS, AND BEHAVIORS

DISSERTATION

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus
Prof. dr. T.T.M. Palstra
on account of the decision of the graduation committee,
to be publicly defended
on Thursday 20 July 2017, at 16:45.

by

Michiel Pieter Joosse
born on 4 October 1989
in Amsterdam, the Netherlands

This dissertation has been approved by:
Supervisor: prof. dr. V. Evers

Summary

As technology advances, application areas for robots are no longer limited to the factories where they perform repetitive tasks behind fences. Robots are envisioned to provide services to us in everyday public spaces - in which they will encounter and interact with people. These types of robots can be considered as social robots, and should interact with people following the behavioral norms people expect of the robot within a specific context.

Compared to interactions people have with social robots in homes and classrooms, interactions with guide robots in public spaces are likely to be of a more incidental nature and shorter duration. This makes it perhaps even more important that users immediately understand how to use the robot. And what intentions and messages the robot communicates. Given that people apply social rules when interpreting the behavior of media (such as computers and robots), we have investigated social norms for social robots; in particular the non-verbal behaviors of (interpersonal) distance and gaze.

In this thesis we have investigated people's preferences for, perceptions of, and behaviors towards social robots through a series (lab and field) studies.

We first conducted a contextual analysis to better understand the future operating environment of the robot. Together with a literature overview of related work in the field of Human-Robot Interaction, the insights we gained the contextual analysis guided the remaining work in this thesis. To address questions about a group's walking speed while carrying luggage we conducted the second study. The results of the first two studies have been used in the design of the third study, in which we investigated people's expectations of a guide robot's behavior in conflicting social situations as it guides small groups.

The next three studies focused on people's perceptions of a robot's approach behavior when approaching a small group of people. The main findings of these three studies are that people adapt the spatial arrangement of their group to the robot. Furthermore, we investigated the perception of approach distances- and directions of Chinese, Argentinian and United States participants. Our results indicate that Chinese participants had a preference for closer approaches over farther approaches. Preferences from people from the United States were the opposite. Finally, we found that projective measures could be used to explore changes in human-robot personal space preferences and behaviors when manipulating an independent variable.

In the seventh and eighth study of this thesis, we investigated people's understanding of a robot's head direction- and turn behavior while the robot guides a small

group of people. We found that a robot's head should face the driving direction when guiding a small group of people, and that occasionally turning towards the people following the robot is interpreted as "caring for them". Finally we conducted a case study in which we investigated under which circumstances the current state-of-the-art in robotic technology could provide added value to airport passengers. This study shows that a social robot at an airport would be a useful service for in particular for people inexperienced with air travel, or to provide information in situations where disruptions of normal operations occur.

The work presented in this thesis provides empirical and methodological contributions to the development of social robots in semi-public spaces. The research in this thesis shows that the behaviors of social robots that provide services such as wayfinding in public spaces (such as airports) should, to an extent, be designed in a human-like way. This thesis offers practical guidelines for designers and developers of social robots. Specifically in terms of which distance robots should keep from groups of people, and in which direction a robot should gaze when guiding (small) groups. The results of this thesis contribute toward the development of social robots in semi-public spaces. The behavior of social robots should suit the context of use. In order to design behaviors for future social robots it is necessary for HRI researchers to study people's interactions with robots in this context of use during the development cycle of a social robot.

Samenvatting

Naarmate technologie zich verder ontwikkeld zal de inzet van robots niet langer beperkt zijn tot fabrieken, waar ze achter hekken repetitieve taken uitvoeren. Van robots wordt verwacht dat ze diensten aan ons aanbieden in de sociale omgevingen - waarmee ze inraking komen met mensen. Dit soort robots kunnen gezien worden als sociale robots, omdat ze met mensen omgaan volgens de sociale normen die mensen van de robot verwachten binnen de specifieke context.

In vergelijking met de interactie die mensen hebben met sociale robot thuis en in klaslokalen, zijn interacties met gidsrobots waarschijnlijk meer incidenteel van aard en korter van duur. Dit maakt het misschien wel belangrijker dat gebruikers meteen begrijpen hoe ze de robot moeten gebruiken. En wat voor intentie en boodschap de robot communiceert. Gegeven dat mensen sociale normen toepassen als ze het gedrag van media (zoals computers en robots) interpreteren hebben wij sozialen normen voor sociale robots onderzocht; in het bijzonder twee non-verbale gedragingen: (interpersoonlijke) afstand en kijkgedrag.

In dit proefschrift hebben wij onderzoek gedaan naar mensen hun voorkeuren voor, percepties van en gedragingen naar sociale robots door middel van negen (lab- en veld) studies.

Als eerste hebben een contextuele analyse uitgevoerd om een beter beeld te krijgen van de toekomstige werkomgeving van de robot. Samen met het literatuuroverzicht van gerelateerd werk in het vakgebied Mens-Robot Interactie hebben de inzichten uit de contextuele analyse richting gegeven aan het verdere werk in dit proefschrift. Om vragen te beantwoorden met betrekking op de snelheid van een groep, terwijl deze bagage meeneemt hebben wij de tweede studie uitgevoerd. De resultaten van de eerste twee studies zijn gebruikt in het ontwerp van de derde studie, waarin we onderzoek hebben gedaan naar de verwachtingen die mensen hebben van een gidsrobot in situaties die een robot tegen kan komen wanneer de robot op een vliegveld begeleidt.

De drie hieropvolgende studies hebben betrekking op de perceptie van naderingsgedrag wanneer een robot op een kleine groep afrijdt. De belangrijkste bevindingen van deze drie studies zijn dat mensen de fysieke ordening van hun groep aanpassen aan de robot. Daarnaast hebben wij onderzoek gedaan naar de perceptie van naderings afstand- en richting van participanten uit China, Argentinië en de Verenigde Staten. Onze resultaten laten zien dat participanten uit China een voorkeur hadden voor een lage naderingsafstand ten opzichte van een grote naderingsafstand. De voorkeur van participanten uit de Verenigde Staten waren tegenovergesteld. Tenslotte

hebben wij onderzoek gedaan in hoeverre het mogelijk is om projectieve meetmethoden te gebruiken om naderingsafstand tussen mensen en robots te onderzoeken wanneer een onafhankelijke variabele wordt gemanipuleerd.

De zevende en achtste studie van dit proefschrift hadden tot doel om een beter beeld te krijgen van een robot's kijkrichting en -gedrag terwijl de robot een kleine groep mensen begeleidt. Onze resultaten laten zien dat het hoofd van een robot in de rijrichting van de robot moet kijken wanneer een kleine groep begeleidt wordt, en dat af en toe naar de groep mensen kijken wordt geïnterpreteerd als inlevend. Tenslotte hebben we een case studie uitgevoerd waarin wij onderzochten onder welke omstandigheden de huidige state-of-the-art in robotica toegevoegde waarde kan bieden aan passagiers op een vliegveld. Deze studie laat zien dat een sociale robot op een vliegveld in het bijzonder van toegevoegde waarde kan zijn voor mensen die onervaren zijn met vliegen, of om informatie te verstrekken in situaties waar onderbrekingen van normale operaties plaatsvinden.

Het werk in dit proefschrift geeft empirische en methodologische bijdragen voor de ontwikkeling van sociale robots in semi-publieke ruimten. Het onderzoek in dit proefschrift laat zien dat het gedrag van sociale robots die diensten in publieke ruimten (zoals vliegvelden) aanbieden, tot op zekere hoogte kan worden ontwerpen op een menselijke manier. Dit proefschrift geeft praktische richtlijnen voor ontwerpers en ontwikkelaars van sociale robots. In het bijzonder met betrekking tot de afstand die robots moeten houden van groepen mensen, en in welke richting een robot moet kijken wanneer deze (kleine) groepen begeleidt. De resultaten van dit proefschrift dragen bij aan de ontwikkeling van sociale robots in semi-publieke ruimten. Het gedrag van een robot moet passen bij de context waarin de robot gebruikt gaat worden. Om gedrag te ontwerpen voor de sociale robots van de toekomst is het noodzakelijk voor onderzoekers om de interactie tussen mensen en robots in deze context te bestuderen tijdens de ontwikkelingscyclus van een sociale robot.

Contents

1	Introduction: The Shape of Things to Come	1
1.1	Introduction	1
1.2	Designing non-verbal behaviors for guide robots	3
1.2.1	Research motivation	3
1.2.2	Research context and scope	4
1.2.3	Research questions	5
1.2.4	Expected contribution	7
1.3	Structure of this thesis	7
2	Theoretical Background & Related Works	11
2.1	Introduction	11
2.2	Robots with social skills	12
2.2.1	Operating context for social robots	12
2.2.2	People's interpretation of robot appearance and behavior	14
2.3	Non-verbal social norms for people and robots	15
2.3.1	Social norms	16
2.3.2	Social norms in Human-Robot Interaction	17
2.3.3	Relevant norms for wayfinding robots	18
2.4	Related work: What we know from previous user studies in HRI on distancing and gaze behavior of robots	22
2.4.1	Approach distance between people and robots	24
2.4.2	Arm- and head gestures for guide robots	37
2.5	Research questions addressed in this thesis	40
2.5.1	Question 1: People's expectations of a robot coping with situations at airports	40
2.5.2	Question 2: Cultural preferences for a robot's spatial behavior (distance and direction) when approaching a small group	41
2.5.3	Question 3: Measuring personal space in through behavioral and projective measures	43
2.5.4	Question 4: Perception of robot head turn behavior while guiding a small group	43
2.5.5	Question 5: Benefits and challenges for a wayfinding robot in "the real world"	44
2.6	Chapter conclusion	44

3	Overview of Studies	45
3.1	Research question I	47
3.2	Research question II	47
3.3	Research question III	48
3.4	Research question IV	48
3.5	Research question V	48
4	Context, Speed & Behavior for a Future Airport Guide Robot	51
4.1	Introduction	51
4.2	Study 1: Contextual Analysis	53
4.2.1	Method	53
4.2.2	Results: observations	55
4.2.3	Discussion	58
4.3	Study 2: Walking Speed Under Time Pressure	58
4.3.1	Theoretical background	59
4.3.2	Method	61
4.3.3	Results	64
4.3.4	Discussion	65
4.4	Study 3: Social Situations at the Airport: How Everyday People Think a Robot Should Act	65
4.4.1	Motivation	66
4.4.2	Hypotheses	66
4.4.3	Method	66
4.4.4	Results	70
4.4.5	Discussion	72
4.5	Chapter conclusion	73
5	Robots Approaching Groups of People: The Role of Spatial Arrangements and Culture	75
5.1	Introduction	75
5.2	Theoretical background: The Spatial Arrangement of Small Groups . .	76
5.3	Study 4: A Telepresence Mystery Investigation	78
5.3.1	Hypotheses	78
5.3.2	Methodology	80
5.3.3	Results	85
5.3.4	Discussion	86
5.4	Study 5: Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People	87
5.4.1	Hypotheses	87
5.4.2	Methodology	88
5.4.3	Results	92
5.4.4	Discussion	95
5.5	Chapter conclusion	96

6	Measuring Patterns in Personal Space with Projective Measures	99
6.1	Introduction	99
6.2	Theoretical Background	100
6.2.1	Techniques to measure personal space	100
6.2.2	A comparison of personal space measurement techniques	101
6.3	The development of two projective personal space measures	104
6.3.1	3D robot	104
6.3.2	2D robot	105
6.4	Study 6: Projective Measures to Assess Patterns in Personal Space	106
6.4.1	Motivation	106
6.4.2	Hypotheses	107
6.4.3	Methodology	107
6.4.4	Results	111
6.4.5	Discussion	113
6.5	Chapter conclusion	114
7	Impact of a Guide Robots' Head Movement on Preferences and Perceptions of Small Groups	117
7.1	Introduction	117
7.2	Study 7: The Effect of Appearance Change on the User Experience	118
7.2.1	Methodology	119
7.2.2	Results	121
7.2.3	Discussion	124
7.3	Study 8: "Are You Still There?": Evaluation of a Guide Robot's Head Turning in the Wild	125
7.3.1	Pre-study: Identifying the effective head turn	125
7.3.2	Main study: Method	128
7.3.3	Results	133
7.3.4	Discussion	135
7.3.5	Conclusion	137
7.4	Chapter conclusion	137
8	Moments of Truth: A Case Study in the Wild	139
8.1	Context of the study	139
8.1.1	The SPENCER platform	140
8.1.2	Operating environment	141
8.2	Methodology	142
8.2.1	Use case and test scenarios	142
8.2.2	Data collection methods & measures	144
8.2.3	Participants	145
8.2.4	Procedure	146
8.2.5	Data analysis	147
8.3	Results	147
8.3.1	Participants' first impressions	148
8.3.2	Positive and negative experiences while being guided	148
8.3.3	Situations where a robot would be useful	150

8.3.4	Customer satisfaction	151
8.4	Discussion	151
8.4.1	Participant bias	151
8.4.2	A niche for wayfinding robots in the airport transfer process . .	152
8.5	Chapter conclusion	153
9	Discussion & Future Work	155
9.1	Discussion of the main findings and contributions	156
9.1.1	Dynamics of people and their expectations of robots at airports	156
9.1.2	Preferences for human-robot approach distance and direction .	157
9.1.3	Design of a robot's head turn behavior	159
9.1.4	Added value of wayfinding robots at airports	159
9.1.5	Preferences for personal space in HRI: behavioral and projective measures	160
9.2	Limitations	161
9.2.1	Sample bias	161
9.2.2	Crowdsourcing	162
9.2.3	Self-reported measures	162
9.3	Implications for theory	163
9.4	Implications for the study of social robots	163
9.5	Future work	164
10	Conclusion	167
10.1	Findings	167
10.2	Implications	168
10.3	The future is bright for social robots and people	169
A	Coding scheme Study 1	171
B	Questionnaire Study 2	173
C	Questionnaire Study 3	177
D	Coding scheme Study 3	183
E	Questionnaire Study 4	185
F	Questionnaire Study 5	191
G	Questionnaire Study 6	199
H	Questionnaire Study 7	205
I	Questionnaire Study 8	207
J	Questionnaires Study 9	211
K	Coding scheme Study 9	215

Bibliography

217

Acknowledgements

This thesis does not mark the end of a four year journey, rather it marks the end of a phase in a journey which started (I think) about six years ago. Doing a PhD, and especially completing a PhD, is a journey filled with numerous obstacles one has to overcome. That being said, it was definitely not always easy, and challenging a lot of times. Over the past four years there are a number of people without whom I would have never completed this project.

First and foremost, my promotor, Vanessa. When we first met, you were an energetic assistant professor, and I was a (naive) first-year bachelor student sitting in the back of a classroom, enrolled in a course called human-computer interaction at the University of Amsterdam. I will never forget the power of paper prototypes. Or the pace with which you went through your slides and (every week) managed to cover the most important contents of 2-3 chapters of the textbook, all within 90 minutes. The consecutive courses you taught; computer mediated communication and research methods, were equally inspiring. When searching for a Bachelor's project I remember a meeting with you where I had some vague idea, something with "the uncanny valley, and perhaps children as well". You mentioned that you were supervising a project, related to the uncanny valley, on which a master student, Aziez, was currently working. That topic turned out to be personal space invasion by robots. I don't recall ever working so hard, or learning so much about doing research, as in those months. You taught me how to do research, analyze data and subsequently write a paper, which was accepted at ICSR'11 in Amsterdam. I was sad to see you go away to Twente, luckily we kept in touch as I still had a paper to present at ICSR, at which I also met Charlotte, Alice, Dennis, and Astrid for the first time. You co-supervised my master thesis from Twente, and offered me a PhD position in Twente. Thank you for giving me this opportunity, for always believing in me, and for keep asking critical questions until the very end.

Manja, thank you for being yourself! Together we embarked on the SPENCER project, and as my daily supervisor you were there every step of the way. In particular I am grateful to you for showing me the value of quantitative research, for patiently reading the 5th, 8th or 11th version of my papers, and of course: teaching me consistency. I will never forget! Your supervision style gave me the freedom to explore different research directions, and the safety of always having you to fall back on when I didn't know what to do. And you taught me what is (perhaps) the most valuable lesson of all: work is important, sure, but there are things which are more important, such as family. Thank you for being there in good times, and in bad times. I couldn't

have done this without you.

I would like to thank my committee members for the reading my of thesis, and your willingness to be part of this committee. I would also like to thank you for your kind words whenever we met, be it Cambridge, Bristol, Bielefeld, Vienna, the DesignLab or just in the hallways of the Zilverling building. Dirk, thank you for making the weekly staff meetings you chaired an enjoyable, at times unpredictable, but welcome start of the week.

Completing a PhD can be lonely at times, luckily for me I have been part of a (growing) group of people at HMI conducting research with social robots. Jorge and Daphne: thank you for making me feel at home at HMI, and all advice you gave over the past four years. Thank you Roelof, Daniel, Cristina, Jaebok, Vicky, Bob, Jeroen, Merel, Pauline, Jamy, and Gwenn for your help. I would like to express my appreciation towards Khiet, Ronald, Dennis, and Betsy for all your advice, even if you were not my supervisors. Over time this group grew to include BMS researchers, in the persons of Maartje and Eduard. In particular I would like to thank Jered: while our backgrounds are different, our research was quite closely related. This made your office my office-of-choice whenever I was looking for honest advice, constructive questions, or just somebody to listen and (when appropriate) tell me that my research indeed made sense. Thank you for being one of my paranympths, it is highly appreciated!

While working on the SPENCER project I have had the pleasure of working with partners from several universities and institutes. The organizational and technical challenges of deploying a robot at an airport were made bearable in large parts due to joint effort of the project partners. In particular I would like to thank Timm, Luigi, Billy, Tomek, Martin, Stefan, Lucas, John, Umer, Harmish, Michaelangelo and Omar for their company during integration weeks in Toulouse, Freiburg and at Schiphol.

At the HMI group I have been lucky to have had supportive office-mates, including Meiru, Bayu and last-but-not-least Roelof. Thank you Roelof, for being an always-supportive sparring partner, and listening to my occasional complaints and frustrations. Thank you Khiet and Randy for accompanying me downstairs every day at around 12:07 for the tomato, mustard or mushroom soup from Sodexo. For four years I have had the pleasure of assisting Vanessa with the organizational aspects of a course I took myself when I was a bachelor student. Though logistically challenging at times, it has been a pleasure to work together with prof. Pamela Hinds, Daisy, Daniela and Ece at Stanford. Also big thank you's to Dennis and Mariët for answering all my questions regarding the administrative and practical sides of teaching. I have had the pleasure of working together with several master students: thank you Robin, Geert, Rutger, Josip, Yannick and Sanne for your work. Miriam, thanks for hosting me in the DesignLab when running my experiments! Andrea, I'd like to thank you in particular for helping me with the 3D-printing and making each visit to the DesignLab worthwhile, thanks to your sense of humor. Thank you Alejandro C., Alejandro M., Angelika, Christian, Danish, Dong, Gijs, Jan, Jelte, Laurens, Mannes, Merijn, Rieks, and Robby for being wonderful colleagues.

I believe that our department would be nowhere without the support of our secretaries. Charlotte, Alice and Wies: the close proximity of my desk to your office made me realize the tremendous workload you are faced with each day. Despite the

many emails, phone calls and double appointments in agendas you always managed to make time to help me with administrative matters. From booking flights, the logic of OFI-numbers, and re-ordering printer cartridges, to advice how on to organize demo's and photo-shoots with robots. I cannot thank you enough. I would like to thank Lynn for reminding me to complete my TAS, the help with EU-bureaucracy and preventing me from introducing new words and expressions to the English language.

There is still a topic that ought to be addressed; and despite my best efforts this topic could not be added to my thesis. In my (humble) opinion, researchers funded by public money have the obligation to disseminate this knowledge back to the public. One way is through open-access publications. Fortunately, social robots in particular have proven to be extremely popular with the public. As a consequence, we have received many opportunities to showcase our research to the public over the years. It has been a tremendous pleasure to contribute to various outreach activities undertaken on behalf of our university. In particular, I would like to mention the opening of the Gallery building, the showcase of FROG at the Hannover Messe, and numerous TV appearance in which I appeared, or, even better, stayed behind-the-scenes to fix the Nao or Giraff (or other robots). I could not have done this by myself, and therefore I would like to thank the Marketing & Communication department for their help. I have always experienced it as a welcome change from my day-to-day desk job. And thank you again, Charlotte, for all your help and support with the demos and talks! Of course, programming demo's for robots always took more time than I had originally planned. This resulted in me especially enjoying the days without any appointments, to "get some work done". Usually these were the Friday afternoons. Thank you, Roelof and Cristina, for reminding me to go home.

At one point in time I decided it would be good do something besides my job (and Netflix'ing). I went to a bunch of people called "Twente Toastmasters". In the past two years with them, I have developed myself. From fearing public speaking, to actually enjoying being "on stage" and telling a story. Learning to listen. And developing skills I did not know I possessed - such as telling a humorous story. In particular I would like to thank Heidy, Huub, Inès, Frederik, Louis, Gonny, Robert, Roberto, Laura, Irfan, Martijn, and Corine.

Last, but certainly not least, I would like to thank my family. I would like to thank my parents for the values with which they raised me. And for your unconditional support, no matter the decisions I made, and if you agreed with them or not. Of course I'd also like to thank my siblings. I think it is special to have a twin brother, and as such I think it is especially fitting that you are my second paranymp, Bas. Thank you! To Bas, Folkert and Wietske: thank you for your visits, your interest in my work, and occasionally worrying that I spend too much time on work.

This is not the end of a journey, rather, it is a milestone. I don't know where my journey will go to, or where it will end. The only thing I know is that I became a better person thanks to all of you - even if I forgot to mention you. You may have noticed that the months preceding the completion of a PhD thesis can be busy. With the successful writing of this thesis, my self-imposed "social exile" is over, and I promise: from now on I will make time.

Michiel Joosse, Enschede, June 2017.

"In the beginning, there is darkness, the emptiness of a matrix waiting for the light. Then, a single photon flares into existence, then another. Soon, thousands more. Optronic pathways connect, subroutines emerge from the chaos, and a holographic consciousness is born. I awaken into this world fully programmed, yet completely innocent, unaware of the hardships I'll endure, or the great potential I will one day fulfill.
Computer, save revisions and open chapter one."

The Doctor:
Star Trek Voyager: Author, Author

1

Introduction: The Shape of Things to Come

"So now you're telling me, umm- Now you're telling me that you're a machine?"
"I'm a woman."
"You're a machine. You're a synthetic woman. A robot."
"I've said it three times now."
"Well, forgive me, I'm having the tiniest bit of trouble believing that, because the last time anybody saw the Cylons, they looked more like walking chrome toasters."
"Those models are still around. They have their uses."

Gaius Baltar, Six.

Battlestar Galactica: The Miniseries, part I

In this chapter I will introduce the main goal of this thesis: studying non-verbal behaviors for guide robots with a wayfinding task in semi-public spaces. Specifically, I will introduce the concept of social robots, the scope of this thesis and the research questions that guide the research reported in this thesis. I will conclude this chapter by providing an overview of the remaining chapters in this thesis.

1.1 Introduction

We use technology in our everyday lives: microwaves, phones, laptops, cars: technology is all around us. The quote above illustrates the current sentiment people are experiencing: robots are seen to leave factory floors and will co-exist with us in our everyday environments. Nowadays, we can buy a vacuum robot which will autonomously drive through our living room, vacuuming the floor, and if we enter the a factory, the factory worker might have been replaced by a robot performing repetitive tasks. Soon, we are told, a hotel receptionist may be replaced by a robot that welcomes and us and checks us into our rooms.

Statistics from the IFR show that the number of service robots sold in 2015 - used for domestic tasks - was 5.4 million units worldwide, an increase of 16% compared to 2014¹. Forecasts predict an additional 42 million units to be sold between 2016 and 2019, ranging from window cleaning robots to robots doing simple tasks in offices. As technology continues to develop, there will come a time when robots evolve from robots with a service task to robots with social capabilities. Interacting besides us. And perhaps with us as well.

This thesis concerns a specific type of robots: guide robots. Historically, robots have been designed to perform tasks which are too dull, dirty and/or dangerous for people to do (Takayama, Ju, and Nass, 2008). Or tasks which can simply be done more efficiently by automating these tasks through the use of robotic technology.

As robots start entering our daily lives, questions about their role in our society arise. Questions about their capabilities and their design, both in terms of appearance and behavior. Robots working closely together with us have until now mostly been known from science fiction, noteworthy examples include Data from *Star Trek*, R2D2 and C3PO from *Star Wars*, and the Cylons from *Battlestar Galactica* (Figure 1.1).

Figure 1.1: Robots in science fiction.

Whether or not one considers the robots in Figure 1.1 heroes or villains, these robots have in common that they communicate seamlessly with people. However when comparing these robots to the current state of the art in robotics there is still a lot of work to be done before robots can seamlessly communicate and interact with people. In this thesis, we will be studying a particular social interaction between people and robots: non-verbal behaviors of robots - in particular interpersonal distance and gaze for the purpose of guiding people.

The development of social robots requires a multidisciplinary team to address the capabilities that such a robot requires. Examples of these capabilities include localization (where am I?), planning (where do I need to go to?) and navigation (how do I get there). Another capability is to interact and communicate with its users. This capability is being studied in the field of Human-Robot Interaction (HRI). HRI is a field of research dedicated to understanding, designing, and evaluating robotic systems for use by or with people (Goodrich and Schultz, 2007) and is heavily rooted in psycho-

¹Source: Executive Summary World Robotics 2016 Service Robots. Retrieved from: http://ifr.org/downloads/press/02_2016/Executive_Summary_Service_Robots_2016.pdf, last accessed 24 April 2017.

logy and computer science (Dautenhahn, 2007). In this thesis we will address the challenge of people's interactions with robots in public space from a HRI perspective.

In the remainder of this chapter we will define the scope of this research, as well as introduce the research questions (Section 1.2). Finally we will present the structure of this thesis (Section 1.3).

1.2 Designing non-verbal behaviors for guide robots

In this section we will introduce the motivation, scope, research questions and expected contribution of this thesis. The goal of this thesis is to investigate "how a robot should behave". Specifically, which non-verbal distancing and gaze behaviors to display within the context of international airports, as we will introduce in Section 1.2.2. Before introducing the research questions, which guide the work reported in this thesis, we will first describe the scope and context of this research.

1.2.1 Research motivation

As robots start entering public spaces, it is important that they adhere to the social norms that exist in the environment where the robot is to provide services (f.e. Bartneck and Forlizzi, 2004; Breazeal and Velasquez, 1999; Michalowski, Šabanović, and Simmons, 2006). In order to sustain long-term interaction between robots and people, social robots should be able to establish, and maintain relationships with people (f.e. Hudlicka, Payr, Ventura, Becker-Asano, Fischer, Leite, and Von, 2009; Klamer, Allouch, and Heylen, 2010). The scope of this thesis is limited to guide robots, as we will further introduce in Section 1.2.2. For guide robots, the duration of the interaction is likely to be shorter and more of an incidental nature as compared to robots which are deployed in domestic environments. When a period of habituation is not possible, it is important that users can easily understand how to use a system and predict how a system will respond to them: the system, in this case a robot, has to be intuitive and predictable.

The way a robot communicates with its users (and bystanders) is not limited to verbal communication. According to Breazeal (2003) social robots have the advantage of sending para-linguistic signals to people which both complement the (verbal) message of the robot, and carry a message of their own. Similarly, Norris (2004, p.2) states that "all movements, all noises and all material objects carry interactional meaning as soon as they are perceived by a person", under which she clusters both verbal and non-verbal behaviors, but also images. For HRI this implies that both the appearance of the robot, the verbal- and nonverbal behavior communicate messages and/or intent towards both the user and bystanders of the robot. In order to design and develop effective robots that assist people in public spaces it is important to optimize the non-verbal behavioral cues a robot communicates. Therefore, we believe it is important to study people's responses to a robot's non-verbal behaviors which communicates meaning or intention.

Having established the importance of studying non-verbal behaviors for guide robots, it is necessary to define which behaviors are to be studied. Different types of nonverbal communication can be distinguished, such as distancing behaviors (e.g.

proxemics and related leaning behaviors), touch (haptics), smell and gaze (Argyle, 1988, p.1). We have not investigated smell and touch. Current technological developments into the artificial generation of smells are not at a level that these can be implemented. Haptics has been investigated in HRI, for example by investigating the most appropriate speed at which a robot should touch a person (Willemse, Huisman, Jung, van Erp, and Heylen, 2016) or the other way around: how people communicate emotions through haptic interaction (Yohanan and MacLean, 2012). Based upon a contextual analysis we conducted (Chapter 4) we observed that the most common non-verbal behaviors people use are distance and gaze. In everyday interactions in public spaces, people maintain an appropriate interpersonal space from other people and use pointing behaviors, for instance through gaze. Therefore, we have chosen to further investigate the non-verbal behaviors of interpersonal distance and gaze in situations which are relevant within the context and scope of the SPENCER project, which we will first introduce.

1.2.2 Research context and scope

The work presented in this thesis has been conducted as part of the EU-FP7 SPENCER project². The goal of the SPENCER project was to develop a mobile robotic platform to guide small groups of people. In particular overseas passengers at Schiphol airport (Triebel, Arras, Alami, Beyer, Breuers, Chatila, Chetouani, Cremers, Evers, Fiore, et al., 2016, p.608). The reason for this target population was as follows: due to an increasing number of passengers, airlines are faced with a higher probability of delays and missed flights. It can be difficult for passengers who are flying for the first time to find their way around, especially at busy airports. At Schiphol airport, passengers have to pass passport control to enter the Schengen area when transferring from an intercontinental to a continental flight. A problem seems to be that passengers do not take into account the time required to pass passport control, and one use case of SPENCER was to guide these passengers to their departure gate. The airport context in which SPENCER will be deployed can be considered as a semi-public space. There are three aspects which make interactions between robots and people in public space special: the interactions are not planned in advance, they take place in an open space, and they could involve co-participants; which we will refer to as bystanders (Weiss, Mirnig, Buchner, Förster, and Tscheligi, 2011).

The SPENCER platform is a 196 cm tall and 250 kg heavy robot (Figure 1.2). SPENCER was designed as an anthropomorphic robot with human-like features, such as a head and eyes. The robot has a somewhat humanlike appearance, though it is clearly not a person. This choice was made because a completely anthropomorphic design could have raised (unfounded) expectations about the robot's cognitive capabilities, which the state-of-the-art in robotics could not produce; thus leading to disappointment of its users (Triebel et al., 2016, p.609). In order to communicate with people through channels other than motion behavior, SPENCER has been equipped with a touchscreen interface, a boarding card reader, a 2-Degrees-of-

²The SPENCER project was funded by the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement no FP7-ICT-600877 (March 2013-2016). The SPENCER consortium consisted of 7 partners. More information can be found at <http://www.spencer.eu>.

Figure 1.2: Artist impression of the SPENCER robot (Figure 1.2a), which was envisioned to provide services to small groups of people in a semi-public space (Figure 1.2b).

Freedom (DoF) head, and a speaker. We will provide a more in-depth description of the SPENCER platform in Chapter 8.

Guide robots, or robots with a guiding task, can be divided into two types: guide robots which provide wayfinding services, and guide robots that focus more on providing context information, for example in a museum. We will refer to the first type of robot as wayfinding robots, and to the second type as explanatory guide robots (Figure 1.3). The research reported in this thesis is limited to wayfinding robots, and as such the studies focus on motion behaviors of the robot. These motion behaviors concern both full-body behaviors of the robot and motion behaviors in terms of motion behaviors of specific component, in this case the head of the robot. We will introduce the research questions associated with these behaviors, and the other research questions, in the next section.

Figure 1.3: The scope of the work reported in this thesis is limited to wayfinding robots.

1.2.3 Research questions

This thesis concerns the study of non-verbal behaviors for a wayfinding robot, to be deployed at an international airport. In the previous section we have introduced the two types of non-verbal behaviors we will investigate in this thesis; interpersonal distance and gaze. In this section we will discuss the research questions which form the basis for the research described in this thesis. We will address the research questions in Chapters 4-8, as further detailed in Section 1.3. Relevant literature and further argumentation for the research questions will be provided in Chapter 2.

Passengers at airports will have needs and goals which could differ from people in other semi-public spaces, such as (f.e.) malls or museums. Therefore, it is important

to understand the context in which the robot, will be used (Maguire, 2001). After identifying situations a robot can expect to encounter when guiding a small group of people, a next step would be to investigate how people expect such a robot to respond. Therefore the first research question is as follows:

RQ-I How do people expect a robot to respond when it encounters social situations while guiding a small group of people? (Chapter 4)

The second research question concerns the investigation of interpersonal distance, in particular into how a robot should approach groups of people. The research presented in Chapter 5 is guided by research question II.

RQ-IIa How should a robot approach a small group of people, in terms of direction and distance? (Chapter 5)

RQ-IIb Do people from different cultures prefer a robot to approach them differently? (Chapter 5)

The third research question is related to the use of behavioral versus projective measures, where research participants have to imagine they are in a certain situation (for instance through dolls). We would like to find out if we can use projective measurement methods - which require less resources - to approximate people's responses compared with a behavioral measure.

RQ-III Can responses to robot proxemics behavior be reliably measured through projective measures? (Chapter 6)

The fourth research question investigates the effect of gaze, or the head turn behavior of a robot on people's perceptions and attitudes towards guide robots. We specifically investigate whether or not a robot should use its head gaze to keep contact with the group being guided, or rather look in the direction of travel.

RQ-IV What head turn behaviors are appropriate for a robot guiding small groups? (Chapter 7)

The research questions II-IV have been addressed using studies in (semi-) controlled environments. In Chapter 3 we will discuss the trade-off between lab and field studies. The dynamics of an airport are difficult to simulate in a lab setting, therefore, in order to gain insights into people's responses to a robot providing wayfinding services in an airport, it is necessary to conduct a field study at an actual airport.

RQ-V How do people respond to and behave when being guided by a wayfinding robot in an actual airport environment? (Chapter 8)

Together these research questions address our overall aim to investigate people's expectations for, perceptions of and behaviors when interacting with a wayfinding robot at an airport. Having provided the research questions to be answered in this thesis, we will continue with the expected contribution in the next section.

1.2.4 Expected contribution

This thesis aims to contribute to an increased understanding of the ways in which a wayfinding robot at an international airport can efficiently use non-verbal distancing- and gaze behaviors. We expect that this thesis contributes to this understanding through three specific contributions.

This thesis contributes to literature on human-robot personal space by providing an overview of existing user studies in the field in which personal space was manipulated or measures. We add to this body of research by presenting two studies in which we investigated human-robot approach distance and directions, specifically in the context of a robot approaching a small group. Finally we present a study in which we investigated if personal space in HRI can be reliably measured through projective measures, beyond the work conducted by Kamide, Mae, Takubo, Ohara, and Arai (2014) and Ohara, Negi, Takubo, Mae, and Arai (2009).

The work reported in this thesis contributes to the development of social robots by providing insights on people's perceptions of different head turn behaviors. Furthermore, we provide insights on people's expectations of a robot's behavior when encountering conflicting social situations while guiding a small group of people.

Finally, this work contributes to development of social robots by presenting a case study in which we report on the opportunities and challenges for a robot deployed in an airport environment. This work complements field studies conducted in other public spaces, such as streets (Weiss, Mirnig, Bruckenberg, Strasser, Tscheligi, Wollherr, Stanczyk, et al., 2015), shopping malls (Kanda, Shiomi, Miyashita, Ishiguro, and Hagita, 2009), stores (Ludewig, Döring, and Exner, 2012), museums (Kuno, Sadazuka, Kawashima, Yamazaki, Yamazaki, and Kuzuoka, 2007b), historic sites (Karremann, Ludden, and Evers, 2015) and transportation facilities (Shiomi, Sakamoto, Kanda, Ishi, Ishiguro, and Hagita, 2008).

1.3 Structure of this thesis

This thesis has been divided into 10 chapters (Figure 1.4). In the current chapter we have presented the motivation, research questions and expected contribution of this thesis.

In Chapter 2 we will present the theoretical background of this thesis. We will introduce and define concepts, such as social robots, and discuss theories related to people's interpretation of a robot's appearance and behavior, such as the media equation and anthropomorphism theories. In the second part of this chapter we will review literature on human-robot distance and gaze, as well as relevant human proxemics literature from among others social psychology. This chapter also contains literature leading to working definitions of groups, culture and social norms.

Chapter 3 will provide an overview of the nine user studies which will be reported in Chapters 4, 5, 6, 7 and 8. In this chapter we will also provide the rationale for the various types of methods applied (experiments, observation studies) and data collected.

In Chapter 4 we will present three studies conducted to better understand the needs of the users of the SPENCER robot: we will present a contextual analysis where

Figure 1.4: Overview of chapters in this thesis.

we observed people’s behaviors at an airport, and from which we defined various guidelines for the SPENCER robot (Joosse, Lohse, and Evers, 2015c). The second study was conducted to investigate how fast a small group of people walks at an airport with luggage, and when under time pressure. The third study was conducted halfway through the project and investigates people’s opinion of appropriate robot behavior in relevant social situations. As cultural differences were expected we involved participants both from China and from the United States (Joosse, Van Waveren, Zaga, and Evers, 2017).

In Chapter 5 we will present two studies related to human-robot distancing. The first study reported is a pilot study conducted in concert with the TERESA project³. We collected data on a telepresence robot’s approaching and positioning behavior in a collaborative task with four people (Vroon, Joosse, Lohse, Kolkmeier, Kim, Truong, Englebienne, Heylen, and Evers, 2015). We collected both objective and subjective data from 14 groups of 4 people. The second study is an online survey in which we investigated cultural differences in human-robot spacing. We collected data from participants in China, the United States and Argentina and found that people from China were more positive to a robot approaching rather closer as compared to participants from the United States and Argentina (Joosse, Poppe, Lohse, and Evers, 2014). Given the nature of this study, it was decided to use crowdsourcing to collect data, therefore asking the participants to project themselves in the situation. In order to study whether or not this method is valid to investigate trends in human-robot personal space a study was conducted, as we report on in Chapter 6.

In chapter 7 we will present two studies related to the anthropomorphic design of the SPENCER robot, specifically participants’ interpretation and assessment of the position of the head while being guided. In both studies people in groups followed a robot. In the pilot study we found participants were more positive about a robot constantly facing forwards compared to constantly facing backwards (Joosse, Knappe,

³The TERESA project was funded by the European Community’s Seventh Framework Program (FP7/2007-2013) under grant agreement no FP7-ICT-611153 (December 2013-2016).

Pingen, Varkevisser, Vukoja, Lohse, and Evers, 2015a). In the second study with the SPENCER platform we manipulated the behavior of the head; we evaluated a behavior where the robot continuously faced towards the participants, versus a behavior where the robot faced forwards and turned around to face the participants at pre-defined intervals. While we did not find significant differences between both conditions in the self-reported questionnaire and distancing data, the open questions suggest that the behavior where the robot turned its head around was appreciated by the participants as it conveyed interest in the group.

In Chapter 8 we will present a qualitative analysis of a case study we conducted during deployment of the SPENCER platform at Schiphol airport (Joosse and Evers, 2017). During this deployment a passengers were guided through the airport, and subsequently participated in interviews. These interviews provided us with insights and directions for future wayfinding robots, and guide robots in general, in public spaces.

Finally, in Chapter 9 we will present a summary and discussion of our main results and the limitations of our studies. We will discuss implications for future research in HRI. Chapter 10 will conclude this thesis with a summary of the contribution of this thesis.

Theoretical Background & Related Works

"We have a saying in our line of work: There's no time like the past."

Captain Braxton.

Star Trek Voyager: Relativity

Part of the work reported in Section 2.5.2.2 has been presented as: Joosse, M.P., Lohse, M. & Evers, V. (2014) Lost in Proxemics: Spatial Behavior for Cross-Cultural HRI. Paper presented at the ACM/IEEE International Conference on Human-Robot Interaction (HRI) Workshop on Culture Aware Robotics, Bielefeld, 2014.

In this chapter we will introduce the background for the research questions which form the basis of the work reported in this thesis. We will first introduce the main concepts and key theories for the purpose of the thesis. Following this introduction, we will present a literature review in which we provide an overview of user studies conducted in HRI which are relevant for the non-verbal behaviors for wayfinding robots in this thesis. Based upon this literature overview we identify the gaps in the literature which we will address in this thesis. To that end we will provide further argumentation for the research questions and working definitions of concepts contained within those questions.

2.1 Introduction

In Chapter 1 we introduced the problem statement and the research questions of this thesis. In this chapter we will discuss the theoretical background and related work on which the research questions are based (Figure 2.1). We will first define social robots beyond the definition already provided in Chapter 1 and in addition we will further discuss the scope of this thesis: social robots in a guiding context (wayfinding robots). We will introduce three theories which show that people evaluate technology in a social way and attribute expectations to the behavior of social robots in a similar way as they would to people (Section 2.2.2). In Section 2.3 we will discuss the concept of social norms which govern people's responses and behaviors in social situations, and which people have consequently been found to apply to social robots as well.

Figure 2.1: Overview of the layout of chapter 2.

In this section we will also argue for the investigation of two non-verbal behaviors (distance and gaze). Subsequently, we will present an overview of user studies in HRI which investigated these two non-verbal behaviors in Section 2.4. Based upon this overview we have identified gaps in the literature we address in this thesis, for which we provide the research questions in Section 2.5.

2.2 Robots with social skills

A robot is *a machine which automatically performs complicated and often repetitive tasks* (Merriam-Webster, 2015b). Typical examples of operating environments of today's robots are factories but also (remotely-controlled robots) far under water and on mars. What these robots have in common is that the tasks they are designed to perform can be considered (too) dull, dirty and dangerous for people to perform. Social robots are designed and expected to conduct tasks beyond dull, dirty and danger (Takayama et al., 2008), and thereby differ from "regular" robots. Bartneck and Forlizzi (2004, p. 592) define a social robot as "an autonomous or semi-autonomous robot that interacts and communicates with humans by following the behavioral norms expected by the people with whom the robot is intended to interact". The definition of Bartneck and Forlizzi (2004) shows that two elements should be taken into account when designing a social robot: adherence to social norms and the target user group. We will define both of these concepts further down in this section. Given that the above definition refers to robot as machines, this hints towards robots that are physically embodied (e.g. have a physical body). In this case we are referring to the definition of embodiment by Lee, Jung, Kim, and Kim (2006a), where a social robot can be disembodied if it does not have a physical body. Within the context of studying the behavior of robots, a disembodied social robots could be considered to be similar to a virtual agent. In the research documented in this thesis we will focus on physically embodied robots, rather than virtual agents. For the purpose of this thesis, we will refer to social robots as follows:

Social robot *"A social robot is a physically embodied robot, which interacts (and communicates) according to the behavioral norms of the people with whom the robot is intended to interact."*

2.2.1 Operating context for social robots

There a number of factors which shape the context in which a social robot operates. Among these factors are the physical environment and the people with whom the robot interacts. Context, task and hardware are three factors which influence each other.

Compared to people, social robots vary in their size, appearance and the modalities they can use to express themselves, and as a consequence this does not automatically make each social robot suited for every task. For example, small, highly-expressive robots such as the Nao robot can be used to explain an exhibit in a museum (Pitsch, Wrede, Seele, and Süssenbach, 2011), and the same robot can interact with children in a classroom environment (Tanaka and Matsuzoe, 2012). However, it would be impractical (for several reasons) to have a Nao give (physically) guided tours. In those cases, a robot such as FROG (Evers, Menezes, Merino, Gavrila, Nabais, Pantic, Alvito, and Karreman, 2014) would be more suited due to the size and mobility of the robot. Therefore, we pose that there is an interplay between a robot's context of use, required capabilities and consequently hardware and appearance, as well as the tasks which a robot can execute.

Social robots have been deployed in different environments. These environments include (semi-)public spaces, such as shopping malls, museums and outdoor public places (e.g. Kanda et al., 2009; Burgard, Cremers, Fox, Hähnel, Lakemeyer, Schulz, Steiner, and Thrun, 1998). Other environments are classrooms (Kanda, Sato, Saiwaki, and Ishiguro, 2007) and domestic environments (Kidd and Breazeal, 2008). Particular tasks for social robots include rehabilitation (social-assistive robots) (Feil-Seifer and Matarić, 2005), robots supporting children with autism, and search-and-rescue tasks (Casper and Murphy, 2003). In Section 1.3.2 we distinguished between two types of guide robots: on the one hand robots which primarily provide wayfinding services (wayfinding robots) and on the other hands guide robots in museums and other historic places which primarily provide information (explanatory guide robots). This thesis focuses on the first type of guide robots.

The scope of this thesis is limited to wayfinding robots in (semi-)public spaces, as a particular type of guide robots. We will briefly provide an overview of guide robots providing services and interacting with people in (semi-)public spaces. Guide robots have been developed over the past decade (e.g. Thrun, Bennewitz, Burgard, Cremers, Dellaert, Fox, Hähnel, Rosenberg, Roy, Schulte, et al., 1999; Tomatis, Philippsen, Jensen, Arras, Terrien, Pignet, and Siegwart, 2002). Where the challenge was first and foremost to develop a robotic platform which could autonomously navigate through a crowded environment, over the past years these robots have evolved, tackling new challenges which has required additional disciplines, such as operating in the outdoor environment, detecting human interest and providing information in ways human guides could not (Evers et al., 2014). Guiding itself can be divided

Figure 2.2: Examples of guide robots: Rhino (1998), Minerva (1999), RoboX (2002), Robotinho (2009) eMuseum (2011) and FROG (2014).

into two different acts: the act of physically guiding people, and guiding in terms of providing information on museum content or directions towards a specific shop. Robots providing physical guiding services have been deployed in various contexts. Examples in museums and expo's (Figure 2.2) include Rhino (Burgard et al., 1998), Minerva (Thrun et al., 1999), RoboX (Tomatis et al., 2002), the PeopleBot-based ATLAS (Shen and Hu, 2006), Robotinho (Faber, Bennewitz, Eppner, Görög, Gonsior, Joho, Schreiber, and Behnke, 2009) and the eMuseum guide robot (Bueno, Viruete, and Montano, 2011). Robots providing guiding services in other contexts than museums have also been investigated: examples include shopping malls (Kanda et al., 2009), train stations (Hayashi, Sakamoto, Kanda, Shiomi, Koizumi, Ishiguro, Ogasawara, and Hagita, 2007) and cultural heritage sites (Evers et al., 2014). This is one of the first studies in which a guide robot is deployed at an airport.

In this section we have so far provided a working definition of a social robot, and we have introduced various environments in which social robots are being deployed. Furthermore we have provided examples of guide robots, as this thesis is about a specific type of guide robot (a wayfinding robot). Especially for human-robot interactions of a short duration, it is important that users understand intuitively how a system works and which response they can expect. Therefore, we will continue this chapter by introducing three theories which help us understand how people interpret the behavior of robots.

2.2.2 People's interpretation of robot appearance and behavior

Having established a working definition of what constitutes a social robot, we will review two theories which help explain how people interpret the appearance and behavior of social robots. These are the Media Equation theory and anthropomorphism theory.

The first theory, or notion, is that people have the tendency to anthropomorphize objects: animals, shapes in clouds and technology, including robots. Anthropomorphism refers to people's tendency to "imbue the imagined or real behavior of non-human agents with humanlike characteristics, motivations, intentions, and emotions" (Epley, Waytz, and Cacioppo, 2007, p.864). According to Epley et al. (2007, p.873) this goes further than the observable actions of an agent, to the level that people make assumptions about the thoughts of agents. Anthropomorphism can be used to rationalize the behavior of an entity (Duffy, 2003).

In HRI, several studies have found that people attribute humanlike characteristics to robots. Examples include experiments which have showed that people attribute personality and attribute characteristics such as intelligence to robots (Lee, Peng, Jin, and Yan, 2006b). People form expectations of a robot's knowledge of famous landmarks based upon the robot's supposed country of origin and name (Lee, Lau, Kiesler, and Chiu, 2005). Eyssel and Kuchenbrandt (2012) used the same social cues to create a feeling of group membership and showed that people respond more favorably to robots belonging to the in-group as opposed to the out-group.

In a similar fashion, the Media Equation holds that people tend to treat computers and other media as if they were either real people or real places (Reeves and Nass, 1996, p.5). More concretely, people unconsciously (or mindlessly) apply social rules

and expectations. For example, people are polite to computers (Nass, Steuer, and Tauber, 1994, experiment I). Experiment participants gave a more polite evaluation of the performance of a computer program when providing the evaluation questionnaire on the same computer as opposed to a separate computer. Another example showing that people treat computers as social actors is that computers can be seen as teammates (Nass, Fogg, and Moon, 1996): participants and computer, part of a blue and green team, conducted a task as teammates or individuals. Participants working with a computer as teammate indicated the information provided was of a higher quality and was more likely to conform to the suggestions of the computer.

So far we have made a case for the proposition that people interpret the behavior of media, including social robots, in a social way. In case of social robots this interpretation is not only caused by the actual behavior but also by its appearance as physical appearance biases interaction (Fong, Nourbakhsh, and Dautenhahn, 2003, p.150). People's initial expectations of the behavior of a robot which looks like a dog will be different from the expectations they may have when they encounter a human-like robot with a head, arms and legs. Next to the initial expectations stirred by the appearance of a robot there is an interplay between a robot's appearance, behavior and the task it conducts. A robot whose human-like appearance conveys expectations which are not effectuated by the behavior could be perceived as dishonest (Walters, Syrdal, Dautenhahn, Te Boekhorst, and Koay, 2008a). In the same way people have a preference for robots whose level of human likeness match the seriousness of the task (Goetz, Kiesler, and Powers, 2003).

Fong et al. (2003) provide a taxonomy of robot morphologies consisting of four major categories (zoomorphic, anthropomorphic, caricatured, and functionally designed robots). Especially when considering guide robots (Figure 2.2) the morphology of these robots can generally be described as more or less anthropomorphic with features such as heads and eyes. The presence of these features may not be without reason. DiSalvo, Gemperle, Forlizzi, and Kiesler (2002) found that presence of specific facial features account for 62% of the variance in the perception of humanness in humanoid robot heads, specifically the nose, eyelids and mouth.

In this section we have introduced theories, or theoretical concepts, which show that the appearance and behavior shown by social robots are interpreted in a social way by people. It therefore seems to be only natural to design the behavior of robots following the rules and conventions people follow in their daily interactions. But herein lies a problem. As Nass and Yen (2010, p. 8) acknowledge there are situations where people do not know the prevailing social norms. Therefore, to develop robots that comply to local social rules it is important to first study the behavior of people in context. We will first take a look at the concept of social norms within the context of those behaviors which could be considered as useful capabilities for a wayfinding robot.

2.3 Non-verbal social norms for people and robots

In Section 2.2 we identified two key concepts of social robots: (1) social robots have to adhere to social norms, and (2) people interpret the behavior of robots in a social

way. This section follows Section 2.2, specifically we will define the concept of “social norm” to be used in this thesis. We will then briefly review the role of social norms in HRI and identify the social norms we will further study in this thesis.

In our daily lives we adhere to all sorts of rules: they govern the way we join a queue when waiting to board a train and how loudly we speak in a library. These rules are (mostly) unwritten, and yet, everybody understands and obeys these rules. Michael Argyle (1973, p.25) states that human social interaction is pre-programmed: “patterns of interaction have been worked out by earlier members of the society, embodied in cultural rules and norms, and are taught to the young. Each of these rules or patterns has been retained because it has been found to be a useful way of handling some situation”.

A general, broad definition is that a norm is *a standard of proper or acceptable behavior* (Merriam-Webster, 2015a). More precisely: a norm can be considered to be a collective understanding of how one ought to behave, an expectation of how a person will behave, and / or expectations of reactions to a particular behavior (Gibbs, 1965, p.589). There are different types of norms, and they are studied by people in different academic disciplines for different purposes. For example, laws can be considered as written norms (Gibbs, 1965; Posner, 1997). In this thesis we are specifically interested in social norms, which we will distinguish from other types of norms, specifically based on whether they are adhered to by an individual or a collective, and whether they are written or unwritten (Figure 2.3).

Norm “a standard of proper or acceptable behavior.”

Figure 2.3: Various types of norms in social psychology.

2.3.1 Social norms

Social norms are unwritten norms, sustained through enforcement within groups (Prentice, 2012), and the prospect of embarrassment and guilt when violated (Elster, 2000, p.100). Social norms are situationally dependent; norms governing appropriate conduct during a soccer game differ somewhat from those which govern a funeral. Aarts and Dijksterhuis (2003) have shown that when confronted with environments associated with specific social norms, people automatically apply rules associated with those environments. For example, when being shown a picture of a library, participants’ voice intensity decreased. Similarly, Axelrod (1986, p.1097)

states that: "A norm exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way". Knowledge of which social norms govern a specific situation can help to more easily predict others' actions (Brauer and Chaurand, 2010).

While different scholars from different fields of research employ different definitions for social norms, one widely-used definition by Cialdini and Trost (1998, p.152) is that social norms are "rules and standards that are understood by members of a group and that guide and/or constrain social behavior without the force of laws". Prentice (2012) provides a similar definition, but focuses somewhat more on the situationally dependent nature of norms. As a working definition, influenced by Cialdini and Trost (1998) and Prentice (2012) we define a social norm follows:

Social norm *"A rule, standard, or convention, understood by members of a group which guides and/or constrains social behavior without the force of written laws in a certain context".*

Thereby a social norm differs from other norms (such as laws) in that the norm is unwritten, applies to a specific social situation, and is culturally dependent. An understanding of what type of behaviors constitutes a violation of the norm can therefore vary from group to group. This distinction is important, because on the one hand robots already have to comply with legal regulations, such as ISO 13482:2014, which specifies safety requirements for personal care robots. On the other hand, there are no laws which state that a robot has to comply to social norms. The inability of a robot to conform to the first could lead to a developer (or researcher) being held accountable, the second can not.

Having provided an understanding of "social norms", we will first present work in HRI which has investigated social norms (Section 2.3.2). Afterwards we will select social norms to be investigated in this thesis, and in the subsequent sections (Section 2.3.3 and 2.4) we will discuss both work from (social) psychology and HRI regarding those particular norms.

2.3.2 Social norms in Human-Robot Interaction

HRI publications frequently refer to the importance for social robots to conform to social norms (e.g. Bartneck and Forlizzi, 2004; Breazeal and Velasquez, 1999; Michalowski et al., 2006). Examples of HRI research investigating social norms include works investigating human-robot approach distances (f.e. Walters, Dautenhahn, Te Boekhorst, Koay, Kaouri, Woods, Nehaniv, Lee, and Werry, 2005a; Mumm and Mutlu, 2011) and the change of social norms when robotic technology is inserted in workplaces (Lee and Takayama, 2011).

Two studies investigated what type of motion behaviors a social robot should display, in order to be perceived as socially acceptable. Qian, Ma, Dai, and Fang (2010) conducted a survey in order to identify important, socially acceptable, safety criteria for mobile robots. Out of eight possible social conventions, four were reported as being the most important by participants of the survey. These conventions were (1) adhere to personal space, (2) maintain visibility when approaching, (3) drive on the correct side of a hallway, and (4) give priority to humans should a robot and human

appear to be on a collision course. Qian et al. (2010) note that the importance of these criteria depends on the scenario and context of the interaction. Pandey and Alami (2009) developed, implemented and tested a framework for a social robot which (autonomously) conformed to different social conventions. These four conventions are: (1) Maintain right-half portion in a narrow passage, (2, 3) pass and overtake a person from his / her left side. (4) Avoid very close sudden appearance from behind a wall. Pandey and Alami (2009) compared this algorithm against a purely reactive algorithm. Three unwanted behaviors of the robot were defined (physical discomfort, unexpected approach and unintuitive behavior). The authors found that when using the algorithm which conformed to the social conventions, an 84.7% reduction in unwanted behaviors was found. These works shows that people have a preference for robots which adhere to social norms.

In the next section we will review which social norms would be relevant for wayfinding robots. Based upon the research of Qian et al. (2010) and the related work in human-robot interaction we have chosen to investigate two aspects of motion and communication behavior for guide robots in public spaces, these being related to human-robot distance and head turn behavior. The reason for these two modalities was further motivated by the affordances of the SPENCER platform and the amount of related work available which provided us with a solid literature foundation.

2.3.3 Relevant norms for wayfinding robots

We have shown that people adhere to social norms in their daily lives, and that adhering to social norms is important for social robots as well in order to ease their acceptance by people. For a wayfinding robot, motion behaviors are essential as the robot guides people physically. As such motion behaviors such as driving speed and the distance a robot keeps from other people would seem to be appropriate behaviors to study. In this thesis we would like to view behavior as more than the physical actions a robot takes in terms of the robot's driving speed and the way it positions itself in an area. In particular, robots which have morphology designed to project a certain image of a robot (e.g. a humanoid or more pet-like robot) can communicate to its users and bystanders, without specifically doing anything. Norris (2004, p.2) poses that "all movements, all noises and all material objects carry interactional meaning as soon as they are perceived by a person". In this she includes both verbal (e.g. speech, sound) and non-verbal behaviors (body movement, facial expressions), and also images. For HRI, this implies that both the morphological design of the robot, the verbal and nonverbal behaviors communicate messages and/or intent towards both the user and bystanders of the robot. We will specifically look at two non-verbal behaviors, which we will introduce in this section.

There are different types of nonverbal communication, such as distancing behaviors (e.g. proxemics and related leaning behaviors), touch (haptics), smell and gaze (Argyle, 1988, p.1). The body of research investigating people's use of non-verbal behaviors started growing in the 1950's, which Patterson (2014, p.172) describes as an "exciting and fluid time for researchers beginning to study nonverbal behavior". Though empirical studies manipulated or measured one non-verbal communication channel at a time, in theories researchers acknowledged the relation between various

channels (Patterson, 2014).

For the user studies in this thesis we have chosen to investigate two non-verbal behaviors to investigate. These behaviors are interpersonal distancing behaviors, and gaze. Distancing behaviors are relevant because (being physically embodied) wayfinding robots are going to have to navigate through space: approaching people and interacting with people. Gaze is interesting because like other robots such as Robovie and Nao, the SPENCER robot has a morphological design which features eyes and a head. This can be used to provide directional cues, or cues showing the intention of (future) movement. For both non-verbal behaviors there is literature available upon which we can build further, both from psychology and human-robot interaction. Other non-verbal behaviors such as touch and smell are also interesting to explore for Human-Robot Interaction, though not as relevant for wayfinding robots. In the remainder of this section we will review literature on the non-verbal behaviors of interpersonal distance (Section 2.3.3.1) and gaze (Section 2.3.3.2). In Section 2.4 we will review the related work concerning these non-verbal behaviors for the topic of wayfinding robots. Having done so we can provide argumentation for the research questions addressed in this thesis as we will do in Section 2.5.

2.3.3.1 Proxemics

The first non-verbal behavior we will review is closely related to proxemics, the way people use space. Proxemics is a term coined by Hall (1966, p.1), which he defines as *the interrelated observations and theories of people's use of space as a specialized elaboration of culture*. We will be investigating interpersonal (approach) distance, and specifically the factors which influence the minimal comfortable distance people keep between themselves and another agent - this being either a person or a robot. The comparison of people's responses to approaches (including comfortable approach distances) by people and robot's has previously been investigated in HRI (e.g. Sardar, Jooisse, Weiss, and Evers, 2012) is not the topic of this thesis.

Research from psychology and related fields (such as anthropology) has investigated people's use of space. In the 1970s and 1980s a lot of research was conducted into this topic, and one reason for this interest identified by Knapp, Hall, and Horgan (2013, p. XVI) was the fear of problems caused by a worldwide population boom. Different related concepts of space have been defined and investigated, and we will first provide an overview of these various concepts.

Sommer (1969, p. 26) used the term personal space to refer to "an area with invisible boundaries surrounding a person's body into which intruders may not come". As this area is situated around us, it is an area we take with us in our daily life. Personal space is therefore dynamic, the opposite of a person's or animal's territory which is usually a static area. Other terms which have been used to describe the (minimum) space people keep around themselves include body buffer zone (Sommer, 1969) and flight distance, introduced by Hediger (1950), an ethologist who studied space requirements for animals in order to improve zoo's.

A few years after Sommer (1959) introduced personal space, Edward T. Hall, an anthropologist, published the book "The Hidden Dimension" (Hall, 1966). This book is heavily cited, also in HRI. Hall defined a series of buffer zones around the body,

like personal space, each suitable for different interactions. These are the *intimate space*, *personal space*, *social space* and *public space*. Based upon observations, Hall provided the approximate sizes of these zones for North American people (Table 2.1). In this thesis we will use the term personal space to refer to *the minimum comfortable approach distance between a person and another agent*.

Zone	Range	Situation
Intimate zone	0.00 m to 0.45 m	Lover or close friend
Personal zone	0.45 m to 1.20 m	Conversation between friends
Social zone	1.20 m to 3.60 m	Conversation between non-friends
Public zone	3.60 m +	Public speech

Table 2.1: Four zones of interpersonal distance as identified by Hall (1966, p. 126)

The size of the personal space bubble has been investigated by many researchers, whether they used Sommer (1959)'s or Hall (1966)'s concept as theoretical basis. Research primarily sought to investigate (a) the size of the personal space bubble, and (b) which factors significantly influenced the bubble, either on the side of the agent (e.g. gaze behavior, gender, age differences) or the context (e.g. crowdedness). In the '70s and '80s literally hundreds of studies and (several) meta-analyses were conducted (e.g. Hayduk, 1978, 1983; Aiello, 1987; Sommer, 2002). These reviews point to the complexity of personal space, for example Hayduk (1983) showed that there are more studies showing partial or conditional support for the influence of gender on personal space than those that clearly support that hypothesis. For this particular factor, Hayduk (1983, p.308) notes that this could be because researchers might expect simple effects from a dichotomous variable. Other factors, which could be influenced by gender, such as status difference and context influenced personal space more often. Factors which were found to consistently influence personal space are age, cultural background (Hayduk, 1983; Aiello, 1987) and participants' height and approach angle (Sommer, 2002). Because of the complexity and individual nature of these factor influencing people's preferences for personal space it should be clear that even for people it is difficult to come to a complete understanding (and codification) of factors which influence people's personal space. In HRI some research has been conducted investigating the influence of some of these factors on personal space, either when a robot approaches or when approaching a robot. We will summarize these in Section 2.4.1.

One of the factors which has been found to influence personal space is culture; the level of analysis used is primarily at the national level. Sussman and Rosenfeld (1982) conducted a study in which 105 students from three different ethnic groups (Japanese, American and Venezuelan) had a (seated) five-minute conversation with a same-sex, same-nationality confederate (Sussman and Rosenfeld, 1982). Either in their native language, or in English. Sussman and Rosenfeld (1982) found that, when speaking English, participants from a low-contact culture (Japan) sat further apart from each other than those from a high-contact culture (Venezuela). Within the ethnic groups male participants sat further apart than female participants. Furthermore, when speaking their native language, high-contact culture participants sat closer together. Other studies looking at cultural differences in personal space include

the work by Little (1968), who used the placement of dolls to infer at which distance people from either the United States, Sweden, Scotland, Italy or Greece would place people in 19 different social situations, and found similar differences between countries. Likewise, Høgh-Olesen (2008) looked at proxemic differences between cultures but also at similarities. Høgh-Olesen (2008) found six cross-cultural proxemics conventions within six cultures (Greenland, Finland, Denmark, Italy, India and Cameroon). For example, people leave more room between themselves and two strangers compared with one stranger, and the personal space is smaller in social spaces (e.g. a cafe) as compared with non-social spaces (e.g. a library). Sorokowska, Sorokowski, Hilpert, Cantarero, Frackowiak, Ahmadi, Alghraibeh, Aryeetey, Bertoni, Bettache, et al. (2017) conducted one of the largest studies investigating cultural differences in personal space so far. The results of this study- a survey conducted in 42 countries indicates that gender and age influence people's personal space preferences, and that temperature in the region could also explain a portion of the variation found. Given the above studies we can assume that there are cultural differences in personal space for people, and a question is if this is also true for HRI (See Study 5 reported in Section 5.4). For now, we will continue our overview of personal space by investigating the consequences of violation one's personal space.

When someone's personal space is invaded, that person can act in several ways in order to respond to this closeness, for example to compensate for the intrusion by stepping back. We will refer to these behaviors as compensatory behaviors. Physical reactions to compensate for the invasion of personal space include flight behaviors which occur within physical proximity: physically stepping away (Hayduk, 1983) or leaning away (Mahoney, 1974). More subtle compensatory behaviors include reducing gaze (eye contact) between two persons (Hale and Burgoon, 1984), and producing more negative facial expressions (Burgoon, 1978). These compensatory behaviors can be considered as objective measurements of people's perception of personal space invasion, whereas subjective measurements include assessment through, for example, questionnaires.

We have provided a high-level overview of the concept of *personal space*; people's psychological and physical distance from other people. People's personal space has been found to be influenced by factors such as gender, cultural background, and the level of gaze between interactors. When personal space is being invaded people compensate for the closeness; they restore the equilibrium, for example, by reducing or increasing gaze and by leaning behaviors, or through physical flight behaviors. In the next section we will take a look at the second non-verbal behavior relevant for guide robots, namely gaze.

2.3.3.2 Gaze & head orientation

The orientation of a person's head and eyes have several functions, some of which we have addressed in the previous section. When navigating space we use our head and eyes to signal (unconsciously) where we are going. For example, we look at other people's eyes to infer their locomotion intent and use to that plan our own movement (Nummenmaa, Hyönä, and Hietanen, 2009). When walking and turning, people use their head to anticipate changes in their current trajectory (Hicheur, Vieilledent, and

Berthoz, 2005); that is to say, they look where they are going to (Hollands, Patla, and Vickers, 2002).

People do not only communicate non-verbally when navigating space, they also use non-verbal communication to regulate psychological distance during conversation - in particular eye contact (or gaze). Argyle and Dean (1965) summarize these functions, which include among others signaling that the channel is open, information-seeking, and establishing and recognizing a social relationship. Gaze can also be interpreted by people (and animals) as a sign of hostility, and as von Grünau and Anston (1995, p.1297) "certainly an indication of potential social interaction".

People infer direction of one's gaze not only by looking at eye direction. Head orientation and full body orientation are also important cues (Emery, 2000). The direction between two (or more) people can be classified in different ways (Emery, 2000):

- Mutual gaze: both people direct their attention towards each other.
- Averted gaze: a person looks at the other person, but his/her attention is directed elsewhere (such as the legs, in Takayama and Pantofaru (2009)).
- Gaze following: a person notices the other person gazes to a different point in space (person, object), and subsequently directs his gaze towards that point in space as well.
- Joint attention: a combination between mutual gaze and gaze following, where both persons direct their attention at a particular other person or object.
- Shared attention: Both persons focus their attention, as in joint attention, and are aware of each others' focus of attention.

Mimicking human gaze behaviors has been shown to elicit social responses from people. A body of research in HRI has investigated how robots with humanoid heads use gaze in interactions with people, for example, to direct or deflect attention, as we will review in Section 2.4.2.

2.4 Related work: What we know from previous user studies in HRI on distancing and gaze behavior of robots

In Sections 2.2 and 2.3 we have provided working definitions of concepts in this thesis, argued for the importance of social robot's to adhere to social norms and we have introduced two non-verbal behaviors that we believe are important to study for an anthropomorphic wayfinding robot in the context of an international airport. These behaviors are interpersonal distance behaviors and gaze behaviors. Before we can define the research questions addressed in this thesis, it is important to have an overview of the related work relevant when studying these two behaviors. Specifically we will review literature related to two topics, which are the factors influencing human-robot personal space, and how a robot's gaze behavior influences the perception of a robot in a guiding task. For this review we defined the following goals:

- Get an overview of all user studies in HRI in which 'personal space' or 'proxemics' was measured or manipulated.

- Get an overview of all variables influencing 'personal space' or 'proxemics' in HRI.
- Get an overview of all user studies in HRI in which the behavior of the head (or gaze) of a guide robot was evaluated or manipulated.

In order to achieve these goals we have conducted a systematic literature review. We constructed two queries (Table 2.2), with which we searched through the proceedings of 13 major HRI and robotics conferences and journals contained primarily within the ACM, IEEE¹, ScienceDirect and Springer databases. We limited our search to include only papers which were published between 2002 and 2016². Table 2.3 shows the initial hits for both queries.

Topic	Query
Distance behavior	robot AND (proxemics OR "personal space" OR "approach distance") AND (experiment OR study) AND (user OR participant OR users OR participants)
Gaze behavior	robot AND (gaze OR "head movement") AND ("guide robot" OR "guiding robot" OR "robot guide") AND (experiment OR study) AND (user OR participant OR users OR participants)

Table 2.2: Two separate search queries were formulated for the literature review.

We systematically went through all 519 hits in a 3-step procedure, consisting of sifting through the papers, reading the papers and categorizing the papers. In the first step we sifted through the results, which consisted of reading the paper's abstract and identifying in which context the keywords were used in the paper. We excluded papers based upon the following nine criteria:

1. *False-positives*: papers in which the keyword is not contained, but somehow turned up in the results list (21).
2. Papers in which the keyword is only contained within the reference list (18).
3. Papers only subjected to a light review, such as HRI Late Breaking Reports (28).
4. Papers in which the keyword is only mentioned in the text (129).
5. Papers in which no user study or experiment with participants (66) or at least 1 physically embodied robot (28) was reported.
6. Papers not fitting the scope of this investigation, including studies with flying robots, socially assistive robots and search-and-rescue robots (23).
7. Papers including only user studies with children as participants (19)³.
8. Papers in which the topic of the user study did not include the area of investigation; distancing behavior (46), guide robots (72) or gaze behavior (13).
9. Papers reporting user studies which were already published before at a different venue (6).

¹IEEE command search in full text & metadata

²Proceedings of the first ICSR conference were unfortunately not available at the time of this review

³We have chosen to exclude user studies with only child participants as interactions between children and robots is distinctly different from interactions between adults and robots as their mental development is still ongoing (Belpaeme, Baxter, De Greeff, Kennedy, Read, Looije, Neerincx, Baroni, and Zelati, 2013). The results of behavioral measures such as personal space cannot be directly compared as the personal space zone of children is under development from the age of 5 until reaching adolescence (Aiello, 1987).

Database	Journal / conference	Distance	Gaze
IEEE	Int. Symposium on Robot and Human Interactive Communication (Ro-Man)	102	36
IEEE	Int. Conf. on Intelligent Robots and Systems (IROS)	33	21
IEEE	Int. Conf. on Robotics and Automation (ICRA)	21	6
ACM	Int. Conf. on Computer Support Collaborative Work (CSCW)	2	3
ACM	Conference on Human Factors in Computing Systems (CHI)	2	23
ACM	Int. Conf. on Human-Robot Interaction (HRI)	29	67
Springer	Int. Conf. on Social Robotics (ICSR)	36	9
Springer	Int. Journal of Social Robotics (IJSR)	49	20
Springer	Autonomous Robots	12	2
ScienceDirect	Int. Journal of Human-Computer Studies (IJHCS)	9	2
ScienceDirect	Computers in Human Behavior	8	2
ScienceDirect	Robotics and Autonomous Systems	14	7
-	Int. Journal of Human-Robot Interaction (IJHRI)	2	2
	Total	319	200

Table 2.3: A systematic literature review conducted in 13 journals and conference proceedings with 2 separate queries resulted in a total of 519 hits.

The sources of the 50 remaining papers are shown in Figure 2.4. When looking at the various publication venues, we note that 80% of these papers originate from conference proceedings; the remaining 20% was published in journals. The most “popular” venues for these type of user studies appear to be Ro-Man (34%, 17 papers), HRI (18%, 9 papers), and ICSR (12%, 6 papers).

While reading the 50 papers we noticed that while all papers were relevant, various categories could be distinguished, which seem to go deeper than the specific query where we found a paper. Therefore we will discuss the user studies reported in these papers based upon the categories. The remainder of this section describes the procedure, in Sections 2.4.1 and 2.4.2 we report the findings.

2.4.1 Approach distance between people and robots

Table 2.4 reports all studies on personal space. For each paper we noted the independent and dependent variables relevant for this overview. Six categories can be distinguished; (in)dependent variables relating to the robot’s *appearance* or *behavior*, the *use of distance* in the experiment, relating to the *participants* recruited, the *situation and/or context* and a *miscellaneous* category. In situations where personal space was a dependent variable, this has been mentioned in Table 2.4 using [Me].

The results will be discussed as follows: we will first provide general descriptive statistics of the results. We will then report on the different (in)dependent variables manipulated on the robot (Section 2.4.1.1) and demographic characteristics of the participants (Section 2.4.1.2) which were found to be of influence on the personal space adhered to. In Section 2.4.1.3 we will report on those papers where exact

Figure 2.4: 50 individual papers were selected out of the original 519 hits, most papers originated from the Ro-Man conference.

distances were reported, and finally in Subsection 2.4.1.4 we will discuss studies studying human-robot proxemics through different robot representations, which is especially relevant for Study 6 in which we investigated if projective measures can be used to measure personal space in human-robot interaction.

A number of papers have reported multiple experiments with different participants. Of the 43 distinct experiments 10 used a within-subjects (23.3%) and 14 a between-subjects (32.6%) design. 11 (25.6%) mixed-design experiments are reported and 4 (9.3%) experiments featured no independent variables. Of the remaining 4 (9.3%) experiments it is unclear how they were designed. Most of the experiments fall in the category “short-term” experiments (76.7%) in which participants in laboratory settings were only exposed to the particular robot and setting once. Three experiments (7.0%) had participants participate in the same experiment over a period of several weeks (Koay, Syrdal, Walters, and Dautenhahn, 2007; Walters, Oskoei, Syrdal, Dautenhahn, et al., 2011b; Syrdal, Dautenhahn, Koay, Walters, and Ho, 2013). In most studies larger robot platforms were used, such as the PeopleBot (used in 25.6% of the reported studies) and Care-o-bot (7.0% of the studies). However, also smaller robots such as Aldebaran’s Nao were used (11.6% of the studies), which have been found to yield different results (Torta, Cuijpers, and Juola, 2013).

Reference & Year & No	N	Platform	Robot appearance \$2.4.1.1	Robot behavior	Distance use \$2.4.1.3	Participant \$ 2.4.1.2	Scenario	Other Part. \$2.4.1.4
Bainbridge et al. (2008)	65	Nico			[Me]			[Ma] Embodiment
Ball et al. (2015)	140	Pioneer		[Ma] Direction	-	[Ma] Group size		[Ma] Group formation
Butler and Agah (2001)	40	Nomad	[Ma] Height	[Ma] Speed	-			
Chidambaram et al. (2012)	32	Wakamaru		[Ma] Sound	[Ma]			
Cruz-Maya and Tapus (2016)	16	Meka M-1	[Ma] Height	[Ma] Smiling	[Ma]	[Ma] Personality		
Eresha et al. (2013) 1	24	Nao			[Me], Placement	[Ma] Culture		
Eresha et al. (2013) 2	24	Nao			[Ma], [Me] Pref.	[Ma] Culture		
Gockley and Matarić (2006)	11	Pioneer			-			
Ham et al. (2012)	85	Nancy	[Ma] Posture		[Me] [HR]	[Ma] Cognitive load		
Hüttenrauch et al. (2006)	22	PeopleBot			[Me] Hall zone		[Ma]	
Joosse et al. (2011)	85	Nomad		[Ma] Speed	-			[Ma] Embodiment
Kamide et al. (2014)	61	Asimo		[Ma] Distance	[Me] [RH] [SD]			[Ma] Embodiment
Kim et al. (2013)	60	Nettoro		[Ma] Sound	[Me] [HR]			[]
Kim and Mutlu (2014) 1	32	Wakamaru		[Ma] Power distance	[Ma]			
Kim and Mutlu (2014) 2	32	Wakamaru		[Ma] Task distance	[Ma]			
Koay et al. (2007) A	12	PeopleBot	[Ma]	[Ma] Direction	[Me] [RH] [SD]		[Ma]	
Koay et al. (2007) B	33	PeopleBot	[Ma]	[Ma] Direction	[Me] [RH] [SD]		[Ma]	
Koay et al. (2014) A	19	Care-o-bot3		[Ma] Direction	[Ma]		[Ma]	[CT] Hand-over
Koay et al. (2014) B	11	Care-o-bot3		[Ma] Direction	[Ma]		[Ma]	[CT] Hand-over
Kosiński et al. (2016)	27	Double		[Ma] Direction	[Me] [SD]			
Kristoffersson et al. (2013b)	21	Giraff			-		[Ma]	[Me] F-formation
Kruse et al. (2014)	17	PR2		[Ma] Algorithm	-			[CT] Path crossing
Lauckner et al. (2014)	35	Bosch Mechanoid		[Ma] Direction; [Ma] Speed	[Me]			

Reference & Year & No	N	Platform	Robot appearance \$2.4.1.1	Robot behavior	Distance use \$2.4.1.3	Participant \$ 2.4.1.2	Scenario	Other Part. \$2.4.1.4
Mead and Mataric (2016)	180	Bandit		[Ma]				
Mumm and Mutlu (2011)	60	Wakamaru		[Ma] Gaze; [Ma] Likeability	[Me]			
Obaid et al. (2016)	22	Nao	[Ma] Posture		[Me][SD]	[Ma] Posture		[Ma] Who approaches
Pacchierotti et al. (2006)	10	PeopleBot			[Ma]			[CT] Passing
Riccio et al. (2016)	75	Turtlebot			[Ma]			[Ma] Embodiment
Rodriguez-Lizundia et al. (2015)	95	Sacarino	[Ma] Posture	[Ma]	[Me]			[Ma] Embodiment
Silvera-Tawil et al. (2015)	4310	Diamandini			[Me] Hall zone	[Ma] Age group, [Ma] Gender		
Syrdal et al. (2006)	42	PeopleBot		[Ma] Direction	-		[Ma]	
Takayama and Pantofaru (2009)	30	PR2		[Ma] Gaze	[Me]	[Ma] Experience	[Ma]	
Torta et al. (2011)	10	Nao		[Ma] Direction	[Me][RH][SD]		[Ma]	
Torta et al. (2013)	9	Nao		[Ma] Direction	[Me][RH][SD]			
Walters et al. (2005a)	28	PeopleBot			[Me][HR][RH][SD] Hall zone			
Walters et al. (2007)	39	PeopleBot		[Ma] Direction	-		[Ma]	
Walters et al. (2008a)	79	Peoplebot	[Ma]		-		[Ma]	
Walters et al. (2008)	58	PeopleBot		[Ma] Sound	[Me][HR][SD]	[Ma] Experience		
Walters et al. (2011b)	7	PeopleBot			[Me][HR][RH][SD]		[Ma]	
Weiss et al. (2010)	52	ACE			[Me] Hall zone		[Ma]	
Woods et al. (2006b)	42	PeopleBot		[Ma] Direction	-		[Ma]	[VHRI]

Table 2.4: Overview of user studies on approaching and distancing :: [Ma] variable manipulated, [Me] variable measured, [CT] context/task of experiment, [HR] human-robot approach, [RH] robot-human approach, [SD] stop-distance task.

2.4.1.1 Manipulations on the “robot”-side

In this section we will discuss the independent variables manipulated on the robot side (columns 4 and 5 in Table 2.4). These factors are approach direction, speed, sound, gaze, and manipulations of a robot’s body.

Approach direction In an experiment by Walters, Dautenhahn, Woods, and Koay (2007) a robot carrying a remote control approached seated participants. A preference was found for a frontal-side approach, with a slight preference from the right side, though the authors indicated this could be due to personal preference. Syrdal, Dautenhahn, Woods, Walters, and Koay (2006); Woods, Walters, Koay, and Dautenhahn (2006b) conducted a study in which participants were approached by a PeopleBot from various directions (front, rear, left-front, right-front, left-rear and right-rear) in four scenarios. In addition, participants shown the same scenario as a video to investigate the feasibility of video-based HRI studies. The scenarios varied and had people either sitting or standing in the middle of a room and behind a table. Participants reported having a preference for frontal-side approaches. Kosiński, Obaid, Woźniak, Fjeld, and Kucharski (2016) conducted an experiment in which participants were approached from five different angles by a Double telepresence robot. During the approaches, participants indicated comfort using hand signals. Results showed that participants preferred approaches from the right side over the left side and front.

As people get used to interacting with robots, preferences can change over time. Koay et al. (2007) conducted a long-term study in which 12 participants were approached by a PeopleBot from various directions in multiple sessions over a 5-week period. Distances were measured using a stop-distance like measure: participants had to press a button on a remote control as soon as they felt uncomfortable. Results indicate that people allowed the robot to approach closer from the frontal sides than the front in week 1, in line with the premise that the personal space zones are not a perfect circle. In week 2 these two distances converged somewhat more, and in week 5 people allowed the robot to approach closer from the front compared with the frontal sides. An experiment with the more anthropomorphic Care-o-bot 3 showed that participants were more comfortable with a robot bringing an object from the side as compared with the front Koay, Syrdal, Ashgari-Oskoei, Walters, and Dautenhahn (2014).

Torta et al. (2011) investigated a robot’s approach distance from multiple directions, though using a Nao robot approaching a seated participant, either from the front (0 degrees) or 35 and 70 degrees from the left or right - thus 5 conditions. Participants were asked to indicate when the robot was at a far distance, an optimal distance and close distance. Torta et al. (2011) found that the farthest directions (70 degrees) were perceived as being the most uncomfortable. No significant differences between the left- and right direction were found. When approaching standing participants similar results were found Torta et al. (2013): approaching from the far left- and right was perceived as less comfortable, and though no significant difference between the left- and right approach direction was found, the data reported showed a preference for approaching from the right-hand side.

Ball et al. (2015) conducted an experiment where people (individuals or dyads)

were approached by a robot from eight different directions while completing a cognitive task. Results for the individual participants were similar to results of previous studies in that participants' rated frontal approaches as more comfortable than rear approaches. Participants seated in a vis-a-vis formation, and on the left side of a side-by-side formation showed preferences just like individual participants. Participants seated in a L-shaped formation, or on the left side of a side-by-side formation, did not show such a preference.

In an experiment relating to approach distance, Lauckner et al. (2014) investigated human-robot approach distances when a robot approached, or when a robot passed by (in front) at different speeds. Lauckner et al. (2014) found that the distance people put between themselves and the robot was smaller when the robot passed by in front as compared to when it approached them.

From these studies we can draw the general conclusion that when approaching participants, with the goal of handing an object people prefer to be approached from a frontal-side direction. The studies reported mainly dealt with domestic tasks in domestic settings.

Approach speed Approach speed has been identified as an important factor in HRI. If a robot approaches too quickly, this action can be considered threatening; if it approaches too slowly, it could signal a lack of social skills. Research in HRI has investigated the appropriate approach speed in various situations. Butler and Agah (2001) manipulated a robot's approach speed and found that participants preferred slower approach speeds (0.25 m/s and 0.4 m/s) over a faster approach speed (1.0 m/s). Lauckner et al. (2014) investigated human-robot approach distances by enabling participants to stop the robot when they felt the robot was too close, when the robot approached at 0.6 and 0.8 m/s. No significant differences between these two speeds were found, and the authors assumed this was due to the small difference between the speeds. Joosse, Sardar, and Evers (2011) investigated differences in reactions to personal space invasion when being approached by a fast (0.4 m/s) or slow (1.0 m/s) robot. Results of this experiment reported in Sardar et al. (2012) indicated that speed did not intensify participants' reaction to the invasion of personal space.

In the previous section we reported on a trial where seated participants were approached by a robot (Walters et al., 2007). In this experiment, the robot approached at a speed of 0.4 m/s, and decelerating to 0.25 m/s when coming close to the participant. Dautenhahn, Walters, Woods, Koay, Nehaniv, Sisbot, Alami, and Siméon (2006) report that 60% of the participants indicated the final speed as generally being correct, while the remaining participants indicated the speed to be too slow.

From these papers we cannot conclude that the approach speed of a robot influences the approach distance - or the distancing at which they would like the robot to stop. The result from Butler and Agah (2001); Dautenhahn et al. (2006) points to a preference for a slower approach speed under 0.6 m/s.

Sound Walters, Syrdal, Koay, Dautenhahn, and Boekhorst (2008b) investigated if the type of voice would lead to different distances, and conducted an experiment in which participants had to approach a stationary robot which either had a male, female,

synthesized or no voice (instructions by experimenter). Walters et al. (2008b) found that people kept significantly more distance from a robot with a synthesized voice. Chidambaram et al. (2012) investigated how a robot could use non-verbal behavior to persuade people, and found that the use of non-verbal cues increased compliance to the robot's suggestions. Furthermore, bodily cues (a combination of distance, gaze and gestures) were more effective than vocal cues. From these two studies it would seem that bodily cues (such as gaze and gestures) might even be more effective in persuading people than vocal cues.

Gaze Takayama and Pantofaru (2009) had participant and robot approach each other, and found that when the robot gazed at participants' legs (averted gaze), women maintain less distance and men more compared to when the robot would look at their face when approaching (mutual gaze). Furthermore, it was found that pet ownership and previous experience with robots also decreases the distance participants put between themselves and the robot. Mumm and Mutlu (2011) also conducted an experiment in which they manipulated gaze, and found that participants compensate for increased robot gaze by increasing distance as well but only if they disliked the robot. In the condition where the robot had a likeable personality, the physical distance did not increase. In general, men positioned themselves further away than women, and they reacted stronger to increased gaze. In a similar fashion, Karreman, Bradford, van Dijk, Lohse, Evers, et al. (2013) manipulated the gaze of an museum guide robot and found that when the robot alternated its gaze between participants and the object participants stood closer to the robot, as compared to when the robot only gazed at the participants.

Robot's body Butler and Agah (2001) investigated the influence of various approach speeds and robot heights, in general participants preferred a smaller, mechanical robot (35 cm) over a variant which was equipped with a humanoid shell (170 cm). Koay et al. (2007) manipulated both robot appearance (humanoid vs. mechanoid) and height (120 vs 140 cm) and found that robot appearance was of more influence to the approach distance than robot height. In a follow-up study, Walters et al. (2008a) evaluated three robot designs (mechanical, basic or humanoid robot) and the influence of participants' personality on participant's rating of the robot. The study showed that participant have a preference for a more humanoid robot.

Obaid, Sandoval, Złotowski, Moltchanova, Basedow, and Bartneck (2016) conducted an experiment in which participants approached a standing or sitting Nao robot and found that participants kept more distance from a standing robot (50.0 cm), than from a sitting robot (35.1 cm). Trials conducted vice-versa, with the robot walking towards the participants (sitting or standing) did not yield significant differences.

Algorithms and models A number of studies reported user studies in which algorithms for robot navigation were evaluated. Oftentimes, these algorithms use the Hall's personal space theory in costmaps. These user studies do not only concern approaching people but also situations in public spaces such as avoiding- and passing people. Kruse, Kirsch, Khambhaita, and Alami (2014) evaluated an algorithm where a robot

avoided collision by slowing down speed, instead of altering its path, this behavior was considered as clearer and less confusing. Pacchierotti, Christensen, Jensfelt, et al. (2006) investigated passing distance during human-robot encounters in the corridor and found people had a preference for an increased lateral distance, which at least kept the robot out of the participants' intimate space zone (Hall, 1966), which would be a distance of 0.4 meter between robot and human. Gockley and Matarić (2006) conducted a study in which a Pioneer robot mimicked participants' behavior in their personal- and social space zones with a random delay, and found that extroverted participants were more open to the robot staying in their personal space as compared to introverted participants.

Miscellaneous Various papers manipulated the type of behavior displayed by the robot.

Mead and Mataric (2016) conducted a study in which the robot's ability to understand instructions (social signals) of 180 participants was manipulated. Before and after the experiment participants approached the robot at a distance which they believed would be ideal for a face-to-face conversation. When asked to approach the robot at the conclusion of the experiment, participants chose locations which were near their perceived location of the robot's maximum understanding. This shows that participants were willing to adapt their behavior - in this case location - in order to help the robot understand them.

Kristoffersson et al. (2013b) conducted a study in which 21 participants operated a telepresence robot in a home-tour, and observed (among others) the spatial F-formations assumed during various interaction episodes. In most of the cases participants assumed the expected F-formation.

2.4.1.2 Manipulations on the "participant"-side

In the previous section we reported on studies where the behavior and appearance of the robot were manipulated.

Culture Eresha, Haring, Endrass, Andre, and Obaid (2013) conducted two studies investigating the influence of participants' cultural background (nationality) on the expectations of human-robot distances, and found that Arab participants positioned themselves significantly closer than German participants in a multiparty conversation. In the second study the authors did not find evidence to conclude whether or not participants had a distinct preference for a robot displaying culturally-normative distancing behavior.

Cognitive load Ham, van Esch, Limpens, de Pee, Cabibihan, and Ge (2012) manipulated approachability of a robot by having participants approach a robot which was positioned either in an open or closed body posture. While body posture did not influence distance, participants under high cognitive load maintained significantly more distance compared with those under low cognitive load.

Prior experience with robots Takayama and Pantofaru (2009) found that people who had prior experience (>1 year) maintained significantly less personal space towards robots as compared to participants who did not have this experience. In a similar way pet ownership influenced personal space. Walters et al. (2008b) conducted an experiment where participants approached a robot with different voice styles, and found that participants who participated in a different experiment with an identical robot⁴ approached the robot significantly closer. Therefore prior experience with robots can influence human-robot approach distances.

Public space Two studies report on deployments of robots in public space. Silvera-Tawil et al. (2015) conducted a study in which a robot was deployed as part of an installation in a museum. People around the robot were classified as interactant, or observer, and child or adult. Results indicated that observers were more likely to position themselves at a “social zone” distance (45 cm - 120 cm) whereas interactants positioned themselves closer, at distances which Hall classifies as personal space zone. Children were observed to position themselves closer to the robot than adults. Rodriguez-Lizundia, Marcos, Zalama, Gómez-García-Bermejo, and Gordaliza (2015) conducted a study in which a robot was deployed in a hotel lobby; also in this study participants who interacted with the robot positioned themselves in the personal space zone of the robot.

2.4.1.3 Personal space as (in)dependent variable

30 studies reported in Table 2.5 manipulated or measured personal space, this could be either through general proxemic zones (Table 2.1) or through the measurement of actual distance, usually by ways of stop-distance measurements. Of these 30 studies 12 reported mean approach distances. Table 2.5 provides descriptives of these studies. Due to the diversity of manipulations, robot platforms and tasks, we will not draw conclusions such as the “mean approach distance”. One general observation that can be made is that - with one or two exceptions - participants in general position themselves in the personal space zone the robot, and expect a robot which approaches them to do the same.

Mumm and Mutlu (2011) investigated how likeability and gaze of a robot influence both psychological and physical closeness to a robot, and found that (1) participants who disliked the robot disclosed less, and (2) compensated for increasing gaze by increasing their physical distance, this in contrast to the likeable robot where participants did not change their physical distance when the robot changed its gaze.

Rather than actual distance, three studies report in which of Hall’s personal space zones participants positioned themselves. Hüttenrauch, Eklundh, Green, Topp, et al. (2006) conducted a study in which 11 participants⁵ gave a home-tour to a People-Bot robot. During various *interaction episodes* both distance and F-formation were analyzed. The findings indicate that during most episodes participants position themselves within Hall’s personal space zone (45 to 120 cm). When looking at the spatial formation, participants predominantly assumed an vis-a-vis formation. Walters et al.

⁴See Walters et al. (2007); Dautenhahn et al. (2006).

⁵22 participants, though data from 11 participants been used in the analysis.

(2005a) conducted an experiment in which participants approached a PeopleBot robot. It was found that 40% approached within the initiate space zone and 60% in either personal- or social space zones. The designs of both studies are too different to draw conclusions regarding this difference. During field trials of the ACE robot, Weiss, Igelsböck, Tscheligi, Bauer, Kühnlenz, Wollherr, and Buss (2010) asked people who interacted with and people who (only) observed the ACE robot, to indicate on a figure how close they would allow ACE to approach them. 37 of the 52 people (71%) indicated they would allow ACE into their intimate space zone (<45 cm), 12 people their personal space zone and the remaining three their social space zone (>120 cm).

The remaining studies used approach distance as independent variable with a robot positioned at a distance based upon Hall's personal space zones. An example of this are the two studies reported in Kim and Mutlu (2014). In these particular studies proxemic distance was used alongside power distance and task distance in a task where participants played a card-matching game with a robot. Both studies showed that the proxemic distance chosen should be congruent with the power- and task distance between robot and person. In a similar fashion Chidambaram et al. (2012), Koay et al. (2014) and Cruz-Maya and Tapus (2016) used distance as independent variable by placing a robot either within or outside participant's personal space boundary. In the study reported by Riccio, Vanzo, Mirabella, Catarci, and Nardi (2016) a robot positioned itself in the intimate-, personal-, or social space zone; participants' scored significantly higher on collaboration attitude towards the robot when the robot positioned itself in the personal space zone. Pacchierotti et al. (2006) investigated the appropriate lateral distance between a robot and person when passing in a hallway, and found that a robot should also stay out of participants' intimate space zone when passing each other in the hallway; the maximum lateral distance of 0.4 meter was deemed most comfortable by participants. Finally, in the second study reported in Eresha et al. (2013) the authors used proxemic distances gathered in the first study to investigate if Arab and German participants' would have a preference for robots showing culturally normative proxemics behaviors. The results as reported indicate that the participants which did notice a difference between the conditions preferred proxemic behaviors which matched their cultural background.

Reference	Setting	N	Platform	Variable	Task		Variable	Level	Distance (SD)
Eresha et al. (2013)	Lab	24	Nao	[B] Participant culture	Placement	n/a	RR	Robot-Robot (Arab)	39.75 (8.34)
								Robot-Robot (German)	42.33 (8.19)
							HRR	Robot-Robot (Arab)	60.17 (18.59)
								Robot-Robot (German)	59.08 (17.68)
								Human-Robot (Arab)	65.79 (19.69)
Ham et al. (2012)	Lab	85	Nancy	[B] Cognitive load [B] Approachability	HR	(stationary robot)		Human-Robot (German)	85.63 (26.30)
							n/a	Overall mean	43
							High cognitive load	Approachable	41.2
								Less approachable	47.2
							Low cognitive load	Approachable	43.3
Kamide et al. (2014)	Lab	61	Asimo	[W] Approach direction [W] Robot embodiment	RH (SD)			Less approachable	44.5
								Frontal	48.3 (31.9)
								Frontal - 30cm	74.7 (26.8)
								Frontal - 60cm	84.9 (34.0)
Kim et al. (2013)	Lab	60	Nettoto	[B] Name calling [B] Speech style	RH (CAD)	WoZ	Name calling	familiar speech style	76.25 (11.53)
								honorific speech style	108.13 (10.89)
							No name calling	familiar style	131.13 (11.53)
								honorific style	95.13 (11.53)
Koay et al. (2007) (L)	Robot house	12	PeopleBot	[B] Robot appearance [W] Scenario [W] Who is in control	HR/RH(SD)	WoZ	Scenario	No-interaction	58.3
								Verbal interaction	56.7
							Mean wk 1, 2, 5	Physical interaction	49.9
							Habituation	Week 1 (overall mean)	56.5
								Week 2 (overall mean)	57
								Week 5 (overall mean)	51.4
Koay et al. (2007) (S)	Robot house	33 ⁶	PeopleBot	[B] Robot appearance [W] Scenario [W] Who is in control	HR/RH(SD)	WoZ	Scenario	No-interaction	60.2
								Verbal interaction	61.2
								Physical interaction	48.9
Lauckner et al. (2014)	Lab	35	Youbot	[W] Approach direction [W] Speed	HR/RH(SD)	(remote control)		Mean frontal	77 (27.5)
								Mean lateral	40 (12.5)
Mead and Mataric (2016)	Lab	180	Bandit		RH(SD)	WoZ		Pre-experiment pref.	112 (51)
								Post-experiment pref.	139 (63)

Reference	Setting	N	Platform	Variable	Task		Variable	Level	Distance (SD)
Rodriguez-Lizundia et al. (2015)	Lab	95	Sacarino	[B] Body posture [B] Who initiates conversation [B] Embodiment	HR	(stationary robot)	Body posture	Awake	56.07
								Asleep	78.5
							Embodiment	Body	69.58
								No body	45.71
							Who initiates conversation	Robot	67.71
								User	56.75
Takayama and Pantofaru (2009)	Lab	30	PR2	[B] Robot experience [W] Gaze	HR/RH		Pet experience	With (> 1 year)	39
								Without	52
Torta et al. (2013)	Lab	18	Nao	[B] Sitting/standing [W] Approach direction [W] Task	RH			Sitting	182
								Standing	173
Walters et al. (2008b)	Lab	58	PeopleBot	[B] Robot voice [B] Previous trial experience	HR	(stationary robot)	Robot voice	Synthesized	80.3
								Male	51.5
								Female	60.3
								No voice	42.4
							Previous experience	With	51.2
								Without experience	73.9

Table 2.5: Overview of user studies reporting minimum or comfortable approach distances (in cm), [B] between-subjects factor, [W] within-subjects factor.

2.4.1.4 Physical (re)presentations of robots

In the previous sections we have provided literature showing that people also keep a similar distance from robots as to people (Walters et al., 2005a) and expect robots to keep a socially normative distance from them as well (Koay et al., 2007; Takayama and Pantofaru, 2009; Mumm and Mutlu, 2011). These studies were all experiments where participants interact with a physically embodied robot. A number of studies have investigated (and compared) methods in which participants did not interact with a physically embodied robot. There are multiple reasons for investigating alternative methods: conducting lab studies are resource-intense (Walters, Woods, Koay, and Dautenhahn, 2005b) and are often biased because that a convenience sample is used (Henrich, Heine, and Norenzayan, 2010).

Three of the studies in this review investigated participants' use of space when interacting with different representations of robots. In a between-subjects experiment were participants received instructions from a robot during a book-moving task, participants could either reach over the robot or walk around it. When interacting with a physical embodied robot participants were more inclined to walk around the robot compared with the virtual robot (Bainbridge, Hart, Kim, Scassellati, et al., 2008). Ohara et al. (2009) conducted two experiments where participants were approached by a HRP-2 robot which was real (a live experiment) or in virtual reality. Participants faced mental stress earlier when being approached by a real robot (at 2.38 m distance) compared with when they were approached by a virtual robot (at 1.74 m). Kamide et al. (2014) conducted a stop-distance experiment with an ASIMO robot, also a real robot and a robot in virtual reality, and found no differences of the type of agent on the amount of personal space.

Different from the other three studies, Woods et al. (2006b) investigated whether people had similar preferences and evaluations of robot behaviors (approach direction). Participants were shown both a video of a robot approaching in a domestic setting, and participated in a trial themselves to compare both methodologies. Results indicated a high level of agreement between the results from both methodologies.

Studies comparing alternative methods to live experiments to investigate robot approach behaviors (distance and direction) include virtual reality and video-based HRI studies. The studies described in this section show promising results (e.g. similar results between the methods). Still, the amount of studies is limited, and this would therefore require additional research.

2.4.1.5 Conclusion

In this section we have provided a review of user studies available using approach distance as (in)dependent variable. From this review we can draw a number of conclusions, not only relevant for this work, but for the broader HRI community as well.

First, we have documented as many as 41 studies in which aspects of the specific navigation behavior of approaching people were studied. The majority of these studies have been conducted in controlled lab settings where individual (or less frequently, multiple) factors were manipulated. The variable distance was either measured or used as an independent variable.

The second conclusion from this review is that most studies investigate the approach distance between robots and individual people. Of interest to the SPENCER project, and for robots in public space in general, would be to investigate how a robot should approach a group. We will provide further argumentation in Section 2.5.

Furthermore, in Table 2.5 we have documented 12 studies in which human-robot approach distances were reported. What is clear from these studies is that people approach robots, and expect robots to approach toward a distance of about 40-60 centimeters, with a small number of exceptions (e.g. Torta et al., 2013; Kim et al., 2013; Mead and Mataric, 2016). This distance is consistent with Hall's personal space zone (Table 2.1), which is according to Hall typically used for conversations with friends. For the broader HRI community this implies that robots should be kept to human social norms, at least regarding interpersonal distancing behavior.

The literature so far suggests that participants show different proxemics behaviors in response to robot behaviors, such as gaze, sound and the combined (multimodal) behaviors of the robot. In particular for the SPENCER project, we deemed it relevant to further investigate the influence of (national) culture on personal space, as this has also found to influence personal space of people rather consistently.

A final conclusion from this overview is that various distinctions can be made between these papers, not only on factors as the type of variable manipulated (which we deemed on the robot, or the participant side), but also a distinction based on the robot platform. For example, the Nao and PeopleBot platforms are distinctly different robot platforms in terms of appearance, size and purpose.

2.4.2 Arm- and head gestures for guide robots

The third goal of this literature review was to get an overview of user studies in HRI in which head turning behavior of guide robots was evaluated or manipulated (Table 2.6). Our review yielded 13 results, and the user studies can be divided by the types of guide robots we defined earlier in Chapter 1; explanatory guide robots (providing information about exhibits) and wayfinding robots (providing directions and/or physically guiding people).

Guide robots have been deployed in various public spaces, such as shopping malls, train stations and museums. Kanda et al. (2009) deployed a guide robot in a shopping mall, which used pointing gestures and natural speech to direct passengers. Compared to information displays, the robot was found to elicit more interest and visits to shops.

One major challenge of deploying robots in public spaces is their ability to understand and interpret verbal communication with people. One of the solutions to this problem is to have the speech recognition done by a human operator (Kanda et al., 2009; Shiomi et al., 2011). Shiomi et al. (2011) evaluated the effectiveness of such an approach by comparing such a semi-autonomous (human operated speech recognition) versus a fully autonomous (automatic speech recognition) system, deployed in a train station. Results indicated that this mixed-method was very effective: in the semi-autonomous condition the system achieved a 69.1% success rate as opposed to a 29.9% success rate in the fully autonomous mode.

Deictic (pointing) gestures have been found as a useful way to emphasize a robot's

Reference & Year & No	Platform	N	Type	Setting
Kanda et al. (2009)	Robovie-M	235	[WG]	Shopping mall
Karremen et al. (2013)	Magabot	57	[EG]	Lab
Knight and Simmons (2013)	Nao	34	[EG]	University building
Koide, Kanda, Sumi, Kogure, and Ishiguro (2004)	Robovie	108	[WG]	Research fair
Kuno, Sekiguchi, Tsubota, Moriyama, Yamazaki, and Yamazaki (2006)	Pioneer II	16	[EG]	Museum
Kuno, Sadazuka, Kawashima, Tsuruta, Yamazaki, and Yamazaki (2007a)	Robovie R-2	46	[EG]	Lab
Kuno et al. (2007b)	Robovie R-2	12	[EG]	Lab
Okuno, Kanda, Imai, Ishiguro, and Hagita (2009)	Robovie	21	[WG]	Lab
Pitsch and Wrede (2014)	Nao	260	[EG]	Museum
Sakamoto, Kanda, Ono, Kamashima, Imai, and Ishiguro (2004)	Robovie	50	[WG]	Lab
Shiomi, Kanda, Ishiguro, and Hagita (2010)	Robovie	269	[WG]	Shopping mall
Shiomi, Sakamoto, Kanda, Ishi, Ishiguro, and Hagita (2011)	Robovie	168	[WG]	Train station
Yamazaki, Yamazaki, Okada, Kuno, Kobayashi, Hoshi, Pitsch, Luff, vom Lehn, and Heath (2009)	Robovie R-2	63	[EG]	Museum

Table 2.6: Overview of user studies on guiding: [EG] = explanatory guide robot; [WG] = wayfinding guide robot

verbal message. Koide et al. (2004) developed a robot which used gaze and pointing and found that these gestures are useful to guide passengers to exhibits. Okuno et al. (2009) developed a model for providing route directions to people, which incorporated both gestures, utterances and timing and found that even though utterances are enough to provide sufficient understanding, adding (deictic) gestures improves understanding. Knight and Simmons (2013) conducted an experiment where a Nao provided information about an university building, and found that deictic gestures combined with utterances are more effective in eliciting the attention of people than utterances alone. The two modalities do need to match as Pitsch and Wrede (2014) found that when gaze- and deictic gesture do not match, for instance because the robot orients to the visitor and points at an exhibition, this causes confusion.

In opposite situations, where a person explains information to a robot, gestures can improve interactions as well (Sakamoto et al., 2004). These gestures are called cooperative body movements, in which a robot repeats gestures made by the person with which it interacts. In situations where the robot used cooperative body movements, participants felt the robot understood them better and they felt more empathy

with the robot.

Next to the use of deictic gestures with arms, research has also investigated the use of head movement and gaze. Kuno et al. (2006) observed museum guides explaining exhibits, and developed a robot which moved its head like a human guide. In a within-subjects experiment Kuno et al. (2006) found participants prefer a robot which moves its head compared with a robot with a head which faces the exhibit continuously. In a follow-up experiment with the Robovie-2 robot Kuno et al. (2007b) found that when a robot turns its head at "transition relevant places" (TRPs), compared with random points, participants nodded more and established more mutual gaze. This suggests that participants understand the explanation better and are more engaged with the robot. Based upon feedback from experts Kuno et al. (2007a) improved the explanation of the exhibit. A person prerecorded the utterances so as to incorporate para-linguistic cues, and participants were asked to remain standing at a predefined position. Furthermore, the experiment was turned into a between-subjects experiment. Results also indicated that the proposed communicative modalities at TRPs lead to significant different non-verbal behaviors by the participants, which can be interpreted as improving understandability. In line with this research Yamazaki et al. (2009) conducted a study with a robot explaining an exhibit in a museum and found that a robot should turn its head at transition relevant places (at the end of a sentence) towards the participants. When asking questions to keep participants involved it should look at the object the robot is asking the question about and turning towards the participants at the end of the question. Also it appears that a robot is an effective tool to keep participants interested as 83% of the visitors stayed during the full 120 seconds explanation.

Head movement can thus be used to keep people engaged and improve understandability when gazing towards participants at transition relevant places - and this research is primarily relevant for what we call tourguide robots where a robot explains exhibits. Head movement can also be used by a mobile robot to attract the attention of people. Shiomi et al. (2010) investigated in which direction a robot should drive in order to attract the attention of bystanders in a shopping mall - people who observe interactions of a robot with other people. It was found that when driving backwards more people were attracted to listen to the robot.

2.4.2.1 Conclusion

The research in this section was concerned with getting an overview of user studies in HRI in which the head movement behavior was manipulated. What we learned from these papers is that gestures - not limited to head gestures - can be used to attract and direct the attention of people. Especially when a robot is equipped with arms pointing gestures appear to be useful in museum settings. Timing of gestures at relevant points within the interaction improves understanding. Specifically relating to the movement behavior of the robot: research indicates that driving backwards is useful for attracting the attention of bystanders. The research however does not give an indication of the head (movement) behavior while guiding a group from A to B, which is one of the research questions which we will address in this thesis.

2.5 Research questions addressed in this thesis

Interactions people have with wayfinding robots that guide them to a specific location are likely to be shorter, and of a more incidental nature as compared to the interactions one might have with for instance domestic robots (such as cleaning robots). It is therefore important that people can easily understand how to use the system. From theory we know that people interpret the actions of media, including robots, in a social way (Section 2.2) people attribute human-like characteristics, such as a personality (Lee et al., 2006b). Furthermore, people apply social rules when interacting with technology. Examples of these rules include being polite to computers (Nass et al., 1994) and maintaining a socially appropriate distance to robots (Walters et al., 2005a). Social robots have to adhere to these social rules (social norms) as well.

In this chapter we have established that people interpret the behavior of a robot in a social way (Section 2.2). Behavior includes all observable signals a robot communicates to its users and bystanders: verbal behavior (e.g. how it speaks), non-verbal behaviors (e.g. how it approaches people) and the cues it transmits through its appearance. People also expect a robot to adhere to human social norms, such as keeping an appropriate distance from people, and driving at the correct side of a hallway (Section 2.3). Because the interactions people have with wayfinding robots are likely to be of a limited duration and focused on visual wayfinding cues, we feel it is important to study a wayfinding robot's non-verbal behavior. Within the constraints of the anthropomorphic design of the SPENCER guide robot we have chosen to investigate the non-verbal behaviors of distance (proxemics) and gaze. In the previous section we have provided a literature overview of user studies in HRI which investigated these two non-verbal behaviors. Having established an understanding of the related work, we can now identify the gaps in the literature which are relevant to investigate within the context of the SPENCER project.

We have identified five research questions. In the remainder of this section we will introduce the research question, and where necessary provide a (brief) theoretical background to define concepts which are important for the specific research question.

2.5.1 Question 1: People's expectations of a robot coping with situations at airports

The first question addressed in this thesis concerns people's expectations of a robot's behavior within the context of an international airport (Chapter 4). From literature on social norms we know that these norms are situation-dependent (Aarts and Dijksterhuis, 2003). We define context as *"the physical environment in which a system (the robot) operates and the background, needs and goals of the potential users"*. The first goal of this chapter is to arrive at a better understanding of what type of situations a wayfinding robot can encounter while guiding a small group of people. Situation, in this case, can also be thought of as specific characteristics which influence the way the system can behave. We will therefore look both at the actions of people and the (physical) characteristics of the operating environment.

From literature we know that robots should adhere to human social norms (e.g. Bartneck and Forlizzi, 2004; Qian et al., 2010), and that the incorporation of typical human social conventions in path-planning algorithms reduces unwanted robot beha-

vivors (Pandey and Alami, 2009). Still, a problem is that we do not know how people expect a robot to respond to situations in terms of what behavior it should display. We will therefore investigate how people expect such a robot to respond in situations which can arise while the robot guides a small group of people to a location in the airport.

RQ-I How do people expect a robot to respond when it encounters social situations while guiding a small group of people?

2.5.2 Question 2: Cultural preferences for a robot's spatial behavior (distance and direction) when approaching a small group

The first part of the review of user studies in HRI was conducted to get an overview of all user studies in which personal space, or proxemics, was measured or manipulated (Section 2.4.1). In addition, we wanted to get an overview of the variables which influence human-robot personal space. This overview yielded two important gaps in the literature, namely the investigation of group size and the effect of cultural background. We have not yet defined our working definition of groups and culture yet, which we will do next before providing argumentation for two research questions, RQ-IIa and RQ-IIb.

2.5.2.1 Groups

Merriam-Webster (2017b) defines a group as “a number of individuals assembled together or having some unifying relationship” or “two or more figures forming a complete unit in a composition”. Research on small groups from a sociological perspective focuses on power and status structures within groups, whereas research from a psychological perspective focuses more on performance processes (Kelly, McCarty, and Iannone, 2013, p.414). Similarly, in the literature, there are different understandings of when a group of people constitutes a group, or more precisely a small group. Though there are phenomena which cannot occur when a group only consists of two persons (a dyadic group), we will follow Williams’ definition in that we consider a group as being “two or more people” (Moreland, 2010; Williams, 2010). We do add to this definition that they share a goal.

Group *“Two or more people who share a goal.”*

For the purpose of this thesis, we do not consider a robot and one person as a group; though a robot can be part of a group as long as there are two or more persons.

The first gap in HRI research on personal space relates to group size: in crowds in public space up to 70% of the people are part of small groups (Moussaïd, Perozo, Garnier, Helbing, and Theraulaz, 2010). However, the research reviewed in Section 2.4.1 is primarily concerned with robots approaching individual people. It is necessary to investigate how groups of people interact with robots, of particular relevance would be to investigate how robots should approach small groups of people. We define RQ-IIa as follows:

RQ-IIa How should a robot approach a small group of people, in terms of direction and distance?

2.5.2.2 Culture

People's interactions are not governed by one specific set of social norms (Section 2.3) as norms can differ across cultures. Culture is an ambiguous concept. Originally, the term culture stems from the Latin word *colere*, and it "[...] usually referred to something that is derived from, or created by the intervention of humans - culture is cultivated" (Hamamura, 2012). Triandis (1972) divided culture into a subjective and material culture. Material culture consists of elements, for instance food, houses and tools. Subjective culture, on the other hand refers to the characteristic way in which a specific group perceives its environment (Triandis, 1972). When referring to culture, we refer to subjective culture.

Cultures can be investigated at different levels of analysis - in this thesis we analyze culture at the national level. Usually, when scholars are looking at a culture - and the differences between cultures, the level of analysis is the nation, or subcultures within a nation. Karahanna, Evaristo, and Srite (2006) defined different levels of cultures, these being supranational, national, and levels within a nation, such as the professional, organizational and the group level. Cultures can form at different places in society, and people can belong to different cultural groups, each group in turn having its own specific norms, conventions and irregularities. Researchers have developed a number of measures to describe differences between cultures, examples include the work by Hofstede and Hofstede (2001), who described (national) cultures on five different dimensions (power distance, individualism, uncertainty avoidance, masculinity and long term orientation). Other studies investigating differences between national cultures include Pelto (1968) and Levine and Norenzayan (1999) who compared the pace of life in 31 cities around the world. Gelfand, Raver, Nishii, Leslie, Lun, Lim, Duan, Almaliach, Ang, Arnadottir, et al. (2011) investigated the appropriateness of behaviors everyday situations across 33 countries and Brauer and Chaurand (2010) compared how people in eight countries rated 46 uncivil behaviors. For the purpose of this thesis, we define culture as follows:

Culture *"Subjective culture measured at the national level."*

Individual national cultures can be clustered. By clustering cultures it is possible to summarize both similarities and differences. For cross-cultural research this can be useful because it ensures that the sample includes enough cultural variability. To cluster nations, three major forces have been used historically to cluster countries, these being geographic (1) geographic proximity, (2) mass migration & ethnic social capital, and (3) religious and linguistic commonality (Gupta, Hanges, and Dorfman, 2002). One example of such a clustering is the work by Gupta et al. (2002) who clustered 61 nations into 10 clusters of cultures (Figure 2.5).

The second gap in the HRI research on personal space concerns the cultural background of the participants. Previous research found that there are differences in the personal space preferences of people (see Section 2.4.1 for an overview). Beyond the

Figure 2.5: Ten clusters of culture, figure based upon Gupta et al. (2002). Legend: ■ Anglo cultural cluster, ■ Latin European cluster, ■ Nordic European cluster, ■ Germanic European cluster, ■ Eastern European cluster, ■ Latin American cluster, ■ Sub-Sahara African cluster, ■ Arab cultural cluster, ■ Southern Asian cluster, ■ Confucian Asian cluster.

research of Eresha et al. (2013) this factor has not been investigated in HRI. Considering the culturally diverse population one can expect to encounter at an international airport this warrants further investigation. We define RQ-IIb as follows:

RQ-IIb Do people from different cultures prefer a robot to approach them differently?

2.5.3 Question 3: Measuring personal space in through behavioral and projective measures

The third research question is related to the use of non-real (projective) measures where research participants have to imagine they are in a certain situation. In personal space research various measures other than live measurement have been used, both in interactions between people (e.g. Sussman and Rosenfeld, 1982; Høgh-Olesen, 2008; Sorokowska et al., 2017) and in HRI (Kamide et al., 2014; Ohara et al., 2009). We would like to find out if we can use this method - which requires less resources - to investigate trends in personal space, or conduct prototype research in early phases of development. RQ-III, presented in Chapter 6 will address this question.

RQ-III Can responses to robot proxemics behavior be reliably measured through projective measures?

2.5.4 Question 4: Perception of robot head turn behavior while guiding a small group

The second part of the review of user studies in HRI focused on user studies in which the turning behavior of a robot's head was evaluated or manipulated. We found that especially the head of anthropomorphic robots can be a powerful communication modality. Previous studies found that head and eye gaze can be used to engage

museum visitors (e.g. Kuno et al., 2006; Yamazaki et al., 2009). Research by Shiomi et al. (2010) investigated how the driving direction of a robot can be used to attract bystanders into listening to a robot. What we do not know yet is how the driving direction, and in particular the direction of the head of a robot is interpreted by people who are following a robot. To further investigate this, we have conducted two studies which are reported in Chapter 7. These studies contribute to RQ-IV.

RQ-IV What head turn behaviors are appropriate for a robot guiding small groups?

2.5.5 Question 5: Benefits and challenges for a wayfinding robot in “the real world”

The fifth research question concerns the deployment and evaluation of the SPENCER robot at an international airport. Research questions 2-4 have been addressed in (semi-)controlled environments; which have as disadvantage that it is difficult to understand under which circumstances people would be likely to use the robot. The dynamics of an airport, or other public spaces, are difficult to simulate in a lab setting. In order to gain insights into the added value of a robot for wayfinding services, it is necessary to conduct a field study at an airport.

RQ-V How do people respond to and behave when being guided by a wayfinding robot in an actual airport environment?

2.6 Chapter conclusion

Social robots are designed to interact with people. In our everyday lives, our behavior is governed by social norms, and research in HRI has shown that social robots should also adhere to these norms to be accepted, although perhaps not in exactly the same way.

We know from previous research that people expect robots to respect their personal space, and that cultural differences in HRI exist as much as they do in our daily lives. We do not know yet, however, if there are cultural differences for people’s personal space preferences when a robot approaches a small group. We will address this in Chapter 5. A second norm, or convention is that people anthropomorphize robots, and for humanoid robots which are equipped with a humanoid head research has shown that people treat it as a head, with eyes. Gaze has been shown to be an effective modality when for example directing people’s attention. What is yet unknown is how people interpret the gaze behavior of a robot when being guided to a specific location. In Chapter 7 we will present two user studies investigating appropriate head turn behaviors of a robot while *guiding a small group of people*.

Before presenting the user studies in Chapters 4-8 we will present an overview of studies and methodologies in Chapter 3.

3

Overview of Studies

“Wie o wie weet toch het antwoord,
of sterker nog: wat is... de vraag?”¹

Youp van 't Hek.

Wat is de Vraag? Oudejaarsconference 2014

The work reported this chapter has partly been presented in: Joosse, M. P., Lohse, M., and Evers, V. (2015). Crowdsourcing Culture in HRI: Lessons Learned From Quantitative and Qualitative Data Collections. Paper presented at the 3rd International Workshop on Culture Aware Robotics.

In the previous chapter we provided the theoretical background and related works for this thesis, which led to the research questions of this thesis. We have conducted nine user studies which are documented in chapters 4, 5, 6, 7 and 8 (Figure 3.1 and Table 3.1). In order to investigate people's perception of non-verbal behaviors for wayfinding robots in the airport, one approach could be to take a robot to an airport, manipulate behaviors, and measure the responses of people. There are a couple of practical and scientific challenges which we would not be able to address using this method, and we therefore had to use a variety of research methods. In this chapter we will provide the rationale for the specific types of studies conducted, as well as a general overview of the studies.

Figure 3.1: In this thesis we will present 9 studies contributing to research questions I - V.

¹The “oudejaarsconference” is a form of a Dutch cabaret show broadcasted on national television on New Years Eve, in which one or more comedians reflect on the past year. The literal translation of this phrase is “Who knows the answer, or better: what is the question?”.

Type / setting			Data	Context	N	Unit of analysis	Design	Factors	Independent variable*
Ch.4	Study 1	Field	Qualitative	Airport	-	-	-	-	-
	Study 2	Lab - experiment	Quantitative	Lab	39	Single participants <i>in groups of 3</i>	Mixed	1 between-subjects 1 within-subjects	[W] Luggage - 2 [C][B] Time pressure - 3
	Study 3	Survey (crowdsourcing)	Mixed- method	Online	118	Single participants	Mixed	2 between-subjects 4 within-subjects	[P][B] Culture (nationality) - 2 [C][W] Situation - 10 [C][B] Time pressure - 2 [C][B] Group size - 2 [C][B] Majority/minority - 2 [C][W] Involvement - 2
Ch.5	Study 4	Lab - experiment	Quantitative	Lab	54	Single participants <i>in groups of 4</i>	Mixed	1 between-subjects 1 within-subjects	[R][B] Group formation - 2 [R][W] Approach direction - 8
	Study 5	Survey (crowdsourcing)	Quantitative	Online	181	Single participants	Mixed	1 between-subjects 2 within-subjects	[P][B] Culture (nationality) - 3 [R][W] Approach direction - 3 [R][W] Approach distance - 6
Ch.6	Study 6	Lab - experiment	Quantitative	Lab	68	Single participants	Mixed	1 between-subjects 1 within-subjects	[B] Group membership - 2 [W] Measurement method - 3
Ch.7	Study 7	Lab - experiment	Qualitative	Lab	25	Single participants <i>in groups of 2-3</i>	Between- subjects	1 between-subjects	[R][B] Head behavior - 3
	Study 8 pre-study	Survey (crowdsourcing)	Quantitative	Online	29	Single participants	Within- subjects	1 within-subjects	[R][W] head behavior - 5 [R][W] range of turn angle - 4
	Study 8 main	Lab - experiment	Mixed- model	Lab	68	Single participants <i>in groups of 2-3</i>	Between- subjects	1 between-subjects	[R][B] head behavior - 2
Ch.8	Study 9	Field (case study)	Mixed- model	Airport	16	Single participants <i>in groups of 1-3</i>	-	-	-

Table 3.1: Overview of studies. Manipulations between brackets: [R] on robot; [P] participant; [C] context; [B] between-subjects; [W] within-subjects.

3.1 Research question I

The first question addressed in this thesis concerns people's expectations of a robot's behavior within the context of an international airport (Chapter 4). In order to answer this question we conducted three studies. In the first study we performed a contextual analysis of the future operation context to elicit relevant social situations which the robot would encounter. In this study we collected and analyzed qualitative data. Qualitative data includes open-ended information, in contrast to quantitative data which includes closed-ended information (Creswell and Plano Clark, 2007, p. 6).

One of the questions we could not answer with the contextual analysis concerned the walking speed of a group of people, and in particular in what way time pressure and carried luggage influenced this. In order to investigate this we conducted a controlled lab study, reported as Study 2, to investigate this.

The third study was a more exploratory study in which we investigate people's expected response by a robot when encountering social situations (such as people entering a store, or taking a selfie). We collected (free) text data from participants, thus qualitative data, which we subsequently coded and quantified.

3.2 Research question II

To address research question IIa we collected a dataset containing among others spatial-temporal data of small groups of participants being approached by a robot from different angles. The setup required extensive hardware for data collection and called for a level of control that could not be achieved in (semi-)public spaces. As such we conducted a lab study which allowed for a high level of control and replicability (Siek, Hayes, Newman, and Tang, 2014, p.121).

Research question IIb set out to investigate (potential) cultural differences in human-robot proxemics. We chose specifically not to conduct a lab study in which we would have invited foreign participants to participate. When investigating cultural differences with real robots, one either has to take the robot to the participants (such as Wang, Rau, Evers, Robinson, and Hinds (2010)) or bring the participants to the robot, for example by recruiting participants with different cultural backgrounds. One disadvantage of the latter method would have been that the ecological validity could have been compromised as at least a part of the sample was not observed in their natural environment (Baldassare and Feller, 1975). This is a disadvantage as a cultural identity not only consists of one's social norms and ideas (subjective culture) but also the way one lives, including a culture's architecture, dress and food (material culture) (Triandis, 1972). A solution to this problem could have been to "take a robot" to participants without using an actually physical embodied robot. An example of such a solution is the video-based HRI method employed by Walters, Lohse, Hanheide, Wrede, Syrdal, Koay, Green, Hüttenrauch, Dautenhahn, Sagerer, et al. (2011a) and Woods, Walters, Koay, and Dautenhahn (2006a) where participants watched videos of a robot interacting with an actor and subsequently rated extrovert and introvert robot behaviors. We chose for the latter option by presenting participants a questionnaire in which they were presented with various situations containing situations in which

a robot had approached a group of people. We recruited participants through crowdsourcing. Crowdsourcing presents an opportunity to run online studies with people from all over the world. This is an advantage as there is a tendency in Western countries to recruit participants who can be described as belonging to Western, Educated, Industrialized, Rich, and Democratic (WEIRD) countries (Henrich et al., 2010).

3.3 Research question III

Research question III addressed the use of non-real (projective) measures where research participants have to imagine they are in a certain situation. For this experiment we borrowed ideas from studies in the '70s and '80s in which this exact phenomenon was investigated - with the exception of the absence of a robot. The study contributing to this research was conducted as a lab study as we wanted to exercise a high level of control over the situation.

3.4 Research question IV

With research question IV we addressed the people's interpretation of the direction the head of a robot gazes towards while the robot guides that group of people. We addressed this question through studies 7 and 8. Both studies were conducted in semi-controlled environments, that is both locations were university buildings also used for other purposes (such as study). The level of control we could exercise over the situations was therefore not as high as in lab studies but an advantage of this method was that the ecological validity of the studies was higher.

3.5 Research question V

The final research question concerned research the deployment and evaluation of the SPENCER robot at an international airport. In order to so, we conducted a field study at Schiphol Airport. This study had high ecological validity, and a low level of control. This means that there were a number of factors which were different over the trials conducted, in particular the crowdedness in the terminal. Field studies have advantages over other research methods - in particular lab studies - for example that they are conducted within the intended context of the tool. Through field deployments it is possible to understand in which circumstances technology will be adopted, and which other technologies in the environment influence the acceptance of the particular technology. Specifically, field studies are preferred over lab studies in HCI when a phenomenon does not fit in the lab setting and/or when the interaction cannot be staged in a convincing manner in the laboratory (Oulasvirta, 2009). Given the above-mentioned arguments, a field study was considered the most suited method to address this research question.

In the next five chapters we will present the methodology and results of the studies. In particular, Chapter 4 contributes to research question 1, Chapter 5 contributes to research question 2, Chapter 6 contributes to research question 3 and Chapter 7 contributes to research question 4. Finally Chapter 8 contributes to research question

5. In Chapter 9 we will summarize the main results and contributions from these studies.

4

Context, Speed & Behavior for a Future Airport Guide Robot

"All the things you probably hate about traveling -the recycled air, the artificial lighting, the digital juice dispensers, the cheap sushi- are warm reminders that I'm home."

Ryan Bingham

Up in the Air (2009)

The work reported in Section 4.2 has partly been presented as: Joosse, M.P., Lohse, M. & Evers, V. (2014) How a Guide Robot Should Behave at an Airport - Insights Based on Observing Passengers. Technical Report TR-CTIT-15-01.

The work reported in Section 4.4 has partly been presented as: Joosse, M.P., Van Waveren, S., Zaga, C. & Evers, V. (2017) Groups in Conflict at the Airport: How Everyday People Think a Robot Should Act. Poster presented at the CSCW'17 Workshop on Robots in Groups and Teams.

This chapter concerns the study of the context in which the SPENCER robot is envisioned to operate; an international airport.

We will present three studies with the goal of getting a deeper understanding of the role of the SPENCER robot in its future operating environment. In the first study we conducted a contextual analysis to investigate people's behavior when navigating through an airport. The second analysis follows up on this contextual analysis and was conducted to gain insights in people's walking speeds in groups, with and without carrying hand luggage. In the third study we investigated through an on-line survey how people believe a robot should respond when encountering conflicting social situations at the airport.

4.1 Introduction

Context plays an important role in our daily lives. What is appropriate in the context of a playground may not be appropriate in the context of a library (Aarts and

Dijksterhuis, 2003): running around and talking loudly are two example behaviors which may not be tolerated in libraries¹. Merriam-Webster (2017a) provides the following definition of context: “the interrelated conditions in which something exists or occurs”. In contrast, context has proven to be a difficult concept to define in academia (Bazire and Brézillon, 2005; Dey and Abowd, 2000). A context can not be defined only by its physical location, rather a context should be seen as a combination of what, when and where a person is doing a task (Erickson and Schultz, 1997, p. 22). On the other hand, in terms of context-aware (software) applications, Dey and Abowd (2000) note that there are different types of context, with the four most important types being location, identity, activity and time. For this thesis the definition of context by Bazire and Brézillon (2005) seems most applicable as it focuses on the constraints imposed upon a person or system which influence its behavior. We therefore define context as follows:

Context *“Context acts as a set of constraints that influence the behavior of a system (a user or a computer) embedded in a given task”.*

In this research we will consider the context in which a person interacts with a robot as being comprised of the physical environment in which the robot operates and the (demographic) background, needs and goals of the potential users. Within the context of an airport we expect that a wayfinding robot requires different capabilities than wayfinding or tourguide robots in other environments as museums (Pitsch, Kuzuoka, Suzuki, Süßenbach, Luff, and Heath, 2009), shopping malls (Okuno et al., 2009; Shiomi, Kanda, Glas, Satake, Ishiguro, and Hagita, 2009), historic sites (Karremann et al., 2015) and train stations (Shiomi et al., 2008). Therefore, the goals of this chapter are as follows: first we would like to better understand what type of situations a wayfinding robot can encounter while guiding a small group of people in an airport. The second goal is to understand how a robot should respond when encountering social situations while guiding a small group.

The three studies reported in this chapter (Figure 4.1) are focused on understanding the current situation at an airport (Section 4.2), identifying a small group’s walking speed while carrying luggage (Section 4.3), and identifying how a future robot should behave when encountering conflicting social situations while guiding a group of passengers (Section 4.4).

Figure 4.1: Overview of the layout of chapter 4.

¹When I was three or four years old I violated this norm in the library too, once or twice.

4.2 Study 1: Contextual Analysis

We use the term contextual analysis to describe a method to discover how people behave in a given environment, which is in the case of the SPENCER robot an international airport - Schiphol Airport. During the contextual analysis we have observed relevant situations for the purpose of deploying a wayfinding robot. The goal of the contextual analysis was to understand how people behave - towards the environment and other people - when transferring at an international airport. In order to attain this goal we specified a set of contexts in which people navigate at the air-side (after passport control) of the airport. These include among others the disembarkation area and lounge.

In the method section we will describe the research design in more detail, and provide information on the specific locations where we collected data. The data includes video recordings from several locations at the airport. In the data analysis section we will specify how we analyzed and structured the video data. In the results section we will describe the results gained from the observations.

4.2.1 Method

The contextual analysis was centered around two topics: passengers' behavior when disembarking (exiting) an aircraft, subsequently transferring to another aircraft, and general passengers' behaviors in the lounge areas. We conducted non-participatory observations. As opposed to a participatory observation, in the non-participatory observation the persons being observed are not aware of, or actively involved in the study.

4.2.1.1 Data

Data was collected by the SPENCER project consortium on 11 and 12 June 2014 at the air-side of Schiphol Airport in Amsterdam (Figure 4.2). We collected data at the area

Figure 4.2: Map showing data collection locations. Legend: ■ Narrow hallway (gates), ■ Lounge, ■ Schengen area; no data recorded, ■ non-Schengen area; no data recorded

where the non-Schengen flights arrive and depart; these are mainly intercontinental flights. Both days were average days in terms of traffic. Two researchers collected video data, consisting of open and hidden observations. In an open observation the participants know that they are, and why they are observed. In a hidden observation, people do not know about it. Again, as the goal was to interfere as little as possible in the situation, hidden observations were preferred though this was not always possible. In total we collected 4.5 hours of video data which includes open observations of arriving passengers and hidden observations of passengers walking in Lounge-2.

4.2.1.2 Data analysis

Data analysis has been conducted in three phases, which were based upon an inductive data analysis approach coding, described among others by Lofland, Snow, Anderson, and Lofland (2006, p.200-201). Aspects distinguishing this method from deductive data analysis methods (such as hypothesis testing) include that the method is data-driven rather than theory-driven, and that the researchers in the data analysis process are central in the process. As such, different researchers could come up with different conclusions from the data, as they may make different choices during the analysis process. A third distinguishing aspect Lofland et al. (2006) note is that the data analysis in particular is time-consuming.

Phase 1 consisted of preparing the raw video data for analysis, which included merging, splitting and renaming video files.

During phase 2, the open coding phase, a first categorization was made using open coding (Strauss, 1987, p.28). Two researchers went through all video materials and noted down their observations of the materials. For instance, “passengers are running”. No quantitative statements were made about the data. Also, how the behavior was conducted was not described at this point of time (e.g. one passenger was running behind the other). Based on what we observed, a coding scheme to categorize all video materials was developed by two researchers. This coding scheme can be found in Appendix A. Based on the categories from the open coding, we watched all videos again looking for situations where we observed examples for each of the categories. We cut out the scenes and put them into a folder for the respective category. It was possible to have certain video excerpts in multiple folders. In terms of qualitative data analysis, this coding was done by 1 researcher. In some cases it was not immediately clear to which passenger(s) a video applied, for example in some of the scenes in the lounge area, where numerous passengers and staff were present at a given moment.

Figure 4.3: Open observations of arriving passengers (left) and hidden observations of passengers in lounge 2.

In those cases, the videos were edited with circles indicating where to watch so as to avoid confusion for the coders during the final phase of the analysis. This resulted in a total of 406 short video clips. As some material fitted multiple categories, these do not constitute 406 unique clips.

In the third step, the focused coding phase, all video material was watched again, in order to get to an understanding of how people behaved in certain situations (e.g. how did they position themselves with respect to others and the environment). Finally, observations from the data were written down in tables, and from this conclusions were drawn, in this case for robot behaviors and capabilities.

4.2.1.3 Exclusion criteria

During the first phase of coding we defined behavior which was considered usual, and thus occurring everywhere. One of the most common examples of normal behavior was passengers who were walking either in a pair or alone in a straight line in the hallway. If there was no interference between these people's behaviors and other passengers, staff or vehicles, this data was not coded as it was not informative for our analysis and identification of potentially relevant robot behaviors. This information about usual passenger behavior was used however to determine the passenger flow which we will later use in this analysis.

4.2.2 Results: observations

From the analysis of the videos we found several patterns of behavior which are relevant for the SPENCER project, and for future research into guide robots in semi-public spaces. These patterns can be divided into two categories: patterns based upon passengers' behavior, and patterns based upon passengers' interactions with staff and other passengers.

4.2.2.1 Passengers' organized walking behaviors in flows and pairs

We observed that people organized themselves in flows (Figure 4.4). These flows, or streams, can be classified according to their direction - we noted for example a lot of unidirectional flows in the gates area (traffic in a single direction); which could also

(a) Passenger traffic divided into two unidirectional flows by ways of the moving walkway.

(b) Bidirectional and crossing passenger flows were observed in the lounge area.

Figure 4.4: Infrastructure in the gates area allowed for a natural separation of passenger flows; in the lounge area various flows converged and merged; thus creating bidirectional traffic flows.

have been due to natural separations through infrastructure such as moving walkways (Figure 4.4a). In the lounge area various flows of passengers converged - for example flows from the security check and the Schengen area (Figure 4.4b). As such, there were bidirectional flows and also at times crossflows. This is relevant for the SPENCER project for two reasons.

First of all, research has found that people typically walk slower when encountering a situation with opposite- and crossing pedestrian flows (Daamen and Hoogenboom, 2006) but also when approaching signs at airports (Young, 1999). It would be especially relevant to investigate how fast people walk - e.g. the upper limit of their walking speed as the robot would either have to attain this speed safely, or the use case and type of passengers being guided should be adapted. As we were unable of measuring the speed of individual airport passengers in this study, we have addressed this question in Study 3 (Section 4.4).

The second reason why these passenger flows are interesting for the project is because we noted that larger groups tend to organize themselves in pairs walking behind each other. This occurs in more crowded areas (Figure 4.4b and 4.5a) but also in situations where there was room to walk with four people next to each other (Figure 4.5b). For the perception capabilities of the robot this necessitates the capability to detect people temporarily occluded behind other people.

4.2.2.2 Passengers' (temporarily) waiting in the lounge

From our observations, participants were not always walking directly to their destination, at times they appeared (somewhat) lost, or just waiting. Especially the lounge area, which is also equipped with shops and facilities, seems to be a place where this occurs. An indication of when people are lost appears to be when people are strolling and sometimes stopping. We further observed that these passengers turned their head more than those passengers not searching. In case of a group looking for directions, hand- and arm gestures were observed. In most cases people were looking, and stopping, while walking in a general straight direction. However, at times the opposite was observed in that passengers continued walking, however, they did so while making a zigzag-like pattern through the hallway. We observed passengers to be waiting for different reasons. For groups it was hard to identify the reason as to why they would be waiting, though we can assume they would be similar to those of individual people. These could be to kill the time but also waiting for travel companions who

Figure 4.5: Groups of people typically breakdown into smaller units of 2- and 3 persons.

might be buying something in a shop or taking a bathroom break. A popular place for individuals to wait appeared to be the trashcan, which could be as it is an obstacle anyway, they would not be standing “in the middle of the hallway”, nor form an extra obstacle.

At times we observed that some people stopped briefly, for instance to re-pack their luggage, or search for documents (Figure 4.6). There were two ways people were found to re-pack their luggage: either they held their bag out in front to search something, or they put it on the floor and kneeled down. The first behavior resulted in shorter delays than the second. So in the first case the robot might just slow down while it might have to stop and wait in the second case.

4.2.2.3 Information from monitors and staff

The lounge featured a series of information monitors providing passengers with information such as departure location and boarding time for departing flights. We have observed how people position themselves when looking at the monitor, and what we found was that the passengers were forming half-circles around the information monitors as can be seen in Figure 4.7. The crowdedness varied greatly. Even when it was not crowded around the monitors it could be observed that there are multiple layers (one very close, one halfway the hallway), and that the people were spread out more or less in a half circle. The boundaries of the area where you “should” watch the information monitors appeared to be given by the surroundings; i.e. there were six monitors determining from where people could read the information. For wayfinding robots, it would be useful to detect this social situation, e.g. people facing information monitors, and following detection planning a path in such a way that the robot would not drive in between these people and the information monitors as this would obstruct their view.

Though the airport does not provide staff to actively search for passengers in need of help, staff do walk around, and from what we noticed they are always willing to direct passengers. Assisting someone generally consisted of establishing contact, exchanging information (multiple times) and saying goodbye; leaving. From our observations we noted that the boarding card is used quite often to assist the passenger, and that, when directing passengers staff use head- and arm gestures. For wayfinding robots it could be important to provide directions with pointing features such as arms or heads as well.

Figure 4.6: Passengers stopped briefly - either to re-pack their hand luggage or get documents from their hand luggage.

4.2.3 Discussion

In this section we have reported a contextual analysis to inform the design of the SPENCER robot's motion behavior. We analyzed these data using an inductive data analysis approach by Lofland et al. (2006). This method consists of three phases, and after the third phase we made inferences about the data. In Section 4.2.1.2 we already referred to a number of characteristics which distinguish the data analysis method (coding video data) from data collection- and analysis methods such as one could employ when using for example questionnaire data. In this study in particular, the researchers who collected the data also analyzed the data. An advantage is that events not captured on video (such as informal conversations, or feelings about the environment) could have been included in the data analysis. While more time-expensive, the inclusion of other stakeholders in the data analysis process (such as airport staff and/or passengers) could have led to a results more grounded in everyday processes of staff and passengers. At the same time, additional qualitative measures, such as semi-structured interviews with airport staff, would also have provided additional insights in the challenges airport passengers face when transferring at an international airport.

We identified airport-specific situations the robot is likely to encounter, including passengers who cannot keep up with the robot and have additional goals (such as shops or bathrooms). Similarly, SPENCER will encounter situations such as cross-flows and people looking at information monitors which have to be dealt with in a socially appropriate manner. We will use the situations identified in the two studies in the remainder of this chapter. In the next section we will describe a study in which we investigated how people think a robot should respond when in situations which are likely to occur when guiding a group of passengers. These situations range from passengers temporarily stopping to passengers entering a store to buy a present. In Section 4.3 we present a study to inform the speed of the robot.

4.3 Study 2: Walking Speed Under Time Pressure

In the previous section we reported on a contextual analysis of an international airport, in order to identify specific motion behaviors and challenges for the robot. One of the observations we made was that in the (longer) hallways with unidirectional traffic (traffic in a single direction), passengers walked faster than in those areas with

Figure 4.7: People standing around the information screen in a formation resembling a half circles

shops and intersecting traffic flows. A question that arises is how fast people typically walk. On the one hand this could provide insights in how fast the robot should drive, provided this is technically safe. On the other hand, and this is especially relevant in the hypothetical scenario that the robot would guide passengers under time pressure, it also provides insights into the limitations of current robotic technology; if a robot cannot drive at a speed required by the passengers then such a use case would make little sense.

To answer this question we will first provide a literature overview of related work (Section 4.3.1), and as the literature did not provide a satisfactory answer for the specific situation at hand we have conducted a small pilot study in a lab (Section 4.3.2), which provides directions for the SPENCER project, as well as Study 3.

4.3.1 Theoretical background

In this section, we will address three questions; (1) what is the average walking speed of a small group of passengers?, (2) what is the average walking speed of this small group of passengers when they are carrying luggage?, and (3) what is the maximum walking speed of this small group of passengers when they are under time pressure?

4.3.1.1 How fast is the average human walking speed?

A small body of research exists relating to human walking speed. Whereas some of these studies were commissioned in order to aid facilities planning, other works in the context of rehabilitation looked at average human walking speed in order to derive a benchmark which is required before one can assume that a person's task performance.

Observation studies, typically employing hidden video cameras, provide results which in general conclude that average walking speed decreases by age, and that usually men walk faster than women on average (e.g. Ye, Chen, and Jian, 2012; Willis, Gjersoe, Havard, Kerridge, and Kukla, 2004; Young, 1999; Finnis and Walton, 2008; Laxman, Rastogi, and Chandra, 2010). These studies all look at the walking speed of individual people. Daamen and Hoogendoorn (2006) provide an overview of studies observing real-life pedestrian walking speed. An average of 1.34 m/s was found. While these real-life studies take all sorts of factors into account, such as density and flow, it is more difficult to assess the free speed, which Daamen & Hoogendoorn define as the "speed a pedestrian walks with when he or she is not hindered by other pedestrians". Based upon a series of (semi-controlled) walking experiments Daamen & Hoogendoorn estimated a free speed between 1.44 and 1.63 m/s depending on the context (narrow hallway, crossing flow etc).

As well as observation studies, lab studies have been conducted to assess average walking speed. Example tests include those assessing comfortable and fast gait speed (e.g. Bohannon, 1997). The second type is one where the time in which people walk over a track (usually laps) is kept constant, and where the distance walked is the measure. This test, called the Six-Minute Walking Test (6MWT), has originally been devised as a test to measure the distance people can walk within six minutes on a normal, flat surface (Steffen, Hacker, and Mollinger, 2002) in order to investigate for example test exercise tolerance in chronic bronchitis (McGavin, Gupta, and McHardy,

1976). In this test participants are typically instructed to “walk from one end of a hallway to the other end at their own pace, while attempting to cover as much ground as possible” (Enright and Sherrill, 1998). Six minutes have been found to be a good compromise between obtaining valid results and not exhausting participants too much (Butland, Pang, Gross, Woodcock, and Geddes, 1982). For the sake of the SPENCER project we will limit ourselves specifically to papers using this method with healthy adults.

Enright and Sherrill (1998) provided reference values for the 6MWT by administering the test to 290 men and woman. Mean walking distance was 576 meter for men and 494 meter for woman. Based upon the results from the sample two reference equations for healthy adults have been made based upon gender, age, height and weight. Tsang (2005) conducted the 6MWT with 548 healthy adults in Hong Kong and found an average walking distance of 635 ± 97 m. for males and 573 ± 91 m. for females. Different from Enright and Sherrill (1998) test-retest ability of the measure was assessed by having participants complete the test a second time, three days after the first test. On average the mean walking distance was 10.8 meters higher, which indicates good test-retest ability. These studies shows that the 6MWT distance is (inversely) correlated with age, height and gender; just like observation studies. Given the age and gender of people being guided, one can form hypothesis regarding the speed a wayfinding robot should at least be able to attain with the group comfortably following. For some factors (such as height) perhaps an average should be taken.

4.3.1.2 Is average walking speed of individuals influenced by luggage carried?

As the SPENCER robot will be deployed at an international airport it is likely that passengers will carry items such as carry-on luggage, or push trolleys to carry their hand luggage. We looked into the literature in order to investigate whether or not this could have a potential effect on the speed of passengers. Young (1999) documented the observation of passengers walking in corridors at Cleveland Airport. The number of observations was approximately 1000, and no significant difference was found between passengers carrying one or more luggage items ($M=82.4$ m/min, $SD=13.4$) and passengers who did not do so ($M=78.7$, $SD=14.6$).

Ye et al. (2012) conducted a video observation of passengers walking in a corridor at a subway station in Shanghai. To exclude travel purpose as a factor of observed walking speed, observations were conducted during commuter times. The authors classified passengers in one of three age categories (young, middle-aged, “older”), gender, and the amount of luggage, which was made up of 5 categories. In line with other studies speed was influenced both by age and gender. Luggage significantly influenced walking speed. People carrying medium, large luggage or a trolley case are reported to have a significantly slower walking speed, compared with those carrying no luggage. Small luggage decreased walking speed by 2-3%, medium luggage by 5-8%, large luggage by 10-14%, and a trolley case by 3-8%. Contrary to the findings of Ye et al. (2012), Willis et al. (2004) found that people with luggage - in this case bags - walk faster than people without bags. A reason could be that this study was conducted in the city center of a city, and the people in this case were shopping. As such, we feel it is safe to assume that in a transportation setting people with luggage

will walk slower than people without luggage.

4.3.1.3 Do groups walk slower, compared with individual people?

The SPENCER robot is envisioned to guide small groups of people. The research described so far mainly dealt with the walking speed of individual people. In this section we will try to answer the question if groups walk slower compared with individual people. Research conducted in various context shows that as group size increases, speed decreases (Willis et al., 2004; Federici, Gorrini, Manenti, and Vizzari, 2014; Bandini, Crociani, Gorrini, and Vizzari, 2014). Finnis and Walton (2008) found that people walking with children walk slower compared with those without children. No findings related to group size in general were reported, however, authors did report that people who were carrying luggage and talking with other walked slower (78.8 m/min) compared with average walking speed (88 m/min). From above literature we conclude that average walking speed is negatively correlated with group size. It is therefore necessary to either compare individual and group walking speeds, or limit ourselves to groups in our sample.

4.3.1.4 Research question

From the above literature we know the average walking speed of individual people when they carry luggage, and when they do not do so. There is also compelling evidence that when group size increases, the walking speed decreases. Therefore, we do not know how fast the walking speed of a small group is when carrying luggage. Given that time pressure is a factor to take into account when transferring at an airport, we will include this in our study design as well. The research question of this study is as follows:

How does the average walking speed of a small group of people change when carrying hand luggage, and, is this speed influenced by (perceived) time pressure?

To answer this question we have conducted a lab-based walking experiment.

4.3.2 Method

The aim of this study is to get normative data on the walking speed of triads in a controlled lab setting, both when walking comfortable and under time pressure. This information was not available in the literature, therefore we conducted this study.

We have chosen for a lab-based walking test with 3-person groups (triads), in which we manipulated the *time pressure* (between-subjects) and *luggage carried* (counterbalanced within-subjects). Participants were asked to walk on a pre-set track for 2x3 minutes, after which a self-report questionnaire was administered.

The track area was designed to follow general guidelines for the 6MWT as close as possible. The area was located at an open lab space at the University of Twente, and contained a line of 15 meters, taped on the floor. At the far end of the lines two rectangular cubes were positioned. Every 3 meters a small line was taped perpendicular

to the main walking line. At the beginning of the track a larger line was taped where the groups had to start the walking test. Figure 4.8 shows a schematic overview of the area.

Figure 4.8: Figure depicting experiment area: groups of 3 participants walked laps around a line of 15 meters.

4.3.2.1 Independent variables

Two variables were manipulated. Time pressure was manipulated between-subjects, and had 3 levels. Whether participants carried luggage or not was manipulated within-subjects; the order of which was counterbalanced.

Time pressure (between-subjects) Participants were instructed to imagine themselves being at an airport. Groups were provided a general instruction, as well as one of three motivational messages indicating the time pressure:

1. No pressure: "You have more than enough time to walk to your connecting flight."
2. Mild pressure: "Your previous flight has been delayed by a few minutes due to bad weather, therefore, you will have to hurry a bit to catch your next flight."
3. Heavy pressure: "Your previous flight has been delayed by an hour due to bad weather, therefore, you have to hurry up in order to catch your next flight."

Luggage (within-subjects) Luggage was manipulated within-subjects. Participants were provided either no luggage, or equipped with one luggage trolley each (Figure 4.9). These trolleys were filled with books, so that the weight was 6kg. Every group was exposed to the condition with luggage in either the first or second run. To avoid the effect of fatigue on the manipulation we counterbalanced the order of luggage-no luggage.

4.3.2.2 Dependent variables

The main variable of interest was the distance walked by the participants. After participants had walked, the experimenter put a piece of tape on the floor at the position of the participants. Afterwards the position of the participants was measured.

A self-reported questionnaire (See Appendix B) consisting of 25 items was administered during the study. This questionnaire consisted of 7 questions regarding the

Figure 4.9: A group with luggage trolleys receives instructions

physical effort required during the walking. These 7 questions were administered twice, after each trial. In addition, participants answered 11 questions which dealt with their physical condition and demographics. In general, the questions in the first part of the questionnaire were rated on a 7-point Likert scale, while the questions in the second part were open questions or yes/no questions.

4.3.2.3 Procedure

Participants were welcomed and provided information on the goal and general procedure of the study. As we video-recorded the participants, participants were requested to sign a consent form. Following that, participants were provided the specific airport situation, after which the first 3-minute test was conducted. After having walked 3 minutes the experimenter marked the locations of the participants using small pieces of tape. While the participants filled out the first part of the questionnaire the experimenter noted down the distance walked. Afterwards the second 3-minute test was conducted in the same fashion. Having filled out the questionnaire participants were debriefed, thanked again and offered a piece of cake in reward.

4.3.2.4 Participants

A total of 21 males and 18 females, aged between 19 and 59 ($M=27.49$, $SD=8.78$) participated in the experiment. These 39 participants were divided into 13 groups of 3 people each. Participants were students and staff from the University of Twente. All participants indicated they had previous flying experience. 28 participants (71.8%) have previously transferred at an international airport, of which 4 have the experience of missing their flight.

Figure 4.10: Groups walked less when carrying luggage; this effect was only significant in condition B (mild pressure). Legend: ■ Without luggage, ■ With luggage. ** indicates significant difference at $p < .01$, *** indicates at $p < .001$.

4.3.2.5 Data analysis

In order to see whether or not the three group members stayed together during the walk we analyzed the mean distance of the group and the standard deviation. With the exception of one group, all standard deviations were less than 1 meter; thus indicating that the group ended up at the same location, and stayed together.

4.3.3 Results

Together the 39 participants walked 15,451 meters. Figure 4.10 shows the mean distance groups walked categorized by time pressure condition, and whether or not they carried luggage.

4.3.3.1 Luggage

In order to analyze the effect of luggage on the distance walked we conducted separate Mann-Whitney tests for each of the three time pressure conditions. Participants walked less when carrying luggage (Figure 4.10), though this effect was only significant in condition B ($U=27.50$, $Z=-2.57$, $r=-.52$, $p < .01$).

4.3.3.2 Time pressure condition

The distance groups traversed different significantly by condition, as can also be seen in Figure 4.10 ($H(2)=52.57$, $p < .001$). Post-hoc Mann-Whitney tests indicated the effect being condition A and B being significant different ($U=.00$, $Z=-6.27$, $r=-.85$, $p < .001$), and condition A and C as well ($U=18.00$, $Z=-5.95$, $r=-.75$, $p < .001$).

Similarly, participants experienced the conditions different in terms of physical demand ($H(2)=17.18$, $p < .001$) and task pace ($H(2)=36.31$, $p < .001$), questionnaire items *tlx1* and *tlx2*. Post-hoc Mann-Whitney tests showed that participants found condition A ($M=1.97$, $SD=1.38$) physically less demanding than condition B ($M=3.38$, $SD=1.41$, $U=156.00$, $Z=-3.66$, $r=-.50$, $p < .001$) and condition C ($M=3.29$, $SD=1.52$,

$U=172.00$, $Z=-3.39$, $r=-.46$, $p<.001$). Participants also believed that the task pace in condition A ($M=2.00$, $SD=1.37$) was lower than in condition B ($M=4.54$, $SD=0.93$, $U=71.50$, $Z=-5.23$, $r=-.71$, $p<.001$) and condition C ($M=4.50$, $SD=1.41$, $U=87.00$, $Z=-4.91$, $r=-.67$, $p<.001$).

This indicates that when provided some form of pressure the groups walked further. The specific type of condition did not seem to influence the results in this study.

4.3.4 Discussion

In this study we found that a group of people walking an unidirectional path do so at a speed between 0.82 and 1.33 m/s, depending on whether or not they are under a perceived mental time pressure. These findings are consistent with the works of Federici et al. (2014) who reported an average speed of 1.21 m/s for triads and Willis et al. (2004) who reported an average speed of 1.36 m/s. In that sense this pilot does not add any new knowledge. In line with literature there was a negative correlation between age and the distance a person walked, $\tau=-.187$, $p<.05$. At the same time the age of the sample was highly skewed, and as such we will refrain from drawing any conclusions from this.

What we did not find was a significant influence of luggage on the walking speed of the groups. When carrying luggage, groups walked between 4-7% slower as compared to when they would carry luggage, which is comparable to the finding of Ye et al. (2012) who found that (individual) people walked between 5-8% slower when they carried "medium sized luggage". Given that the influence of luggage did not significantly influence the walking speed in this pilot study, we have not included the factor into the design of Study 3.

The third finding of this study is that the simulated time pressure, only using verbal instructions, seems to have been successfully influenced participants, as they walked further. This type of manipulation has been included in the design of Study 3.

4.4 Study 3: Social Situations at the Airport: How Everyday People Think a Robot Should Act

In this section we report on an exploratory study in which we investigated what people think a robot should do when conflicting social situations arise; for example in cases where a part of the group urgently needs to use the bathroom.

Based upon the contextual analysis we conducted (Section 4.2) we defined 10 social situations which are likely occur at an airport. We manipulated variables, such as group size and time pressure, based upon the results gathered in the previous studies reported in this chapter.

We developed and distributed a web-based survey to people living in the United States and China as previous research in HRI has shown that there are cultural differences in people's attitudes towards (behaviors of) robots (e.g. Eresha et al., 2013; Li, Rau, and Li, 2010), including the two countries under investigation in this study (Joosse et al., 2014). We collected and coded qualitative data from 118 participants who described how a robot should act in a variety of situations.

4.4.1 Motivation

In Section 4.2 we presented a contextual analysis, with the goal of getting a better understanding of the future operating environment of a wayfinding robot at an international airport. In particular in the lounge area we found that people encounter all sorts of distractions such as shops and bathrooms. Also, we noted that this area could get particularly crowded, with numerous pedestrian flows intersecting one another. The original use case of the SPENCER project called for a robot to collect (and guide) passengers who share the same arriving and connecting flight, but do not necessarily know each other. These passengers could be considered subgroups within the larger group which is being guided by the robot.

We are unaware of other studies in Human-Computer Interaction (HCI) or HRI which investigated the behaviors a robot should display when encountering situations relevant for the SPENCER context. Therefore we have chosen to conduct an exploratory approach rather than providing participants' with a pre-defined set of answers from which they could choose the "best answer".

4.4.2 Hypotheses

The study reported in Section 4.3 showed that people behave differently when under (perceived) time pressure with respect to how fast they walk. One of our assumptions is that participants under time pressure are less inclined to have the wayfinding robot stop (or slow down) when they are being guided to their departure gate. We thus expect that:

H1 Under high time pressure, participants prefer a robot which does not slow down.

In a similar fashion to time pressure, we reported in the literature section preceding Study 2 (reported in Section 4.3.1) that people's walking speed decreases as group size increases. As this could have an influence on the time it takes the group to reach their target we pose the following hypothesis:

H2 In larger groups, participants are less likely to prefer the robot to slow down.

4.4.3 Method

We designed a survey containing descriptions of situations where a robot guides a group with an interruption of the guiding process, and which are likely to occur at an airport. Based upon the contextual analysis (Section 4.2) and the researchers' own experiences transferring we defined 10 situations which could occur when transferring:

1. A member of the group cannot keep up with the robot.
2. A member of the group forgot his/her passport (or other important documents), and thus has to return to the collect them.
3. A member of the group enters a store to buy a present.
4. A member of the group urgently has to go to the bathroom.

5. An elderly person, not part of the group, crosses the path of the robot, requiring the robot to slow down.
6. A member of the group stops to tie his/her shoelace.
7. A member of the group is lost, and cannot find the robot.
8. A member of the group stands in the way of the robot to take a selfie.
9. A member of the group wants to smoke a cigarette.
10. A member of the group meets someone familiar, stops and starts a conversation.

4.4.3.1 Independent variables

In this section we will discuss the independent variables of the study. We manipulated five independent variables. The first variable we manipulated within-subjects was the *situation*, which have been described in the previous paragraph.

The second and third manipulations relate closely to the hypotheses; both *time pressure* and *group size* (Figure 4.11) were manipulated in a 2-level between-subjects fashion.

The fourth and fifth variables we manipulated concerned the *applicability* of the social situation. The situation either concerned the minority or majority of the group, and the participant filling out the questionnaire was also *involved or not involved* in the problem (e.g. part of the group who has to go to the bathroom).

To summarize these manipulations: this study with a hybrid model contained 2 between-subjects factors (1 10-level, 1 2-level) and 3 within-subjects factors with each 2 levels. Therefore, 1 of 8 designs was presented to each participant. An example of a situation which we provided to participants is the following text:

Please imagine yourself being at a new airport where you transfer from one flight to another. The robot in the picture guides you and 2 other passengers, whom you haven't met before, to the same gate where the connecting flight will depart. You have more than enough time to walk to your connecting flight. Please imagine the following scenario:
The minority of other group members falls behind and can't keep up with the speed of the robot guide and the rest of the group.

As stated above participant involvement in the situation was manipulated within-subjects, with the exception of situation 5, therefore all participants were provided with $(20-1=)19$ situations. The wording of each situation (given a manipulation) is available in Appendix C.

4.4.3.2 Dependent variables

We collected data through a web-based survey. The survey started with an overview of the survey, and technical data of the SPENCER robot, including the height and weight, capabilities and interaction modalities. Following this information, participants were provided with the situations at hand. After having provided their expectations of the robot's response to 19 situations, we included a number of questions about participants' experience with robots, as well as demographic questions. Specifically, we included questions to get insights in participants' demographic information (4 items),

Figure 4.11: Participants were asked to imagine they were either part of a small (left) or larger (right) group.

flying experience (5 items) and experience with robots (11 items). Two manipulation checks were included to assess if participants correctly recalled the size of the (fictional) group they were part of and (fictional) level of time pressure they were under. We further collected data on participants' personality (10 items, Rammstedt and John (2007)) and individualistic and collectivistic orientations (16 items, Triandis and Gelfand (1998)). This resulted in a survey containing a total of 66 items. The complete questionnaire is available in Appendix C. The survey itself was hosted on SurveyMonkey, and participant recruitment and payment was conducted through the CrowdFlower platform.

4.4.3.3 Participants

The survey was distributed to participants in China and the United States. A total of 225 participants completed the survey. We excluded 107 participants (47%) who satisfied one or more of the following three exclusion criteria:

1. Participants whose answers did not consist of text (words, sentences) but rather random numbers or strings of characters.
2. Participants who did not provide complete sentences, or the same answer for all 19 situations)
3. Participants who failed to answer the manipulation check question.

Data of 118 participants was included in the data analysis. Table 4.1 provides information on the distribution of participants in each condition. A total of 72 males and 45 females participated in the study, aged between 17 and 66 ($M=32.33$, $SD=11.99$). The sample had a diverse background and consisted of students (16.9%), unemployed people (11.0%), and people with diverse occupational backgrounds such as education (11.0%), construction (5.1%), health care (5.9%) and finance (5.1%).

	Time pressure	Group size	Applicability	United States	China	Total
A	high	small	minority	9	10	19
B	high	small	majority	11	9	20
C	high	larger	minority	8	5	13
D	high	larger	majority	8	5	13
E	low	small	minority	6	5	11
F	low	small	majority	11	7	18
G	low	larger	minority	7	5	12
H	low	larger	majority	6	6	12
Total				66	52	118
Gender			male	37	35	72
			female	28	17	45
			(unknown)	1		
Age				$M=35.29$ $SD=12.77$	$M=28.63$ $SD=9.85$	$M=32.33$ $SD=11.99$

Table 4.1: Participant distribution over conditions, including basic demographic information

4.4.3.4 Data analysis & dataset descriptives

The data were coded by two coders (31% of the data was coded by both coders), and each answer was initially assigned one or multiple codes out of a potential 27 (Appendix D). As it proved impossible to provide a meaningful analysis of these data, it was eventually decided to narrow the analysis, and group a number of codes into 1 of 10 categories. We then assessed intercoder reliability based upon that category, which was moderately-high and deemed acceptable ($\kappa=0.71$).

Due to the nature of the data (categorical data) we could not run tests such as mixed-model ANOVAs. We have chosen not to run complex tests such as log-linear analysis or generalized linear modeling but to run a series of chi-square tests, and specifically look at the standardized residuals, which shows whether or not gathered values differ significantly from the expected value (Field, 2009, p. 699).

One manipulation (involvement in the situation) was manipulated within-subjects: participants were provided 10 situations in which they were not actively involved in the cause of the situations, and 9 situations in which they were. Running the aforementioned chi-square tests on these data would violate the assumption of independence of data (Field, 2009, p. 691). To remedy this problem we only analyzed data of participants' answers to the 10 (fictional) situations in which the participants were not actively involved situations. The dataset therefore contained a total of 1180 answers², ranging from short answers of 4 characters to stories of 255 characters.

Before presenting the results, we will first present some characteristics of the coded dataset. Most of the data was coded in either the "wait" (31.0%) or "continue" (25.9%) categories (Table 4.2). Roughly 16% of the data could not be coded into one of eight categories describing actions for the robot. Data from female participants was significantly less frequently coded as "other" ($z=-2.0$) and "unclear" ($z=-2.4$);

²The sample contained 118 participants, from each we included the answers to 10 questions (1a, 2a, 3a, 4a, 5, 6a, 7a, 8a, 9a, 10a in Appendix C).

thus less ambiguous.

	Code	%
1	Continue	25.9%
2	Slow down	6.0 %
3	Wait	31.0%
4	Assist person	3.2%
5	Attract attention	1.9%
6	Go with group	3.5%
7	Provide information	8.0%
8	Involve other	3.5%
9	(other answer)	8.4%
10	(answer unclear)	8.6%

Table 4.2: Data was categorized into 1 of 10 categories. Two coders coded 30% of the data ($\kappa=0.71$). Most data was coded as "Wait".

4.4.4 Results

None of the chi-square tests in combination with standardized residuals yielded significant results (all residuals $\leq \pm 1.90$). For the purposes of this chapter we took a look at the results, and we visualized the most relevant results in Figure 4.12. The way the data is reported is as follows: we report the percentages of the top-3 categories in which most answers (sorted from left-to-right following a majority voting principle). Where relevant, we have also included a division based upon one of the manipulations. This is not an exhaustive list of the various differences within the conditions, however, we feel these are the most interesting for the purposes of the thesis. We will discuss some of the results and future directions for research in the remainder of this section.

4.4.4.1 Let's move along now

In 6 out of 10 situations most of the participants indicated the robot should continue or perhaps slow down. Two of four situations where the robot was expected to wait were actions which do not necessarily have to take up a lot of time; these are the situations where a group members ties his or her shoelace and when someone takes a selfie. Though we have not asked participants why they would consider an action appropriate, it could be that in the other two cases (buy a present and go to the bathroom) participants might have found these actions more important. At the same time we should note that while "wait" was coded the most, this category represents far from an absolute majority. The data suggests that when people enter a store to buy a present, or have to go to the bathroom, conflict could arise within the group as a part of the participants indicated the robot should continue, contrary to waiting. An explanation can be found in Figure 4.13a for the scenario in which a portion of the group has to use the bathroom, about 30-40% of the participants indicated the robot should wait. However, when the situation applied to the majority, this number almost doubled to 60%.

Figure 4.12: Top-3 of coded answers, per question. Where relevant divisions have been made on e.g. group size or situation applicability.

Figure 4.13

4.4.4.2 The needs of the many outweigh the needs of the few

In a similar fashion, when conflict arises it would seem that the needs of the many might outweigh the needs of the few. Figure 4.12 shows that in three situations participants were less likely to have the robot continue when a situation applied to the majority of the group as opposed to when the situation applied to the minority of the group. Interestingly, these appear to be situations of which the expected duration is unpredictable. Participants expected the robot not to wait but continue in situations where (#2) someone lost his/her passport and has to go back to retrieve it (Figure 4.13b), (#9) someone wants smoke a cigarette, and (#10) when a person meets someone familiar, stops and starts a conversation.

4.4.5 Discussion

In this section we presented the results of a study we conducted to investigate how people expect a guide robot to act when encountering conflicting social situations, while guiding a group in an airport. We collected quantitative data which we coded and analyzed. Despite not finding significant results, we do believe the results can provide a starting point for future investigations in appropriate robot behaviors, for example by providing a number of categories in which we coded answers.

A second way of using these results could be by designing scenario's featuring a number of robot responses to an event (such as a group of people entering a store), and subsequently having these responses evaluated by people. Examples of these video-studies include the works by Woods et al. (2006b); Srinivasan and Takayama (2016); Read and Belpaeme (2014). Similar to the current study every participant would see the same stimulus, and actually see the stimulus instead of reading the description of a situation, which participants could interpret differently. Lichtenthäler (2014, p.42-43) however point out that video-based studies are not suited when investigating behavioral or psychophysiological measures as feelings (such as fear, anxiety, joy) would be less distinctive when watching a video as opposed to experiencing

something in real life.

We acknowledge that some of the situations chosen would not be easy for a robot to detect, e.g. if people walk away a robot does not know if they have lost their passport, or believe they should take a different route. Simple rules could be implemented such as slowing down when people cannot keep up, and continuing when situations only apply to a minority of the group. Straightforward rules guiding people's everyday behavior could indeed be the default mode for a robot when being deployed in complex social environments.

4.5 Chapter conclusion

In this chapter we studied the context in which the SPENCER robot is envisioned to operate. We studied this through three studies with which we addressed research question I: "How do people expect a robot to respond when it encounters social situations while guiding a small group of people?"

In Study 1 we identified passengers' behaviors and routines at the airport, which primarily contributed to recommendations for the SPENCER robot's evaluation scenario (Study 9). In Study 2 we investigated how fast the walking speed of a group of people is when carrying luggage, under high or low time pressure. We found that the use of a simulated time pressure increased participants' walking speed but that carrying luggage did not significantly influence walking speed. The results of Studies 1 and 2 contributed to the design of Study 3, in which we investigated how people believed a wayfinding robot should respond when it encountered various conflicting social situations while guiding a small group at the airport. The analysis of the coded data led us to conclude that in general a robot should continue, as long as the situation is not applicable to the majority of the group. The take-home messages of this chapter are as follows:

- When navigating through airports, passengers spatially organize themselves in various ways. Specifically, participants' form a half-circle around information monitors, and organize themselves in pairs of 2-3 when walking. Taking into account the perception capabilities of the SPENCER robot, these findings influence maximum group size which may be tracked by a robot.
- The walking speed of a small group increases when under (simulated) time pressure.
- When encountering conflicting social situations while guiding a group through an airport, people believe a wayfinding robot should tend to the needs of the many, e.g. in general continue guiding the remainder of the group if they are in the majority.

These results influence the design of the SPENCER robot's behavior, which will be evaluated in a case study with airport passengers in Chapter 8. Together with the literature previously presented in Chapter 2, this chapter identified a number of relevant behaviors for wayfinding robots, such as a robot's distancing behavior

when interacting with a small group (Chapter 5) and head turn (gaze) behavior when guiding people (Chapter 7).

5

Robots Approaching Groups of People: The Role of Spatial Arrangements and Culture

"It's a primitive culture. I'm just trying to blend in."

Deanna Troi.
Star Trek: First Contact (1996)

The work reported in Section 5.3 has partly been presented as: Vroon, J.H., Joosse, M.P., Lohse, M., Kolkmeier, J., Kim, J., Truong, K.P., Englebienne, G., Heylen, D.K.J. & Evers, V. (2015) Dynamics of Social Positioning Patterns in Group-Robot Interactions. In Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (Ro-Man 2015).

The work reported in Sections 5.2 and 5.4 has partly been presented as: Joosse, M.P., Poppe, R.W., Lohse, M. & Evers, V. (2014) Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People. In Proceedings of the 5th ACM International Conference on Collaboration across Boundaries: Culture, Distance & Technology (CABS 2014).

In this chapter we will present two studies investigating how a robot should approach a small group, with respect to approach distance and -direction. The first study in this chapter reports on a data collection of small groups' interactions with a robot which approached them several times from different directions. In the second study we investigated whether people from different cultures have different preferences for robot approach distances.

5.1 Introduction

In Chapter 2 we reported a comprehensive overview of user studies investigating personal space in HRI. One of our conclusions was that this body of research is primarily concerned with the investigation of single persons interacting with a robot. When operating in public space, it is likely that robots will have to approach small groups of people (between 2 and 4 persons), if only because up to 70% of the people in crowds

are part of small groups (Moussaïd et al., 2010). Therefore it is necessary to investigate how groups of people interact with robots. With this chapter we aim at answering RQ-II, which consisted of two questions:

RQ-IIa How should a robot approach a small group of people, in terms of direction and distance?

RQ-IIb Do people from different cultures prefer a robot to approach them differently?

In this chapter we will present two complementary studies in which we investigate proxemics behavior for a robot interacting with a group (Figure 5.1). The first study is an explorative study conducted jointly with the TERESA project¹, and was conducted to collect subjective and objective data on good and bad robot approaches when a robot interacts with a small group of people (Section 5.3). In the second study we used a questionnaire to investigate cultural differences in preferences for human-robot approach distances and directions, also when approaching a group (Section 5.4). These studies complement each other as the second study allowed us to collect data from a culturally diverse sample. The disadvantage of this method is that we could not take the dynamic nature of groups into account, which we managed to do in the first study. In both studies we studied the approach behavior of a robot when approaching a small group of people. When people interact in groups - for example to have a conversation - they arrange themselves in a spatial formation, which can have different arrangement. Literature on these arrangements has not been discussed in prior chapters, therefore we will first do so in Section 5.2.

Figure 5.1: Overview of the layout of chapter 5.

5.2 Theoretical background: The Spatial Arrangement of Small Groups

People organize themselves spatially not only through interpersonal distance but also in terms of their spatial arrangement when they are part of small groups. Kendon (1990, 2010) introduced the concept of F-formations to capture this phenomenon. According to Kendon, activity is always located in a space. This space can be called the “transaction segment”. When two or more people form a temporarily stable group, they arrange the spatial formation of the group in such a way that the individual transaction segments overlap; thus creating a shared transactional space. This space is also known as an F-formation, and can have different spatial arrangements - according to Marshall, Rogers, and Pantidi (2011) this is short for *facing formation*. Within an

¹The TERESA project was funded by the European Community’s Seventh Framework Program (FP7/2007-2013) under grant agreement no FP7-ICT-611153 (December 2013-2016).

F-formation different spaces can be distinguished (Figure 5.2): participants interact around shared inner space, called the O-space (Kendon, 2010). The O-space is surrounded by a narrow space in which people are located (the P-space). Therefore, in order to join a group a new member has to get access to the P-space. The remainder of the group will then adapt the spatial arrangement to accommodate the new member.

Figure 5.2: Circular F-formation around an o-space

The spatial arrangement of an F-formation thus depends on the number of group members and also on the type of interaction (Figure 5.3). When a third member joins an existing L-shaped formation (Figure 5.3a), it can become a circular formation 5.3e.

The role different people take on in the F-formation could be related to their spatial position. For instance, Kendon observed that speaking rights are reflected in the formation. In a circular arrangement, rights tend to be equal. In other arrangements such as a rectangular arrangement (Figure 5.3f) the one in the spatially differentiated position (i.e. the one person positioned opposite others) has the right to speak more

Figure 5.3: Groups in F-formations can assume different spatial arrangements. Based upon Kendon (1990) and Marshall et al. (2011).

compared with others. Marshall et al. (2011) report on observed interactions in a tourist office, for example when one member of a group of tourists would present a plan to the group the group could form a semi-circular formation around the member presenting.

Rehm, André, and Nischt (2005) reported the six most frequently occurring formations, and divided these six into open and closed formations. People in open formations (Figures 5.3a, 5.3c and 5.3d) are said to allow others to join the conversation while this is not the case with closed formations (Figures and 5.3f). In their experiment with virtual characters, Rehm et al. (2005) found that participants were more likely to join an open formation (84% of the trials) than a closed formation. All participants positioned themselves at a social distance, half in the close-social, and half in the far-social distance. However, the authors found that two Arabian participants positioned themselves in the close-social space, which is consistent with findings in cross-cultural research in that Arabian people generally stand closer to each other.

We now know that people organize themselves into spatial arrangements which allow all participants access to the joint transactional space. We do not know yet if people form a similar spatial arrangement when a robot is part of the group, or becomes part of the group when it approaches the people. It is likely that robot will have to approach groups when operating in public space (Moussaïd et al., 2010). We have therefore conducted two studies in which we investigate human-robot approach distance- and direction in situations where a robot approaches a small group.

5.3 Study 4: A Telepresence Mystery Investigation

This section describes an explorative study we conducted in order to collect data on people's motion behaviors when approached by a robot. The goal of this study was two-fold. The main goal was to generate hypotheses on suitable social positioning behaviors for a robot interacting with a small group of people. The second goal was to confirm findings found in related work on human-robot approach distances and directions, for a group setting.

To achieve these goals, we created a setting in which groups of four participants had to solve a (fictional) murder mystery. One of the participants was present through a telepresence robot, and had to drive around the room to collect information necessary to solve the mystery. We will describe the setup of the data collection and the results which are relevant for this thesis.

5.3.1 Hypotheses

This study was set up as an explorative study. However, to further specify the second goal, we specified three hypothesis concerning the patterns (or results) we would expect to see based upon related work on human-robot approach distances and behaviors. Additionally we specified one research question. These were based on the results from overview of previous studies we reported in Section 2.4.1.

People joining an F-formation have to be granted access to the P-space in a formation by the other group members (Kendon, 2010). One way of allowing people to join is by re-arranging the formation in a way that creates an opening in the P-space

where the new member can join. Rehm et al. (2005) found that people who are in an open formation are more likely to let other people join the conversation as opposed to people in a closed formation. Therefore we will test whether people will be more positive when a robot approaches the group at a position where there is space. To operationalize this we will manipulate the group formation; in one condition participants will be arranged in a formation with a distinct open space in the P-space (Figure 5.4, green area), in the other formation participants will be arranged in a circle with congruent angles (Figure 5.3e). We can then objectively test where the robot operator joins the group and evaluate how the group appreciates this behavior.

H1 People rate a robot's approach in an open space as more positive compared with an approach in a closed space.

Figure 5.4: A circular formation with a distinct opening (green area) in the P-space

Research in HRI on approach directions found that both individual (Syrdal et al., 2006) and dyads (Ball et al., 2015) have a preference for an approach direction where they face the robot. We therefore hypothesize that individual participants will rate the approach as more comfortable when the robot starts its approach from a position where the participant can see the robot.

H2 People will find an approach of a robot more comfortable when the robot starts from a position where they can see the robot as opposed to a position where they cannot see the robot starting its approach.

Previous research with this particular robot platform indicated that participants found this platform easy to operate, among others due to the user interface and the (lack of a steep) learning curve (González-Jiménez, Galindo, and Ruiz-Sarmiento, 2012). Presuming that the robot operators become familiar with the robot themselves at a steady pace, we hypothesize that this skill is also visible in their handling of the robot.

H3 The telepresence robot operator is perceived as becoming more skilled over time.

The final question under investigation relates to the approach distance towards a group. In Section 2.4.1.3 we reviewed literature in which (minimally comfortable) human-robot approach distance was measured and reported. Without exception these studies report on interactions between individual people and a robot - in the case of

Eresha et al. (2013) two robots. With this study we would like to get an initial idea of the distance the operator of a telepresence robot deems appropriate to keep from a group.

RQ1 How much distance does a telepresence robot operator keep from a group when approaching the group?

5.3.2 Methodology

We created a situation where groups of three participants were approached multiple times by a telepresence robot controlled by the fourth participant (we call this person the visitor). To assess the visitor's learning curve all participants were provided (short) questionnaire in between the approaches. We manipulated the group formation of the participants in the experiment room (between-subjects) and approach direction of the robot (within-subjects).

5.3.2.1 Task: solving a mystery

In order to collect data across multiple conversations, we designed a task which motivated the visitor to approach the group, have a conversation with the group, and leave the group several times. Participants were provided with a (fictional) murder case and assigned to identify the most likely culprit. To identify the culprit, the visitor had to drive around the room and collect clues, return to the group, and share the clues with the group. To eliminate effects of the specifics of the murder mystery, groups were randomly assigned to one of three murder mysteries. The victims and murder locations included respectively a man in a school building, a rich philanthropist in his house and an office worker in her office. All murder cases were equal in that they contained eight clues, had one person murdered, and multiple possible suspects. For all cases, all eight clues were necessary to identify the culprit. Preliminary analysis did not indicate any effect of the different murder mysteries and therefore this variable has been excluded from the analysis. Though the context of this task differs from an airport context, the results gathered can be applied in any situation where a robot has to approach a group of people, including airports as we discovered in Study 1 (Section 4.2).

To make sure that participants would not try to rush they were instructed not to collect the next clue until a buzzer had sounded. The buzzer sounded 75 seconds after the previous clue was presented, which gave ample time for both approach and conversation. Each group of participants was part of a total of eight approach/converse/retreat cycles, separated by the visitor having to go to a marker to collect the next clue and all participants completing a brief questionnaire.

5.3.2.2 Independent variables

We manipulated the formation of the group and the approach direction of the robot, as can be seen in Figure 5.5. Two different spatial formations were chosen for the participants in the experiment room, based upon the work of Kendon (1990). In our study we have chosen for two formations, both of which can be seen in Figure 5.5. The

first formation is a circular arrangement, with every participant occupying an equal amount of space (Figure 5.8a). In the second formation, a semi-circular arrangement, there is an open space where the robot operator could potentially position the robot (Figure 5.8b). For the sake of simplicity we refer to these two formations as closed formation (circular arrangement) and open formation (semi-circular arrangement).

Figure 5.5: Participants in the experiment room either assumed a circular or semi-circular formation

The approach direction of the robot was manipulated by having the visitor pick up each of the clues at a different marker, positioned around the experiment room (see Figure 5.5). To eliminate possible ordering effects, the order of the markers was counterbalanced with a balanced Latin square design.

In the study, participants in the experiment room were allowed to re-position themselves, and make small steps when the robot approached. However, the participants were requested not to approach the robot under the guise of limited tracking capabilities. In between the different approaches the formation of the participants was reset by asking the participants to assume to original formation which was temporarily projected on the floor.

5.3.2.3 Dependent variables

During the experiment we collected objective data with various sensors and subjective data with questionnaires.

Objective measures All three participants in the experiment room were equipped with uniquely identifiable markers (one on the back of the chest, one on a cap), which were tracked by an OptiTrack² motion capture system using eight infrared cameras. The system used allows sub-centimeter level precision tracking of both position and orientation of each marker. We optimized tracking for the center of the experiment room, to make sure we could properly capture the interaction. Markers near the edges of the experiment room could often not be tracked. To ensure proper tracking of the actual interaction, we informed the participants about this and asked them to not get too close to the edges of the experiment room. Two cameras recorded the experiment

²<http://www.naturalpoint.com/optitrack/>

Item	Participants	Visitors
How comfortable were you with the approach of the robot?	x	
How comfortable you were with the distance between you and the robot during the conversation?	x	
How comfortable were you with how the robot drove away?	x	
How much effort did you have to put in driving the robot?		x
How successful do you think you were in driving the robot?		x
Did you feel discouraged, irritated and/or stressed while driving the robot?		x
How useful was the clue you have just been provided with?	x	x
How close do you think you are to solving the puzzle (0-100%)?	x	x
Total	5	5

Table 5.1: Items of the in-between questionnaire for the participants in the experiment room and the participants operating the robot (the visitors). All items were administered on a 5-point Likert scale.

room. One camera recorded a side view, the other camera a (fish eye) top down view of the interaction.

Subjective measures After each approach/converse/retreat cycle, all participants were given a short questionnaire (referred to as “in-between questionnaire”). After the session a post-experiment questionnaire was administered.

The in-between questionnaire consisted of five questions; two related to the usefulness of the clue and task progress. The remaining questions measured comfortability with the robot operators’ driving behavior during approach and retreat, and with the distance to the robot during conversation. For the robot operator, we instead assessed *workload* (based on Hart and Staveland (1988)). See Table 5.1 for the items used.

The post-experiment questionnaire consisted of 49 items. The complete post-experiment questionnaire is available in Appendix E. Among others we measured *co-presence* ($\alpha=0.78$) and *attentional engagement* ($\alpha=0.74$) (Biocca and Harms, 2003). These measures were rephrased to reflect the context of the situation following Kristoffersson, Coradeschi, Eklundh, and Loutfi (2013a). Two items from Kristoffersson et al. (adapted from the TPI’s perceptual realism scale by Lombard, Ditton, and Weinstein (2009)) were used to measure *person realism* ($\alpha=0.55$). To measure if the participants were *predisposed to like robots* or not we designed four items that capture various aspects of robot acceptance. These items were based on the Almere framework (Heerink, Kröse, Wielinga, and Evers, 2006), as standardized questionnaires (such as Nomura, Kanda, Suzuki, and Kato (2008)) are not specifically designed for telepresence robots. *Workload* was measured by the 6-item NASA-TLX (Hart and Staveland, 1988). These items were slightly rephrased, and asked on a 10-point Likert scale. *Demographic questions* asked for general demographic information such as gender, age, nationality and closeness with peers (Aron, Aron, Tudor, and Nelson, 1991). We assessed *personality* using a 10-item Big Five Personality Inventory (Gosling, Rent-

Figure 5.6: Three participants worked with a fourth participant, who was present through a Giraff telepresence robot to solve a murder mystery (a), both the participants (b) and robot (c) were equipped with motion tracking equipment.

frow, and Swann Jr, 2003).

5.3.2.4 Procedure

Participants were recruited as part of an undergraduate research methods course. After a briefing, participants were randomly assigned to either be the visitor (1 participant) or be a participant in the experiment room (3 participants). This was followed by task-specific instructions from one of the experimenters. The participants were equipped with everything required for the data collection (such as the caps with markers while the visitor was given a brief training on controlling the Giraff (changing position, orientation and head tilt).

The visitor approached the group for a total of 9 times³ (Figure 5.6). The first eight times the visitor approached the group from one of the eight markers as shown in Figure 5.5. To eliminate possible ordering effects, the visitor had to go to the different markers in one of eight randomly assigned counterbalanced orders. The final approach was from the same marker as the first approach.

At the end of each cycle, before being given the next clue, we asked participants (individually) to fill in the in-between questionnaire concerning the robot behavior during that cycle. The next clue was presented after all participants had finished filling in the questionnaire. While filling in the questionnaire at the end of each cycle, the participants in the experiment room were asked to assume either the circular- or semi-circular formation. The projections were not shown during the cycles and we explicitly told our participants that they were allowed to move around during the cycles.

At the end of the interaction part of the study, and after the group had reached a consensus on their primary suspect, we asked all participants individually to fill in the post-experiment questionnaire. After having completed the post-experiment questionnaire, all participants were debriefed.

³The final approach included approaching the group and having a conversation; as there was no retreat involved this was not a full cycle.

5.3.2.5 Data analysis

With the in-between questionnaires we collected data on participants' subjective ratings of a robot's motion behaviors during approaches, conversations and retreats. With the exception of the final question, all questions relate to how the participant perceived the behavior compared to the behavior in the previous cycles. The raw data are only meaningful if we were to interpret them in light of the mean and standard deviation of each individual participant. An example is provided in Figure 5.7). If two participants (P and Q) were to rate comfortableness as a "4" their raw score is the same. However, when looking at the entire distribution of scores for a single participant this actually shows that participant P believed the approach was very comfortable, whereas participant Q believed this was not comfortable at all.

Figure 5.7: A raw score by itself does not provide information about the location of the score in the distribution.

In order to make sense of the data we transformed the data in a way that for the distribution for each participant the mean and standard deviation would be respectively 0 and 1. To do so, we have standardized the raw ratings into Z-scores⁴ using the following formula: $z = \frac{X - \bar{X}}{s}$.

5.3.2.6 Participants

A total of 56 participants participated, divided into 14 groups of 4 persons. All participants were students, aged between 18 and 32 ($M=20.09$, $SD=2.20$). 43 participants were male (76.8%), 13 female (23.2%). Most participants had the Dutch nationality (85.7%). All participants were enrolled in a technical major; 9 students were enrolled in Business & IT (16%), 27 in Computer Science (48%), and 20 in Creative Technology (36%). Only 10.7% of the participants had no prior experience with robots; 25% had seen robots before; 50.0% had played or worked with robots before and 14.3% had built robots before. Coincidentally the level of experience with robots of the Visitors who operated the robot was higher ($M=3.57$, $SD=1.09$) than the participants in the experiment room ($M=2.74$, $SD=1.15$), $U=180.00$, $Z=-2.242$, $p<.05$. When asked towards their specific experience with telepresence robots, 51.8% of the participants had never seen a telepresence robot before, 23.2% had seen a conversation through a telepresence robot on television or the internet. 25.0% of the

⁴An illustrative example is provided in Field (2016, Chapter 6).

Figure 5.8

participants had seen a conversation through a telepresence robot in real life. Most participants indicated they knew the other group members relatively well.

5.3.3 Results

In this section we will present the results of this study; we will address the three hypotheses and research question previously defined in Section 5.3.1.

5.3.3.1 Approaching an open formation not rated higher

A Mann-Whitney test showed that there was no significant difference between the rating of the open ($M=5.87$, $SD=1.10$) and closed formation ($M=5.85$, $SD=1.191$), $U=17794.5$, $Z=-0.065$, $p=.95$. Therefore, we have not found support for H1 in which we stated that participant would have a preference for a robot to approach in an open spot in the semi-circular arrangement.

5.3.3.2 Participants appear to be more comfortable when they can see the robot

In order to analyze the second research question we compared ratings of participants between approaches from markers where they could see the robot and where they could not see the robot. We first identified from which markers a participant could not see the robot approaching (Figure 5.8). For participant 1, these are the approaches from markers B and C in the circular formation, and marker C in the semi-circular formation. For participant 3, these are the approaches from markers G and H in the circular formation, and marker G in the semi-circular formation. For participant 2 this only concerns marker E.

Participants standardized ratings were lower when approached by the robot from an angle they could not see ($M=-0.18$, $SD=0.92$) as compared to situations where they could see the robot approaching ($M=0.03$, $SD=0.92$), $U=6726.00$, $Z=-1.692$, $p=.09$. Though this was not a significant difference, in future research we can take this into account as a potential confounding variable.

5.3.3.3 Robot operator became more skilled over time

Participants' standardized ratings of the robot's approach were significantly affected by the approach number, $H(8)=70.498$, $p<.001$. A Jonckheere-Terpstra test revealed a significant trend in the data: as the participants went through the experiment, they rated the approaches as more comfortable, $J=42030.00$, $z=8.45$, $r=.44$. The participants thus perceived that the robot operator operated the robot in such a way that they felt more comfortable.

5.3.3.4 Approach distance towards a group

The average distance between the robot and the center of the group was 1.13 meter ($SD=0.42$) at the end of the approach. This is useful up to a certain extent, as this number does not tell us anything about the distance between the robot and individual group members (Figure 5.9). Therefore we also report the average distance between the robot and the interact targets closest to and furthest from the robot. The average distance between the Interaction Target closest to the robot ranged between 0.25 and 2.68 meters ($M=0.96$ m, $SD=0.34$ m). The average distance between the Interaction Target furthest away ranged between 0.45 and 3.61 meter ($M=1.56$ m, $SD=0.42$ m).

Figure 5.9: The same robot-group center distance does not automatically imply that the distance between the robot and individual participants is the same.

5.3.4 Discussion

We presented a study in which we collected data on a group of participants' motion behaviors when being approached by a robot from a variety of directions. We set out to investigate if people would appreciate robot approaches differently if participants were part of an open or closed circular formation, and found that the formation of the participants in this case did not influence the ratings of the participants. We further investigated approach direction by looking at the ratings of the three individual participants in the cases where the starting point of the robot's approach was behind the participants. We found a non-significant trend which indicates that participants rated those approaches as less comfortable, which is similar to related research by Karreman, Utama, Joosse, Lohse, van Dijk, and Evers (2014) and Ball et al. (2015). The third research question showed that the robot operator was perceived to become more skilled over time.

The setup of this study allowed us to collect data on the dynamic behavior of a small group when being approached by, and having a conversation with a robot. As a consequence the sample was limited to participants who could conveniently be recruited; in this case the sample consisted mainly of Western-European students. We will apply the insights gained in this study to inform the design of the next study where we investigated cultural preferences for human-robot approach distances- and directions.

5.4 Study 5: Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People

The goal of this study is to identify optimal approach and placement position for a robot which is seeking to gain the attention of a small group of people. As social robots are envisioned to operate in contexts in which they have to interact with people having different cultural backgrounds (such as airports and fairs), we are particularly interested in finding out if a robot requires different spatial behavior depending upon the cultural background of its users. Therefore we have conducted a study in which we distributed an online survey to three different cultural regions in the world through a crowdsourcing platform.

5.4.1 Hypotheses

Work on personal space in HRI has mostly focused on the personal space of individual people, and while numerous works call these zones “elliptical” (e.g. Ciolek, 1983; Rios-Martinez, Spalanzani, and Laugier, 2014), only one distance is commonly reported: the distance to the front of the participant. The diameter of the different zones can be estimated, but has not been researched extensively up till now.

Figure 5.10 contains two different F-formations: a circular formation with congruent angles between participants, and a circular formation with incongruent angles. There are three figures along a circle with a diameter of 122 cm (or 4 feet). The

Figure 5.10: A circular F-Formation with congruent (left) and incongruent (right) angles.

circles around the participants represent our hypothesized proxemics zones, these being the intimate zone, close personal and far personal space zone, respectively. The initial position where a person places him-/herself to join a group can be found more appropriate or inappropriate.

Based upon the proxemics theory, we hypothesize that participants will find the approach of a robot which stays out of their intimate zone more appropriate. Our first hypothesis is therefore:

- H1 People will rate an approach by a robot as more appropriate when the robot stays out of every group member's intimate space zone.

People often have preferences to join a group at a particular position where there is a person we know, or that seems otherwise appropriate. We are interested in small groups such as families (father, mother and child). It may, for instance, be seen as more appropriate to approach a group in between the mother and father as compared to in between the child and one of the parents, essentially cutting off a child from one of the parents. This leads us to the second hypothesis.

- H2 People will rate a robot approach as less appropriate when a robot approaches in between a child and a parent, as compared with approaching in between both parents.

Given that different cultures exist, and that research by Rehm et al. (2005) found that participants from high-contact cultures stand closer to a group of people compared with people from low-contact cultures, we hypothesize a similar cultural dependent preference will exist when a robot approaches.

- H3 People from a high-contact culture (China, Argentina) are more comfortable with a closer approach by a robot than participants from a low-contact culture (United States).

5.4.2 Methodology

We conducted a 3 (nationality - between-subjects) x 3 (position in the group - within-subjects) x 6 (distance from the center of the group - within-subjects) online study. Participants provided ratings of situations of a robot having approached a group of people and indicated the appropriateness of the situation. The ratings were collected twice in order to calculate test-retest reliability. After having provided these ratings, participants answered a number of questions in order to collect demographic information.

5.4.2.1 Materials

Participants were shown images of small groups of people and a robot, an example can be seen in Figure 5.11b. These groups were composed of three people: a man, a woman and a child. The survey consisted of an introduction that contained detailed instructions as well as a picture of the family (Figure 5.11a). Participants were asked to indicate how appropriate they believed the position of the robot was, imagining

(a) The height of the fictional family members were scaled to average international dimensions.

(b) Example top-down still as shown to participants.

Figure 5.11: Participants were provided with situations showing top-down view of a robot and a small family group. Participants were asked to indicate how appropriate they believed the position of the robot was, imagining that the position was the position after the robot had completed its approach.

that the position was the position after the robot had completed its approach. The position of the robot was manipulated two-fold within-subjects (see next section).

For the groups, a circular formation with congruent angles was chosen. We are aware of the fact that people will often stand in non-congruent angle formations, however, if we were to introduce a formation with non-congruent angles and/or people spaced differently we would introduce a multitude of factors that would be hard to control for and that would make the study overly complex. The diameter of the o-space was set to 122 cm, which corresponds to Hall's social space zone. The height of the participants was based upon average international height. The male was scaled to 178 cm, the female to 152 cm, and the child to 140 cm. The height of the robot was scaled to 140 cm, as can be seen in Figure 5.11a.

5.4.2.2 Independent variables

Two variables were manipulated within-subjects: approach position (the position between which family members the robot approached), and the approach distance of the robot. We refer to the combinations of position and angle as *scenes* (Figure 5.12).

For each of the three different approach positions, the robot was placed at six different distances, measured from the center of the circle. These distances were 20, 40, 60, 80, 100 and 120 centimeter. As a control method, participants were exposed to each scene twice. Thus: participants were asked to rate (2 ratings x 3 approach directions x 6 distances) = 36 scenes.

Circles 1 and 2 (20 and 40 cm) are within participants' intimate zone, circles 3 and 4 (60 and 80 cm) in the personal zone, and circles 5 and 6 (100 and 120 cm) lie in the social zone.

5.4.2.3 Dependent variables

The dependent variables were measured using a 98-item online questionnaire (Appendix F). The questionnaire was divided into three blocks: appropriateness rating of the robot-group scenes, questions regarding participants' cultural background and personality, and general demographic questions. Internal reliability of all validated scales used in the second block was assessed by calculating Cronbach's α .

In the first block, participants were asked to rate the 36 "robot-family" scenes that have been described in the previous section. To avoid ordering effects, the order of all scenes was randomized. Participants were provided with the instruction: *"The robot approached the family and has come to a halt between particular family members at a particular distance. Now it will interact with them"*, and asked to indicate on a 7-point Likert scale how appropriate the position of the robot was.

Four items were included in this block to measure how participants themselves would approach the family. Two items were included to check the approach position- and distance manipulation. Here participants were provided with statements such as *"the robot generally approached from the same direction"* and *"the robot generally approached from different directions"*. Participants were forced to choose which of the statements was true. A final item was included in which we asked participants if they could indicate where they thought the family they had seen in the situations originated from.

The second block of the questionnaire consisted of a series of validated scales measuring four dependent variables. An indication of whether participants were members of a high-contact or low-contact culture was assessed by measuring *closeness* as people from a high-contact culture have been found to sit significantly closer to

Figure 5.12: Participants standing in a circular F-Formation with a diameter of 122 cm. Dark grey indicates possible location of the robot. Grey: intimate zone, light grey: personal space zone.

each other compared with members from a low-contact culture (Sussman and Rosenfeld, 1982). Five items from the IPROX (iconic proximity) questionnaire were used (Høgh-Olesen, 2008) to measure whether people belonged to a *high-contact* or *low-contact* culture ($\alpha=0.83$). *Attitude towards robots* was measured by the Negative Attitude Towards Robots (NARS) scale, measuring participants' negative attitude towards interaction with robots (6 items, $\alpha=0.69$), towards the social influence of robots (5 items, $\alpha=0.61$) and towards emotional interactions with robots (3 items, $\alpha=0.69$). The final construct in this block was *personality* as we figured this could influence people's preference for a robot position (e.g., more extrovert people preferring the robot to come closer or to approach at their side of the group). We measured the Big Five personality traits extraversion ($\alpha=0.73$), agreeableness ($\alpha=0.69$), conscientiousness ($\alpha=0.65$), neuroticism ($\alpha=0.614$) and openness to experience ($\alpha=0.75$) using the 20-item Mini-IPIP scale (Donnellan, Oswald, Baird, and Lucas, 2006). Hofstede and Hofstede (2001) identified five dimensions of culture, one of these being Individualism-Collectivism. One way to explain cultural differences is by measuring *individual and group self-representations*. Individual self-representations refer to whether the self is represented as "a separate, unique individual" whereas group-self representation refers to one who is "an interchangeable part of a larger social entity" (Brewer and Chen, 2007). Individualism and collectivism were measured with respectively three and four items, unfortunately the internal reliability of both measures was too low to be used in the analysis.

The final block consisted of 13 questions, and included *demographic* questions such as gender, age, nationality, and level of education. Social-demographic questions like nationality of ancestors, marital status and number of children were also included.

5.4.2.4 Data analysis

The results presented in this paper focus on the ratings of the scenes and on the closeness scale (five items from the IPROX questionnaire, see Dependent variables). Internal reliability of all scales was assessed by calculating Chronbach's Alpha (α), and deemed acceptable for all scales. As stated in the previous section, the participants rated all scenes twice as form of control method. After having excluded participants these ratings were averaged per participant and scene.

5.4.2.5 Participants

Participants were recruited from three different countries: China, Argentina and the United States. People from these three countries are generally considered culturally different; not only because they are geographically on different continents, but also because various studies have shown cultural differences (Gupta et al., 2002; Hall, 1966; Remland, Jones, and Brinkman, 1995) with respect to for instance societal values.

Participants were recruited through the CrowdFlower platform. 244 participants completed the questionnaire; each being paid \$1 for completion of the survey. Responses were limited to one per IP address. Participants who failed to correctly an-

swer the two manipulation checks were excluded from the sample. A second control method was the analysis of the robot-scene questions, which were 18 situations rated twice by each participant on a 7-point Likert scale. Participants who rated four or more situations with a difference of 3 or more points were also excluded from the survey. In total 63 participants (26%) were excluded. After applying the exclusion criteria, the total sample contained 181 participants, as specified in Table 5.2. Participants were aged between 16 and 73 ($M=37.50$, $SD=12.54$). The occupations of the participants varied: 23.2% indicated they were currently unemployed, 14.4% of the participants were students and 5.5% indicated they were retired. The remaining 56.9% held jobs in areas such as education (7.7%), retail (5.5%), arts (5.5%) and government (5.0%).

	N	mean age (SD)	male / female
United States	86	43.27 (12.25)	26 / 60
China	29	30.48 (8.93)	19 / 10
Argentina	66	33.06 (10.90)	48 / 18
Total	181	37.50 (12.54)	93 / 88

Table 5.2: The sample consisted of 181 participants.

5.4.3 Results

In this section we present the results of the survey that we acquired from the analysis of the ratings of the scenes and the closeness scale.

To determine whether the participants found an approach more appropriate if the robot stayed out of every group member's personal space zone (H1), we conducted a repeated-measures ANOVA with one within-subjects variable (being intimate- or personal space zone), and two between-subject factors (nationality and gender). For the purpose of analysis of this hypothesis, ratings for circles 1 and 2 (intimate space zone) were averaged as well as the ratings of circles 3, 4, 5 and 6 (outside intimate space zone).

To analyze whether an approach between the child and one of the parents was rated as being less appropriate (H2) and whether participants from higher contact cultures were more comfortable with a closer approach (H3), we conducted factorial repeated-measures ANOVAs with two factors as within-subjects variables; these being the average ratings of the position in the group (3) and distance from the group (6). Nationality and gender were used as between-subject factors.

5.4.3.1 Participants prefer a robot that stays out of our intimate space zone

In H1 we hypothesized that participants would rate it as more appropriate if the robot was positioned out of every group member's intimate space zone. A repeated mixed-model ANOVA revealed that participants rated the robot positions in the intimate space zone as significantly less appropriate ($M=3.14$, $SD=1.25$) compared with those positions where the robot was positioned outside the intimate space zone ($M=4.61$, $SD=0.99$), $F(1, 100.658)=109.57$, $p<.01$. We therefore accept H1: a robot which stays out of the intimate space zone of each of the group members is considered to

be more appropriate. These ratings were neither affected by gender ($p=.87$) nor by nationality ($p=.60$).

5.4.3.2 Appropriateness of a robot's approach is not always affected by its position relative to the family members.

A factorial repeated-measures ANOVA with two independent variables (distance and position) and two between-subjects factors (gender and nationality) was conducted. Mauchly's test indicated that the assumption of sphericity had been violated for the main effects of distance, $\chi^2(14)=613.9$, $p<.001$, and angle, $\chi^2(2)=76.37$, $p<.001$. Sphericity had also been violated for the interaction effect (distance*direction), $\chi^2(54)=183.55$, $p<.001$. The degrees of freedom were corrected using Greenhouse-Geisser estimates of sphericity ($\epsilon=0.42$ and $\epsilon=0.74$ for the main effects, and $\epsilon=0.81$ for the interaction effect).

There was a significant main effect of the approach distance ($F(2.09, 365.19)=54.37$, $p<.001$), and a non-significant effect of approach angle on the appropriateness of the robot's position ($F(1.47, 258.25)=2.6$, $p=.075$). Post-hoc contrasts revealed a significant difference of appropriateness between the "Mother/Child" and "Child/Father" approaches: the appropriateness of the "Mother/Child" ($M=4.16$, $SD=0.10$) approach was significantly higher compared with the "Child/Father" approaches ($M=3.93$, $SD=0.10$), $F(1, 175)=11.71$, $p<.001$. The appropriateness of the "Father/Mother" approaches ($M=4.11$, $SD=.095$) were equally appropriate as the "Mother/Father" approach, we therefore only partially accept H2, in which we hypothesized that participants would rate a robot approach as less appropriate if a robot approached in between a child and a parent, as compared with approaching in between both parents. Instead, participants indeed found an approach between parent and child less appropriate but only for the position between father and child. The most appropriate approach position was generally thought to be in between the mother and the child (see Table 5.3).

5.4.3.3 Influence of cultural background on appropriateness

To check whether the countries that we chose actually differed in the low-high contact dimension, we analyzed the items from the closeness questionnaire. There was a significant difference between the ratings, $F(2)=15.53$, $p<.001$. As can be seen in Figure 5.13, participants from the United States gave significantly higher ratings on the closeness measure ($M=4.96$, $SD=1.05$), which indicates they put more distance between themselves and other people. This effect was vice-versa for Chinese people, as expected ($M=3.88$, $SD=1.20$). The Argentinian participants rated in between ($M=4.11$, $SD=1.19$). Therefore, we can assume that the national groups included in this sample can indeed be considered to have different cultural backgrounds concerning the low-high contact dimension.

Our third hypothesis was that participants from high-contact cultures (such as China and Argentina) would rate a close approach as more comfortable than participants from a low-contact culture (United States). There was a significant three-way interaction effect between the nationality of the participant, distance, and position

Figure 5.13: Participants from what are considered low-contact cultures scored indeed significantly higher on the “closeness” construct. *** indicates significant difference at $p < .001$. Mean scores provided in bars.

Figure 5.14: Mean appropriateness ratings for the Chinese, U.S. and Argentinian sample. Appropriateness was measured on a scale from 1 to 7.

of the robot on appropriateness of the scene, $F(16.20, 1417.24) = 1.92, p < .05$. This effect can be seen in Table 5.3 and Figure 5.14.

United States and Argentinian participants gave similar appropriateness ratings for the approach distances, but that one particular approach in between the “Father/Mother” was considered most appropriate by the United States participants, whereas the Argentinian believed the “Mother/Child” position was more appropriate. The Chinese participants’ ratings were generally lower and a notable difference was that the closer approaches (within the intimate zone) were actually considered to be quite appropriate. Like the U.S. and Argentinian participants, the Chinese also had a preference for a further stop distance (80-100 cm), though this difference was much less pronounced.

We therefore partially accept H3. We hypothesized that participants from high-contact cultures (such as China and Argentina) would rate a close approach as more comfortable than participants from low-contact cultures. Chinese participants saw a closer approach as more appropriate. However, we expected similar results for Argentinians, which we did not find. Interestingly enough the ratings the Argentinians provided were quite similar to those provided by the U.S. participants. We will reflect on this in the discussion.

Distance	China			United States			Argentina

20 cm	3.05 (1.56)	3.86 (1.87)	3.74 (1.75)	2.72 (1.56)	3.02 (1.63)	2.72 (1.52)	2.70 (1.54)	2.91 (1.45)	2.53 (1.36)
40 cm	3.35 (1.75)	4.02 (1.74)	3.31 (1.79)	3.55 (1.69)	3.84 (1.68)	3.41 (1.47)	3.48 (1.52)	3.61 (1.23)	3.04 (1.18)
60 cm	3.86 (1.74)	4.16 (1.90)	3.76 (1.74)	4.31 (1.48)	4.55 (1.60)	4.13 (1.47)	4.08 (1.58)	4.44 (1.38)	3.69 (1.20)
80 cm	4.26 (1.77)	4.47 (1.73)	4.41 (1.55)	5.05 (1.49)	4.99 (1.54)	4.71 (1.52)	4.80 (1.48)	5.00 (1.55)	4.22 (1.36)
100 cm	4.52 (1.83)	3.67 (1.65)	4.41 (1.75)	5.34 (1.43)	4.98 (1.44)	4.99 (1.61)	5.14 (1.42)	5.22 (1.47)	4.64 (1.32)
120 cm	3.72 (1.89)	3.10 (1.75)	3.60 (1.66)	4.91 (1.72)	4.69 (1.74)	4.84 (1.59)	4.92 (1.69)	4.84 (1.81)	4.69 (1.58)

Table 5.3: Mean appropriateness ratings for the Chinese, U.S. and Argentinian sample. Distance indicates distance between the center of the circle and the robot (in cm). Mean appropriateness ratings on a scale from 1 to 7, standard deviations between brackets.

5.4.4 Discussion

In this section we presented the methodology and results of a survey investigating cross-cultural HRI proxemics preferences. Our results show that there are indeed cultural differences in spatial behaviors in HRI. Thus, taking culture into account is an important next step for HRI if social robots are designed to operate all over the world in various cultural contexts or in environments where people from different cultures are around (such as airports, fairs and museums). We will now discuss both the results and the methodology to retrieve directions for future research.

5.4.4.1 Implications for theory

We hypothesized that participants would find approaches in between the parents more appropriate compared with the approaches where a child is cut off from one of the parents. The reason for the unexpected finding that approaches between mother and child were found quite appropriate could be a pragmatic one, which we had not considered. By approaching in between the mother and child the robot directly faced the father of the family. It could be that a robot's frontal approach to a male is seen as more appropriate. Even though previous work by Walters et al. (2007) did not confirm this notion, our results could warrant a further investigation into differences in gender preferences and stereotypes. A limitation of this work is the skewed sample, both in terms of gender and cultural background. Both female and participants from the United States are over-represented. This warrants caution when interpreting the results of this study. In future work, it is recommended to find alternative ways of recruiting participants. One solution may be to recruit students from different countries. Although this presents a new bias in terms of generalizability (as previously discussed in Section ??), it would provide for a sample which could be comparable in terms of such as life experience and common interests.

Figure 5.14 shows similarities in the appropriateness ratings of the U.S. and Ar-

gentinian samples respectively, despite the fact that Argentinian's closeness scores indicate a higher-contact culture. Therefore, we expected they would find it more appropriate if the robot approached closer. Thus, it could be that the high-low contact culture dimension is too simple and did not completely capture the subtleties of high-low contact cultural backgrounds and that there are more factors at play. One possible explanation can be found in the work of Hofstede and Hofstede (2001, p. 90). On the Individualism dimension, the U.S. scores high (91 points), and China scores relatively low (20). Argentinians scores 46 points. This is still closer to China than the U.S., however, if we look at other Latin American countries, such as Ecuador (8), Venezuela (12), Colombia (13) and Chile (23) it appears to be that Argentina is a rather individualistic country. This might partly explain why Argentinian participants showed a preference for a further positioning of the robot. However, this issue deserves further investigation.

5.4.4.2 Implications for methodology

To analyze cross-cultural differences in HRI proxemics, we used an online questionnaire as this allowed us to distribute the survey to geographically dispersed samples. The survey contained static images, and while the results do support most of our hypotheses, the ecological validity of our research is limited because groups are dynamic entities. The formation of the group changes when a new member joins the group, and our images might very well not have been able to capture these subtle dynamics.

Furthermore, participants viewed the robot-group scenes from above. This may - unintentionally - have caused a limitation as participants were not able to take the height of the actors into account. In retrospect it is possible that participants would provide different ratings had they been provided with different camera angles next to the top-view.

In social psychology methods other than real-world experiments have been used to investigate the size of a person's personal space. These measures, called *projective measures*, have advantages over real-world experiments which we made use of in this study as well. One example is the ease at which data can be collected. Studies validating these projective measures in social psychology yielded different results, varying from negative (Love and Aiello, 1980) to positive (Greenberg, Strube, and Myers, 1980; Knowles, 1980) as we will show in Chapter 6.

5.5 Chapter conclusion

We have presented two studies investigating how a robot should approach a small group in terms of direction and distance. The secondary goal of this chapter was to investigate whether cultural differences exist for this behavior. These studies contribute to two research questions:

RQ-IIa How should a robot approach a small group of people, in terms of direction and distance?

RQ-IIb Do people from different cultures prefer a robot to approach them differently?

In the first study we found that the spatial arrangement of a group did not influence the approach direction of a person present through a telepresence robot: rather than choosing a position with most space available, robot operators chose to enter at the nearest possible place, during and after which the group members adapted their position. This should be interpreted in the context of the study: the robot was a telepresence robot, through which a group member was present. Indeed this member had crucial information for the completion of the task. This could have made the group members physically present more inclined to allow the telepresence robot operator to join the formation in the easiest way possible. In the second study we investigated cultural differences in preferences for human-robot approach distances. We collected data through a survey and in support of RQ-IIb, we found indications of national cultural differences in human-robot approach distances. In particular participants from China were more positive towards close approaches as compared to participants from the United States and Argentina. The conclusions of the second study rely on the assumption that the method we used - a survey - is a valid measurement. Related work on measurement methods for human-human personal space measurement investigated the validity of similar measures. We will review this literature in Chapter 6 and present a study which investigates the validity of projective measures to investigate human-robot personal space.

To return to the research question of this chapter. Based on the results of the two studies presented in this chapter, we conclude that a robot should approach a group - that is, enter the P-space - at the earliest opportunity. With this we mean that a robot should not drive around the group in order to enter the P-space at a place where there is space to physically enter. We observed that usually the group would make space for the robot to enter, and adjusted its location. We found support for the secondary goal of this chapter: there appears to be a difference in preferences for human-robot approach distances (and directions) among people of different nationalities. In particular Chinese participants were more comfortable with close approaches as compared to participants from the United States and Argentina.

6

Measuring Patterns in Personal Space with Projective Measures

"He plays with toys."
"It's a model"
"With little figures..."

Worf, Ezri Dax.

*Star Trek Deep Space Nine: The Changing Face of
Evil*

In this chapter we will present a study in which we investigated whether we can reliably use projective measures to discover patterns in human-robot proxemics. We will first provide an overview of previous studies in psychology which examined the validity of projective measures. Following this overview we will describe how we developed the projective measures used in this study.

6.1 Introduction

In the previous chapter we described how we distributed an online survey to participants in different countries, with which we investigated cultural differences in people's perception of appropriate approach distances for robots. We utilized the advantage of an online survey (ease of access to a diverse participant pool) even though studies in psychology reported results both in favor and against so-called projective measures. With projective measures we refer to those measurement techniques which require a participants to imagine a certain situation and to project how they would behave (Jones and Aiello, 1979). In this chapter we aim at answering RQ-III: "Can responses to robot proxemics behavior be reliably measured through projective measures?"

With computer technology, including virtual reality, rapidly advancing, a question arises whether people would behave in a similar way to a real robot as to a virtual (representation of a) robot. Advantages of projective measures include reduced costs

(Kamide et al., 2014), the inclusion of larger samples (Woods et al., 2006a) and more diverse samples, for example in terms of cultural background (Henrich et al., 2010). Therefore the goal of this chapter is to discover whether we can use projective measures to reliably find patterns in human-robot proxemics. As research in social psychology has investigated the validity of projective measures to measure personal space we will first review this literature (Section 6.2). Following this review we will describe the design and first evaluation of two projective measures (Section 6.3) which we subsequently used in an experiment to validate the use of projective measures to discover patterns in human-robot proxemics (Section 6.4).

The research reported in this chapter contributes to the existing work in HRI that investigated measures to investigate proxemics behaviors in HRI (Kamide et al., 2014; Ohara et al., 2009), and related behaviors such as approach directions which were evaluated using video-based HRI studies (Woods et al., 2006a).

Figure 6.1: Overview of the layout of chapter 6.

6.2 Theoretical Background

Previous research in social psychology investigated the applicability of measures other than live experiments to investigate approach distances by people. The goal of the study reported in Section 6.4 is to investigate the applicability of projective measurement techniques in HRI in comparison to live measurement. In this section we will provide an overview of projective measurement techniques previously used in social psychology as well as studies comparing the different projective measurement techniques.

6.2.1 Techniques to measure personal space

In the terms of Hall's (1966) book the personal space zone is one of four (elliptical) zones around the body, and each zone suitable for different kinds of social interactions. Hall provided approximate sizes of the various proxemics zones, with the note that these were based upon observations, and under influence of factors such as for instance culture. In the years since publication of Hall's work hundreds of studies have investigated these factors. In a similar fashion, a scala of different methods have been used, and compared with one another also yielding inconsistent results. These methods can roughly be divided into three categories, these being live measurements, estimation and projective measures (Forston, 1975). In recent years, technology using virtual reality environments has also been used to investigate the use of space by virtual human agents (e.g. Bailenson, Blascovich, Beall, and Loomis, 2001). For

the sake of clarity we will use the set of umbrella terms by Forston (1975), thus live measurement, estimation and projective measures.

Live measurement techniques involve one or multiple participants and/or confederates walking up to each other until one of them says “stop”, or indicates the current distance is the minimum comfortable distance, after which the experimenter measures the distance between the two. Live measurement experiments include both informed and disguised measurement (where participants are unaware that they are being tested). Usually this disguised measurement involved choosing one of a number of chairs in a room such as in the experiments by Pedersen (1973), Knowles (1980), Love and Aiello (1980) and Greenberg et al. (1980). Forston (1975) defines estimation techniques as those in which trained observers observe spatial distances in a real world. Projective techniques involve various measures but have as a common feature that the subjects are supposed to imagine they are in the place of a doll, a stick figure or another scaled object which they have to place in a (simulated) environment (Jones and Aiello, 1979), examples include the studies by Tolor and Salafia (1971) and Walkey and Gilmour (1979).

6.2.2 A comparison of personal space measurement techniques

We will provide a brief overview of papers in which different measurement techniques (e.g. real-life measures vs. stick figure placement) are compared to each other in order to find out how valid measures other than real life measures are. Table 6.1 shows this comparison. In those cases where a within-subject design was used, authors have reported that the order of the different measurement methods was counterbalanced. Methods have been grouped in three categories; these being live measurement, estimation and projective measurement.

Different types of projective measures can be distinguished. However, all of these satisfy the requirement that “they require mental projections from the spacings subjects predict they would adopt” (Hayduk, 1983). Examples of projective measures are:

1. 2D Silhouette placement (also known as felt-figure placement), for example the studies by Little (1965), Tolor and Salafia (1971) and Lomranz (1976).
2. 3D Doll placement, used in a cultural context by Little (1968).
3. Paper-and-pencil studies; such as the Comfortable Interpersonal Distance measures (Duke and Nowicki, 1972), Pedersen Personal Space measure (Pedersen, 1973) and the IPROX questionnaire (Høgh-Olesen, 2008).
4. Video studies (Walkey and Gilmour, 1979), where participants had to imagine they were the person being approached in the video, and rate their level of comfortableness.

Dosey and Meisels (1969) compared live measurement, silhouette placement and chair placement measures and found no consistent correlation between the three measures; there were slight positive correlations between live approach and the silhouette and chair placement measures though these correlations were not significant.

#	Study	Design	N	LM	E	P	Specific measures	Result
1	Dosey and Meisels (1969)	Within-subjects	186	x		x	(1) Disguised approach, (2) Silhouette placement [P], (3) Chair placement, and (4) questionnaire	No consistent correlation between measures
2	Pedersen (1973)	Within-subjects	170	x		x	(1) Chair placement task, (2+3) Pedersen Personal Space Measure; retested after 1 hour [P], (4) Direct measurement task	No correlation between Pedersen measure and other measures.
3	Haase and Markey (1973)	Within-subjects	36	x	x	x	(1) stop-distance task [LM], (2) ranking of observation of actors (live) [E], (3) 2D silhouette placement task [P], (4) judgment of photographs [P]	Observation of actors and felt-board techniques were the most representative of actual behavior.
4	Walkey and Gilmour (1979)	Within-subjects	50	x		x	(1) Newly tested video-IPD [P], (2) Pedersen personal space measure [P], (3) Rawls personal space measure [P], (4) Comfortable Interaction Distance Scale [P], (5) Stop-distance task [P]	Newly tested video-IPD was better approximation of stop-distance measurement compared with the other 3 methods.
5	Greenberg et al. (1980)	Within-subjects	50	x		x	(1) Stop-distance measure [LM], (2) 2D silhouette placement [P], (3) Comfortable Interaction Distance Scale [P], (4) disguised approach measure [LM].	Silhouette placement correlated significantly with the stop-distance measure.
6	Love and Aiello (1980)	Within-subjects	30 pairs	x		x	(1) Disguised stop-distance measure [LM], (2) 2D silhouette placement [P], (3) 3D doll placement [P], (4) unspecified projective measures [P]	No significant correlation between projective measures and stop-distance measurement.
7	Knowles (1980)	Within-subjects	91	x		x	(1) Disguised stop-distance measure [LM], (2) figure placement [P], (3) Comfortable Interaction Distance Scale [P]	All measures showed similar results for acquaintance: friends placed closer than strangers.

Table 6.1: Overview of seven studies comparing projective measures [P] with live measurement methods [LM] and estimation measures [E].

Pedersen (1973) also compared stop-distance measures with a chair placement task¹. In addition Pedersen developed a set of four 2D silhouettes from different positions (facing left, right, front and top-down), which participants had to move over a line towards another profile representing "another male approximately your same age" - called "Pedersen personal space measure". This measure did not correlate significantly with the other measures. Love and Aiello (1980) compared distance between pairs when having a discussion with an unspecified projective measure, 2D and 3D silhouette placement and found no significant correlations between the three measures. The lack of significant correlations would indicate that the projective measures are not related to the actual behavior of the participants when they had (in this case) a conversation with each other.

On the other hand there is research showing positive correlations between these measures; which would indicate that the projective measures are to an extent related to the actual distance behavior (Greenberg et al., 1980). Examples of these include the work by Little (1965) who found a positive correlation between participants' placement of live actors and 2D silhouettes. Haase and Markey (1973) conducted a similar study in which participants participated in a stop-distance task, placed 2D silhouettes, and ranked different positions of (live) actors and photographs of seated participants in order of preference. The observation of live actors ($r=0.75$, $p<0.01$) and 2D silhouette placement ($r=0.30$, $p<0.01$) correlated the highest with the actual stop-distance task. Greenberg et al. (1980) found that silhouette placement was significantly correlated with stop distance ($r=0.33$). Contrary to previous research (e.g. Dosey and Meisels (1969); Pedersen (1973)) significant correlations were found for the CID measure (a pen-and-paper measurement) and silhouette placement ($r=0.43$) and the CID measure and stop distance ($r=0.52$). Finally, Knowles (1980) conducted an experiment comparing various measures and found that though the correlations between the measures were low, there were significant results for acquaintance: with every measure participants placed representations of themselves closer to "friends" than to "strangers".

We reviewed the results of seven studies in which live measurements were compared to various types of projective measures, with the goal of investigating to which extended these projective measures relate to the behavior participants displayed during the live measurements. At the onset there seems to be little grounds to warrant an investigation of the validity of projective measures in HRI. However, we believe it could be interesting to investigate these measures. The lack of (significant positive) correlations could be seen as proof that projective measures have no to little external validity (Sundstrom and Altman, 1976). In favor of projective measures, Knowles (1980, p.241) argues that the measure used (correlations) does not assess whether or not the findings generalize, but rather if the position in the sample remains consistent (far or close). Furthermore, Haase and Markey (1973) suggest that the actual participation (e.g. moving figures) is one of the features due to which these measures correlate. Along the same line of reasoning Walkey and Gilmour (1979) found that

¹In a chair placement task participants typically get to choose on one of multiple empty chairs - the chairs are placed at different distances from the experimenter, which thereby serves as a measure if participants would take a place at a chair further away or closer to the experimenter (Sommer, 1959, p.248-251).

the more features measures share, the higher they correlate.

Rather than investigating whether projective measures would yield the exact same results, we would like to investigate if it is possible to use projective measures as a reliable source to discover patterns (or trends) in proxemics, for example in the same way of Knowles (1980) who found that over different measurement methods friends were placement consistently closer than strangers. Therefore the research question of this chapter has been formulated as follows:

“Can responses to robot proxemics behavior be reliably measured through projective measures?”

To this end we conducted Study 6. We will report on this study in Section 6.4 after presenting the design of the two projective measures in the next section.

6.3 The development of two projective personal space measures

Based upon the conclusion of the literature review in the previous section we went forward to design an experiment with which to compare projective measures to investigate patterns in personal space behavior in HRI. For organizational and logistical reasons it was not possible to use the SPENCER robot platform, as used in the subsequent studies. We used a 163 cm. high Giraff telepresence robot (Figure 6.2), which is smaller, both in terms of height and total volume, than the SPENCER robot. We developed two projective measures, the first being a scaled 3D figure (inspired by the research of Little (1968)), the second measure a 2D side-view of a robot and person, which could be manipulated using a slider bar. In this section we will describe the design process and detail the specifications of these measures.

Figure 6.2: We used a 163 cm. high Giraff robot for the experiment. The projective measures were based upon this robot.

6.3.1 3D robot

We created a 12.2 cm 3D version of the Giraff robot; 7.5% of the size of the real robot. A high polygon mesh of the original robot was adapted for 3D printing using Google Sketchup; this process involved among other things removal of the wheels and merging of faces which would not be visible on the printed model (Figure 6.3a). In order to reduce the amount of support material required, it was decided to separate the robot's base and head during printing. A polylactic acid prototype was printed on

an Ultimaker2 3D printer². Minor modifications were made to the prototype, which was followed by printing the final version which consisted of two parts: a base and a neck/head (Figure 6.3b and 6.3c). After printing, a number of layers of paint were applied to the model. We glued the two parts together to create the final model (Figure 6.3d).

Figure 6.3: Design process of the 3D robot in four steps.

As well as the robot we printed a 13.2 cm tall figure which was used to represent the participant. This figure was designed to be a neutral representation of a general person.

6.3.2 2D robot

We used the previously-made 3D model of the Giraff robot and participant figure to create a series of 500 2D situations of the robot and participant. These situations were presented using a slider bar on a web-page, administered on a laptop. The web-page was made in such a way that with every step the person moved 1 centimeter closer to the robot; therefore allowing the participants to move the figure at a distance from anywhere between 500 and 0 centimeters from the robot. The distance data was automatically stored on a server.

Figure 6.4: The 2D projective measure consisted of 500 2D situations of the robot and participant; participants could use the slider bar to adjust the position of the robot.

²Ultimaker, <https://ultimaker.com/en/products/materials>. Last accessed 21 March 2017.

6.4 Study 6: Projective Measures to Assess Patterns in Personal Space

In Section 6.2 we reviewed literature on the validity of projective measures to assess personal space in human-human interaction, and we drew the conclusion that even though there is evidence either in favor of or against projective measures, a small majority of studies argues in favor of these measures. In this section we report on an experiment we conducted to investigate the reliability of projective measures as a method to discover patterns in HRI proxemics (comfortable approach distance between robots and people) using the projective measures previously developed in Section 6.3. We set out to answer the research question “To what extent can projective measures be used to investigate proxemics patterns in HRI?”.

6.4.1 Motivation

We will motivate why we operationalize “pattern” by investigating whether people interacting with an in-group robot position themselves further away from a robot as compared to people who interact with an out-group robot (or vice-versa).

Previous work on personal space in psychology and HRI (See Section 2.4.1) identified a number of confounding variables which could potentially influence people’s personal space preferences with the robot in question. Factors relevant for the current experiment include gender (Takayama and Pantofaru, 2009; Mumm and Mutlu, 2011), experience with robots (Takayama and Pantofaru, 2009), cultural background (Eresha et al., 2013) and previous trial experience (Walters et al., 2008b). The post-experiment questionnaire (Section 6.4.3.3) contains measures for these factors. In his review on factors influencing personal space, Aiello (1987) concluded that people have a preference for shorter distances between themselves and people they like. The potential influence of this factor on human-robot personal space has not been investigated so far.

The in-group favoritism hypothesis holds that people judge in-group members as being more positive than out-group members on perceptual, behavioral and attitudinal measures (Tajfel, 1970). Research has shown that people distance themselves differently based on group membership while being seated. Tedesco and Fromme (1974) conducted an experiment where 18 participants played a game with a confederate. Post-hoc seating measures indicated that participants sat closer to the cooperative player, than to the competitive player. Glick, DeMorest, and Hotze (1988) found that (seated) people being approached by a close in-group member were more willing to comply with requests than out-group members; though this difference disappeared when positioned far away. Glick et al. (1988) suggested that the out-group members caused anxiety, as they would have required a larger personal space, and therefore participants complied less with their requests. Similarly, Novelli, Drury, and Reicher (2010) showed in two experiments that (seated) participants expecting to interact with an in-group member required less personal space than those who expected to interact with an out-group member.

6.4.2 Hypotheses

In Section 6.2 we reviewed work in psychology comparing various measures to assess personal space yielded different results; some experiments found (positive) correlations between the measures, while others did not. In this study we will compare three measurement methods: people's behavioral responses to a physically embodied robot (*real*) versus the two projective measures we developed in 6.3, the *2D robot* and *3D robot*.

In line with our research question and the research of Greenberg et al. (1980), Knowles (1980) and Walkey and Gilmour (1979), we hypothesize that participants' responses over the three measurement methods are consistent (show convergent validity) as we assume they measure the same underlying variable.

H1 Participants provide a consistent response to (a) the 3D figure placement measure and (b) the 2D slider measure compared to the real measurement.

From the literature we also know that the more components live measurement and a projective measurement have in common, that projective measurement becomes a better predictor of actual behavior; as compared to a projective measurement which does not have that many components in common with live measurement (Haase and Markey, 1973).

H2 The more components a projective measurement method has in common with real measurement, the better it will predict participants' behavioral response in real life.

Based upon the research by Novelli et al. (2010) we predict people interacting with an in-group robot (or identifying themselves as such) will position themselves closer to a robot compared to people interacting with an out-group robot.

H3 People interacting with an in-group robot will position themselves closer to the robot as compared to people interacting with an out-group robot.

6.4.3 Methodology

In order to test the hypotheses we have conducted a 2 (in/out-group) * 3 (measurement method) mixed-method experiment. In/out-group membership was manipulated between-subjects, and the measurement method within-subjects in a counter-balanced order. The three measurement methods were administered twice in order to calculate test-retest reliability. The measurement methods were first administered once, after which the experimenter repeated the administration of these measures in the same order.

6.4.3.1 Participants

Participants were recruited from the premises of the University of Twente. Data from 68 out of 70 participants was included in the analysis; one participant was excluded

due to technical failures of the recording equipment, data from a second participant was classified as being an outlier, and therefore removed from the dataset.

Participants were aged between 19 and 59 ($M=24.81$, $SD= 6.41$); 67.6% of the sample was male, 32.4% female. Most participants had the Dutch nationality (57.4%), followed by German (8.8%) and Mexican nationalities (7.4%). Most of the participants described their current occupation as being a student (79.4%) or working at a university-level education facility (10.3%). Most participants were already in the possession of a bachelor's (52.9%) or master's degree (16.2%). 23.5% of the participants indicated the highest level of education they had completed was high school.

In general, participants were considered to be experienced with robots: 35.3% of the participants indicated they had seen robots before and 54.5% had worked with or built robots themselves. 41.2% of the participants had seen the particular robot platform before; we have tested whether this influenced participants' distancing behaviors: this was not the case.

6.4.3.2 Independent variables

Participants approached a static robot, and were either part of the same group (with the robot), or a different group. Before the start of the experiment, participants were informed they would be assigned to a group based upon their answer to a 6-item questionnaire. This questionnaire contained six questions of the "tendency to form parasocial relationships"-questionnaire used by Lee et al. (2006b). In reality the answers did not contribute to the assignment of either group; nor did the questionnaire reveal any information which could have influenced the participants. Participants wore a cap with the respective color of the group; this was a blue cap for in-group members and a white cap for out-group members.

Participants placed (representations of) themselves in relation to the robot. In the experiment we used three different representations:

- Physically approach a real robot (Figure 6.5a).
- Place a scaled figure representing him/herself in front of a 3D robot (Figure 6.5b).
- Place a scaled figure representing him/herself in relation to a 2D robot (Figure 6.5c).

In order to make sure that participants approached from the front, we taped a line on the floor of the experiment room, and in the other stimuli, and instructed participants to *"Please move over the black line towards the robot, as far as you would feel comfortable if you were to have a conversation with the robot"*. The other stimuli were designed to reflect the real situation as closely as possible; that is to say the background color was similar to that of the floor in the lab and all three measurement methods featured the same black line.

6.4.3.3 Dependent variables

Previous work on personal space in psychology and HRI identified a number of confounding variables which could potentially influence people's personal space prefer-

ences with the robot in question. Relevant factors for the current experiment include gender (Takayama and Pantofaru, 2009; Mumm and Mutlu, 2011), experience with robots (Takayama and Pantofaru, 2009), cultural background (Eresha et al., 2013) and previous trial experience (Walters et al., 2008b). We have included measures to assess these factors in the post-experiment questionnaire.

The dependent variables included the distance between the robot and participants. After having completed the six distancing tasks, participants were provided with a post-experiment questionnaire (available in Appendix G). This 42-item questionnaire measured *attitude towards robots* based upon three subscales (Nomura et al., 2008): negative attitude towards interaction with robots (5 items, $\alpha=0.67$)³, negative attitude towards the social influence of robots (5 items, $\alpha=0.51$) and negative attitude towards emotional interactions with robots (3 items, $\alpha=0.41$). The internal reliability of these measures is not high.

We also measured *inclusion of others in self* (1 item, Aron et al., 1991), *closeness* (5 items, Høgh-Olesen, 2008, $\alpha=0.73$) and individualistic (3 items, $\alpha=0.47$) and collectivistic (4 items, $\alpha=0.54$) *group representation* (Brewer and Chen, 2007). Manipulation checks (2 items) and a control question (1 item) were included. Finally we asked participants about their *opinion of the task* (2 items), *experience with robots* (2 items) and *demographic information* such as gender, age and nationality (8 items).

6.4.3.4 Materials

For this study we used a 163 cm tall Giraff telepresence robot (Figure 6.5). In order to create three measurement methods with identical robots the robot was static and powered off. Instead, a set of static eyes were placed on the robot.

Figure 6.5: Approach distance was measured using three measures: approach distance to a real robot, placement of a 3D figure and placement of a 2D figure.

The 3D robot placement task was on a laminated piece of dark paper, on which a 23 cm black line was printed (7.5% of the length of the real line). The distance was measured by putting a line at the base of the feet of the participant figure, and measuring afterwards with a ruler. The 2D robot placement task was administered on a laptop.

³Item “The word ‘robot’ means nothing to me” was removed due to low internal reliability.

6.4.3.5 Procedure

The experiment was conducted in a laboratory of the University of Twente (Figure 6.6). Participants were welcomed by the experimenter and informed of the general procedure of the experiment. Following signing of the consent form, participants received the first questionnaire, after which they were assigned into the white or blue team. Participants were provided with a cap in the color of their team, which also contained a tracker for the OptiTrack system.

Figure 6.6: Layout of the experiment lab.

The experimenter then continued the experiment by introducing the participants to the robot, and told the participants the robot was part of the blue team. To participants who were part of the blue team the in-groupness was strengthened by adding that “the robot is thus a teammate”. Following this instruction, participants completed the measurement methods in a counterbalanced order. With the exception of the physical approach task the measures were arranged in such a way that the participant could complete them by walking from the left to right through the experiment room (Figure 6.6b). After completing the tasks for the second time, participants were invited to complete a post-experiment questionnaire. After having completed the post-experiment questionnaire participants were debriefed, and any remaining questions were answered. The total duration of the experiment was approximately 20 minutes.

6.4.3.6 Data analysis

The manipulation checks and control questions showed participants correctly identified themselves with their respective team, and all identified the robot as being part of the blue team. Results of the *inclusion of others in self* question showed that participants who were part of the in-group indicated they felt “closer” towards the robot ($M=3.27$, $SD=1.49$) as compared to participants in the out-group condition ($M=2.89$, $SD=1.64$), $U=488.50$, $Z=-1.13$, $p=.26$. One-way ANOVA's showed that there were no ordering effects on the distance measures.

In order to assess test-retest reliability of the measures we calculated correlations between the three measurement methods. Test-retest reliability was in general high (Table 6.2). This shows that participants gave the same responses, or in this case positioned (representations of) themselves at roughly the same position when repeating

the specific measurement.

Real	3D placement	2D slider
0.89**	0.86**	0.90**

Table 6.2: Test-retest correlations between the distance measures was in general high. * indicates the correlation is significant at $p < .05$ (2-tailed), ** at $p < .01$ (2-tailed).

6.4.4 Results

In this section we will first test the three hypotheses, after which we will provide some general remarks on the data.

6.4.4.1 Correlations between measurement types

The first hypothesis relates to the different measurement methods employed. In support of hypothesis 1 we found significant positive correlations between the measurement methods, as can be seen in Table 6.3. This shows that the three measurements are highly related to each other. It further shows that the two projective measures have convergent validity and provide an indication of participants' personal space behavior in real life. When comparing these correlations with the studies reviewed in Table 6.3, our correlations are pretty high. For example, Knowles (1980) reports correlations between $r = -.16$ and $r = .49$, Love and Aiello (1980) reports an average correlation of $r = -.23$ and Walkey and Gilmour (1979) reports correlations between $r = .45$ and $r = .83$.

	Real	3D placement	2D slider
Real	-		
3D placement	0.82**	-	
2D slider	0.84**	0.86**	-

Table 6.3: Correlations among measurement methods. * indicates the correlation is significant at $p < .05$ (2-tailed), ** at $p < .01$ (2-tailed).

Together with discriminant validity, the above-mentioned convergent validity is a subcategory of construct validity. In short: convergent validity measures whether concepts that should (theoretically) relate to each other do so, and discriminant validity measures whether concepts that should (theoretically) not relate to each other also do not relate to each other (Trochim). A limitation of this thesis is that it does not include a factor definitely unrelated to human-robot approach distance. The factor *negative attitude towards robots* (NARS), which we measured through the three scales, is at least conceptually different. Relations between the three measurement methods and the three NARS-scales were all weak (all r 's $\leq .23$ and all p 's $\geq .06$), providing an initial suggestion of discriminant validity.

Hypothesis 2 stated that the more two components have in common, the higher the correlation between the measures would be. In particular we believed that the real live measurement and 3D placement task would have the most in common. Table 6.3 shows that we have to reject this hypothesis as the 3D placement task and 2D slider

Figure 6.7: Participants in the in-group condition positioned themselves slightly closer to the robot (distances in cm). Legend: ■ Real, ■ 3D placement, ■ 2D slider.

task have the highest correlation. However, in general all three methods correlate with each other. This implies that there is a relationship between the three measures.

6.4.4.2 A group membership pattern in personal space

The third hypothesis posed that people interacting with an in-group robot would position themselves closer to the robot as compared to people interacting with an out-group robot. We conducted separate ANCOVA analyses with the three measurement methods as dependent variables, and *gender*, *height*, *nationality*, *perceived task intimacy*- and *pleasantness* and *robot experience* as covariates.

As can be seen in Figure 6.7 and Table 6.4, participants approached the robot closer if they were part of the in-group, as compared to the out-group. This result was similar for each of the three measurement methods. As these effects were also similarly non-significant for the measurement methods, we cannot accept hypothesis 3. Figure 6.7 shows that the direction of the effect is the same for each of the measurements. Therefore, each measure gives the same indication of the direction of the effect of group membership, though the exact distance cannot be predicted by the projective measures. In the discussion section we will provide a more elaborate discussion of the implications of this result.

Method	In-group	Out-group	Test statistic
Real	$M=62.3, SD=26.3$	$M=70.1, SD=31.6$	$F(1,60)=0.78, p=0.38,$ $partial \eta^2=0.01$
3D placement	$M=51.3, SD=19.9$	$M=55.0, SD=23.9$	$F(1,60)=0.17, p=0.68,$ $partial \eta^2=0.003$
2D slider	$M=76.7, SD=36.8$	$M=80.5, SD=39.1$	$F(1,60)=0.001, p=0.97,$ $partial \eta^2=0.00$

Table 6.4: Participants in the in-group condition positioned themselves slightly closer to the robot (distances in cm).

6.4.4.3 Personal differences & preferences

The ANCOVA analyses revealed that the distances between the participants and robots over the various measurement methods were related to a number of covariates. Task intimacy was significantly related to the approach distance in real life, ($F(1,60)=4.82$, $p<.05$, $r=.27$), and when placing the 3D figure, ($F(1,60)=4.58$, $p<.05$, $r=.27$). A non-significant trend was found for the 2D slider task, $F(1,60)=3.94$, $p=.052$, $r=.25$.

Furthermore, the ANCOVA analyses showed that gender was significantly related to approach distances when placing the 3D figure ($F(1,60)=4.49$, $p<.05$, $r=-.26$), and for the 2D slider task ($F(1,60)=6.27$, $p<.05$, $r=-.31$). In particular this shows that female participants positioned the representations of themselves closer to the robot than male participants. This effect was non-significant for the real life approach distance ($p=.13$). The covariates of height (all p 's $\geq .10$), nationality (all p 's $\geq .40$), experience with robots (all p 's $\geq .30$) and task pleasantness (all p 's $\geq .28$) did not significantly relate to the approach distances.

6.4.5 Discussion

The goal of this study was to investigate to what extent projective measures can be used to discover proxemics patterns in HRI.

We tested a total of three measures which showed high test-retest reliabilities between $\alpha=0.86$ and $\alpha=0.90$. Correlations between the measures were also high ($r=.82$ to $r=.86$). The pattern we have tested for showed to be consistent over the three measurements: participants part of an in-group with the robot stood slightly closer towards the robot than participants in the out-group.

Our results support the results found by Greenberg et al. (1980) and perhaps even more so the results of Knowles (1980) - the main difference being that we replaced the construct of acquaintance with group membership. These results are no argument to replace lab studies by studies solely using projective measures, as we found a significant difference between the distance of the three measurement methods (Figure 6.7). The only conclusion we can draw is that projective measures indeed seem to be a reliable indicator of people's behavioral response in live experiments; in other words projective measures could be used to investigate proxemics patterns in HRI. Because the two projective measures yielded different results (in terms of human-robot distance in centimeters), projective measures can not replace lab studies to answer these types of research questions. Research questions associated with data other than continuous data, such as "which approach direction would people prefer in certain situations", might be easier answered using projective measures. Examples in HRI include the video-based HRI studies (Woods et al., 2006a,b).

One of the HRI studies which is (design-wise) close to our study is the study of Kamide et al. (2014) in which participants were approached by a real robot and a robot in virtual reality - until the participant indicated the robot had approached to a point where they felt uncomfortable with the distance between them. In that experiment there were no significant differences between the approach distance between the real and virtual reality robot. Bailenson et al. (2001) also replicated results previously found in psychology on personal space in a virtual reality environment. While

we did not include virtual reality as a measurement method the work of Bailenson et al. (2001) and Kamide et al. (2014) would lead us to conclude that the two projective measurement methods would be an inferior way as compared to virtual reality. The clear advantage of the projective measurement methods we used however was that they do not require a calibrated virtual reality environment. We would therefore like to view these results in the following way: we sacrificed precisenesses of the data for a higher ease of use and practicability of the measurement method.

Group membership was not found to significantly influence participants' distancing behaviors. There could be two reasons for why we did not find significant differences. The first reason could be that the manipulation was not convincing enough - despite all participants responding correctly to the manipulation check. We based the manipulation upon the work of Nass and Moon (2000) in which participants were assigned to a green or blue team. The main difference was that participants were not dependent on the robot: there was no collaborative task which featured some sort of communication between the robot and participant. Moreover, we used a cap instead of a colored wristband. A second reason for why the manipulation was not effective could be that group membership does not contribute significantly to distancing behaviors when walking, given that we reviewed the literature about the influence of group membership on personal space was exclusively based on interactions between seated individuals (Glick et al., 1988; Novelli et al., 2010; Tedesco and Fromme, 1974).

To truly establish the value of projective measures to investigate personal space in HRI a follow-up would be required, not only to replicate these findings but rather to address a number of shortcomings in the current study. Personal space was measured using a single (behavioral) measure and especially in order to assess convergent validity it would be necessary to include a measure that is conceptually related to personal space and at the same time a different (type of) measure. This study has also not addressed the predictive validity of the projective measures, as we only measured the projective measures and real live measures in a single session. In order to assess predictive validity in future work we would ideally design an experiment in such a way that participants complete a second session at a later point in time.

6.5 Chapter conclusion

In Chapter 5 we investigated national cultural differences in preferences for human-robot approach distances. As we were unsure as to whether a projective (non-real) measure would be an accurate representation of real-world behavior, we investigated the effectiveness of projective measures to discover patterns in human-robot proxemics in this chapter. We first provided an overview of literature in psychology comparing various types of projective measures to measure personal space in human-human interactions. Following this overview we described the development of two personal space measures to assess personal space in HRI.

The goal of the study presented in this chapter was to investigate to what extent projective measures can be used to discover proxemics patterns in HRI. We have shown that the two projective measures used in this study - the placement of a scaled 2D and 3D figure - provide similar results compared with when participants

approached a real robot: participants in the in-group maintained less distance compared with participants in the out-group. The results gained through the projective measurement methods did not map exactly to the “real world” measurement. This sacrifice of preciseness for a higher ease of use and practicability limits the applicability of projective measures in that they can potentially be used to discover patterns in HRI (proxemics) research but not predict actual behavior of participants to centimeter-level precision.

In support of RQ-III projective measures could be a potentially valuable method to investigate patterns in human-robot proxemics. In both previous and the current chapter we looked at preferences and evaluations of human-robot distancing behavior. In the next chapter we will take a closer look at the evaluation of different (head) gaze behaviors of a robot while it guides a small group of people.

7

Impact of a Guide Robots' Head Movement on Preferences and Perceptions of Small Groups

"No plan and Stannis Baratheon sacks this city, takes the Iron Throne, puts your pinched little head atop a gate somewhere. It might be quite amusing, except that my head would be up there, too. I've never much liked my head, but I don't want to see it removed just yet."

Tyrion Lannister.

Game of Thrones: Blackwater

The work reported in Section 7.2 has partly been presented in: Joosse, M.P., Knuppe, R.A., Pinggen, G.L.J., Varkevisser, R.A., Vukoja, J., Lohse, M. & Evers, V. (2015) Robots Guiding Small Groups: The Effect of Appearance Change on the User Experience. In Proceedings of the 4th International Symposium on New Frontiers in Human-Robot Interaction.

In this chapter we will present two studies. The goal of these studies is to get a deeper understanding of people's preferences- and perceptions of a robot's head turn behavior. In the first study we have evaluated how people perceive two types of content and head direction displayed by a robot. In the second study we evaluated a more dynamic head movement behavior in a semi-controlled study in a hallway of a university building.

7.1 Introduction

Social robots are designed to interact with humans in everyday environments in a socially meaningful way (Lee et al., 2006b), in contrast to robots in environments such as factories. As a logical consequence, the design of social robots often includes human-like features, e.g., heads or arms in order to generate social responses. It has been found that by using such anthropomorphic cues, people automatically generate expectations of the robot's behavior (Lohse, 2011). These expectations people generate can be contrary to the future behavior (or intentions) of a social robot. Especially

in semi-public spaces it is important that people are able to understand and communicate with the robot without having to undergo specific training or possess a certain technical background. The capabilities of social robots differ from those of humans and allows them to use the anthropomorphic cues in different ways. For example, a social robot's eyes can face to the back while driving forward because the robot has other means (e.g., laser range finders) to detect obstacles. Thus, social robots (in the remainder of the chapter we refer to a "social robot" when using the word "robot", in terms of the definition provided in Chapter 2) could drive in a way which may look like as driving backwards.

In Chapter 2 we provided an overview of user studies investigating the effect of a robot's gaze on the perception of social robots in the context of guiding. Previous literature pointed out the effectiveness of deictic gaze (pointing towards people to maintain visitors' attention (e.g. Kuno et al., 2007a; Pitsch et al., 2009). Gaze has also been found an useful modality to attract visitors (Shiomi et al., 2010). When a robot has to guide groups of people, questions arises such as which direction the head and screen should face while guiding and what content should be displayed on the screen. The work reported in this chapter deals with these questions and specifically aims to address research question RQ-IV: "What head turn behaviors are appropriate for a robot guiding small groups?".

The goal of the this chapter is to get a deeper understanding of people's preferences for and perceptions of a robot's head turn behaviors in the context of being guided by a robot (Figure 7.1). To this end we present two studies in which we evaluated people's perception of a wayfinding robot. In both studies we asked small groups of 2-3 persons to follow the robot. The first study in this chapter (Section 7.2) investigates participants' preference for the direction of the head and type of information provided by a robot when being guided. We found that participants' prefer a robot to gaze towards the driving direction of the robot rather than backwards towards the group it guides. The second study follows up on this finding, and investigates a small group's response to constant and dynamic head turn behaviors. We found that a dynamic head turn behavior conveys feelings of "looking at me" and "caring for me" to participants.

Figure 7.1: Overview of the layout of chapter 7.

7.2 Study 7: The Effect of Appearance Change on the User Experience

In this section we present an exploratory study in which a robot guided 9 small groups of two to three people ($N=25$). We manipulated the appearance of the robot in terms

of the position of the “head”, a tablet providing information, facing forwards towards the driving direction or backwards towards the group being guided. The second manipulation was the type of information displayed (eyes or route information). The study was carried out in the DesignLab of the University of Twente. This is an open flexible workspace-like lab, where the location of the furniture (e.g. couches, tables, chairs) was changed from time to time.

7.2.1 Methodology

We conducted a between-subjects study with three conditions in which small groups were given a short guided tour by a robot through the DesignLab (Figure 7.2. At three waypoints the robot provided information through pre-recorded audio messages. The robot was equipped with a tablet representing the “head” of the robot, which faced either forwards (in the driving direction) or backwards (facing the participants). The type of information was the second condition manipulated; the tablet either displayed a set of eyes or relevant route information (distance towards the next way-point). A condition having the tablet mounted on the front of the robot, while displaying route information was deemed unnecessary as this would neither provide information for the participants following the robot, nor for other people present in the laboratory. Therefore we conducted the study with three conditions. We were specifically interested in finding out which combination of tablet direction and type of information provided (eyes or route information) would yield the most positive user experience.

7.2.1.1 Robot platform

For this study we placed a shell on top of a remote-controlled Festo Robotino robot platform. The height of the robot was 170 cm and it drove at a speed of approximately 0.7 m/s. The robot was operated by an operator who steered the robot through the DesignLab using a joystick connected to a laptop. Participants were not made aware of this before participating in the experiment.

7.2.1.2 Manipulation

We manipulated the direction of the tablet mounted on top of the robot and the information displayed on the tablet (Figure 7.3 and Table 7.1). We created three conditions, called $\text{Eyes}_{\text{forward}}$, $\text{Eyes}_{\text{backward}}$ and $\text{Distance}_{\text{backward}}$. In conditions $\text{Eyes}_{\text{forward}}$ (Figure 7.3a) and $\text{Eyes}_{\text{backward}}$ (Figure 7.3c) a set of blinking eyes was displayed

Figure 7.2: Small groups of participants were guided by the robot.

on the tablet either facing the participants or the walking direction. In condition Distance_{backward} we programmed the tablet to display route information, i.e., the remaining distance to the target (Figure 7.3e).

Figure 7.3: The appearance of the robot in the three conditions, showing the front and back side of the robot.

7.2.1.3 Measures

In a post-experiment questionnaire participants' experience was assessed using a variety of measures. All questions (except demographic- and open questions) were formulated as 5-point Likert-scaled items, and can be found in Appendix H. The number of questions varied by condition as particular questions were not applicable due to the specific manipulation. *General experience* questions (items V1 through V12) was assessed with measuring among others if participants trusted that the robot knew where it was going, if it was clear where the robot was going and whether or not the robot was helpful in guiding someone. In this set of questions also the speed of the robot and volume of the audio messages were evaluated.

Five questions related to the *physical appearance* assessed responses to the design (items A1 through A5), and specifically the height of the robot. *Usability questions* (items U1 through U7) included questions related to users' expectations of system capabilities and whether or not they were satisfied with the overall performance of the robot.

We collected demographic information with eight questions. The main information we collected was participants' age, gender and familiarity with robots (items D1 through D8). A manipulation check (item MC1) concerning the position of the tablet was included.

Overall, this resulted in 30 to 32 questions (see Appendix H). Next to these closed questions, we included three open questions to assess what participants liked and disliked about the robot, and what features they like to see in a guide robot.

7.2.1.4 Participants

We invited small groups of two to three people to participate in the study. Nine groups, in total 25 participants (14 males, 11 females), participated in the user study. Parti-

Condition	Eyes _{forward}	Eyes _{backward}	Distance _{backward}
Tablet direction	Front	Back	Back
Tablet display	Eyes	Eyes	Time to destination
N	9	8	8
Group composition	3x 3-person	1x 2-person 2x 3-person	1x 2-person 2x 3-person

Table 7.1: Overview of conditions and participants. We manipulated the direction of the tablet and the type of information provided to the participants.

cipants were aged between 17 and 40 ($M=23.76$, $SD=5.93$). All participants were students and staff from the University of Twente, primarily of Dutch (68%), German (8%) and Greek (8%) nationality. Participants indicated they had average experience with robots in general ($M=2.84$, $SD=0.90$) and little experience with social robots ($M=2.12$, $SD=1.09$).

7.2.1.5 Procedure

Small groups of participants were recruited to participate in a guided tour of the DesignLab. Participants were given a briefing, after which they were given a tour of about five minutes through the lab. Participants were requested to follow the robot. No specific instructions were provided regarding the distance they should keep to the robot. In this iteration the tour went past two points of interest (Figure 7.4, point B and C) where the robot provided a brief statement about the point of interest using a text-to-speech engine. Afterwards the robot returned to the starting position where participants were requested to fill out the post-experiment questionnaire (Figure 7.4 point A). Following the debriefing, participants were provided a candy bar as reward for their participation.

7.2.1.6 Data analysis

We calculated means for all items. To compare between conditions, the data were first tested for normality. The data was significantly different from a normal distribution, therefore Kruskal-Wallis and post-hoc Mann-Whitney tests are reported.

7.2.2 Results

Overall, participants reported a positive attitude towards the robot: they believed the robot was helpful ($M=3.84$, $SD=0.94$ on a 5-point scale. 1-sided binomial test against midpoint: $p<.05$), it moved at a comfortable speed ($M=3.12$, $SD=1.36$, 1-sided binomial test against midpoint: $p<.05$), and participants trusted that the robot knew where it was going ($M=4.48$, $SD=0.77$, 1-sided binomial test against midpoint: $p<.01$). These ratings did not differ significantly between conditions. Participants were moderately positive concerning the usability of the system: they felt moderately comfortable using it ($M=3.67$, $SD=1.05$, 1-sided binomial test against midpoint: $p<.01$) and were moderately satisfied with its performance ($M=3.56$, $SD=0.77$, 1-sided binomial test against midpoint: $p<.05$). No main effects or correlations were

Figure 7.4: Layout of the laboratory showing start/end position (A) and two points of interest (B and C).

found for the factors of gender, age, robot experience and/or educational background.

With respect to the robot's appearance, participants felt that the body design matched the robot's function (question A1, $M=2.72$, $SD=0.94$, 1-sided binomial test against midpoint: $p=.44$). Participants indicated the height was appropriate (question A3, $M=4.20$, $SD=0.82$, 1-sided binomial test against midpoint: $p<.01$). This is interesting because informal feedback gathered during the development indicated the robot was quite tall.

7.2.2.1 Closed questions

Kruskal-Wallis tests indicated there were significant differences between the three conditions for four questions. These were question V2 ("I was at a comfortable distance to the robot"), $H(2)=6.017$, $p<.05$, question V9 ("felt that the direction in which the screen was facing was natural"), $H(2)=7.843$, $p<.05$, question A2 ("The robot's design was intimidating"), $H(2)=10.529$, $p<.005$, and question U6 ("I enjoyed the guiding process"), $H(2)=6.335$, $p<.05$. Kruskal-Wallis tests revealed no further differences between conditions for the remainder of the questions concerning participants general experience and perception of the appearance of the robot and system usability.

Post-hoc Mann Whitney tests indicated the differences which were found were mostly due to the location of the tablet and the type of information displayed, thus between conditions Eyes_{forward} and Distance_{backward}, versus condition Eyes_{backward} where the tablet was mounted on the front of the robot.

Regarding question V9 in Figure 7.5, participants felt the direction of the screen was more appropriate in condition Eyes_{forward} ($M=3.89$, $SD=0.99$) compared with

Figure 7.5: Participants were more positive about the direction of the screen in conditions Eyes_{forward} and Distance_{backward} as compared to condition Eyes_{backward}. Additionally the robot in condition Eyes_{backward} was seen as more intimidating, and less enjoyable. Legend: ■Condition Eyes_{forward}, ■Condition Eyes_{backward}, ■Condition Distance_{backward}. * indicates significant difference at $p < .05$, ** at $p < .01$.

Eyes_{backward} ($M = 2.25$, $SD = 1.28$), $U = 11.5$; $Z = -2.46$, $r = -0.60$, $p < .05$. A similar effect was found between conditions Distance_{backward} and Eyes_{backward} ($M = 4.00$, $SD = 1.20$) in that the direction of the screen in the Distance_{backward} condition was perceived as more appropriate, $U = 10.00$, $Z = -2.36$, $r = -0.59$, $p < .05$.

Furthermore, the design in condition Eyes_{backward} was considered to be more intimidating (question A2, $M = 3.00$, $SD = 0.97$) compared with condition Eyes_{forward} ($M = 1.78$, $SD = 0.68$), $U = 11.50$, $Z = -2.51$, $r = -0.61$, $p < .05$ and condition Distance_{backward} ($M = 1.50$, $SD = 0.54$), $U = 6.00$, $Z = -2.89$, $r = -0.72$, $p < .01$. Participants in condition Distance_{backward} enjoyed the guiding more (question U6, $M = 4.13$, $SD = 0.35$) compared with those in condition Eyes_{backward} ($M = 3.25$, $SD = 0.71$), $U = 10.50$, $Z = -2.62$, $r = -0.66$, $p < .05$.

7.2.2.2 Open questions

The participants provided a total of 33 positive and 33 negative statements about their experiments with the robot (Figure 7.6). Three coders categorized the statements. The voice and information provided by the robot were both mentioned seven times as a positive aspect. Six participants liked the eyes. Furthermore, positive statements were made regarding the appearance (4), the experience of being guided (4), the speed (3), turning behavior (1), navigation capabilities (1).

There were also several noteworthy aspects people disliked. Some of these aspects have also been mentioned by other participants as being positive; we will reflect upon this in Section 7.2.3. Nine participants disliked the speed of the robot. Eight participants disliked the appearance of the robot, and 6 commented on the voice. Also comments were made regarding the turning behavior (3), the goal (1), and amount of information given (1).

Participants also provided some recommendations on what they would like to see in a guide robot. A total of 42 statements were made by the participants, as

Figure 7.6: Participants self-reported likes (green) and dislikes (red) about the tour and the robot during iteration 1.

with the other open questions we categorized the statements. Among these were the possibility to ask questions (7), get more visual information at the way-points (6) and while underway to way-points (4), a more personalized tour (5) and a map (5). Additional recommendations were made towards improvement of the design of the robot (4) and hardware (5).

7.2.3 Discussion

In this exploratory study we have examined the influence of a robot's appearance and input modality on users' experience in a guiding scenario. Neither gender, age nor experience with robots influenced the evaluation of the robots significantly, though this could also be due to the small sample size.

Small groups of people were provided a short tour by a robot. Results indicate that the location of the screen can be either forward or backward, depending on the information displayed. In the case of eyes facing participants, our results showed that this was considered to be very unnatural and intimidating. On the other hand, when the tablet faced participants and route information was provided this was again evaluated as more useful. This might seem to be in contrast with the results of Shiomi et al. (2010) who found that eyes facing participants are more effective to attract bystanders. However, we think this could be explained because in our setup the participants had already been introduced to the robot and asked to follow it.

Points for improvement include the speed of the robot as participants found the speed of the robot in the study too slow. In the next study we will use the insights gained in this study in an experiment where investigate the evaluation of different head turn behaviors on users' experience in an experiment with the SPENCER robot.

7.3 Study 8: “Are You Still There?": Evaluation of a Guide Robot's Head Turning in the Wild

Social robots, and robots in general, come in all kinds of shapes and sizes. A social robot, operating autonomous or not, differs from a “regular” robot in that they operate in an environment specifically designed for humans. One feature many social robots have in common is that they have a design which evokes certain anthropomorphism characteristics, e.g. people attribute humanlike characteristics to the robot (Duffy, 2003).

This study concerns the topic of a robot's head behavior while in a guiding situation. Previous work has investigated where robots should gaze when attracting people's attention (Shiomi et al., 2010) and when guiding people (Study 7, reported in Section 7.2). The SPENCER robot is a robot with a rotating head, which is thereby capable of seemingly looking over its shoulder. A question which arises is whether a robot should occasionally look back over its shoulder when guiding people, for example to seemingly see if the group is still there.

Therefore, we have conducted two studies. In a pre-study we have investigated which head turn behaviors out of a set of 16 were considered the most natural way for the robot to look back towards the participants (Section 7.3.1). In the main study we implemented the behavior considered to be the most natural on the robot. This was compared to a head behavior sequence constantly facing the participants (Section 7.3.2). We then asked 69 participants to follow the robot in small groups.

The contribution of the research presented in this paper is that we investigate to what extent the behavior of the head during guiding results in a different evaluation of the robot as found in previous research in HRI. This section is structured as follow: we will first discuss the pre-study in which we tested which out of 16 head turn behaviors is most normal (Section 7.3.1). This head turn behavior was then implemented on a robot platform, and evaluated in a between-groups experiment in the wild. We will present the method (Section 7.3.2) and results (Section 7.3.3) of this experiment. We will conclude with a discussion of the results found (Section 7.3.4).

7.3.1 Pre-study: Identifying the effective head turn

In Study 7 it was suggested that people guided by a robot have a preference for a robot which looks in the direction of travel while guiding them. In a different context Shiomi, Kanda, Ishiguro, and Hagita (2006) was suggested that a robot looking in the opposite direction, the direction of the people, would be more successful in attracting people's attention.

To select the most appropriate head turn for the main study we designed 16 head turns, which varied along two dimensions: the focus of the head (on the user, to the left or right) and the range of the turn of the robot (narrow to wide). We operationalized this in 5 types of head turn, and 3 different fields of vision (or turn angles). Inspired by Read and Belpaeme (2014), we created short video sequences of the head

Figure 7.7: We designed 16 head turn behaviors, which varied along two dimensions: the type of turn (made up on number of times looking at the participant, and the environment) and the range of the turn angle (narrow - wide range).

turns from the perspective of people that are following the robot. In an online survey participants rated these video sequences. This survey had a within-subjects design.

7.3.1.1 Stimuli

We designed five distinct *head turn sequences*, which differed in the robot's focus point in the environment (people guided by the robot, looking left or looking right). For example, one sequence included "focus on the people being guided, look at environment right, look at environment left, look at people", while another sequence only included "focus on the participant, environment right". A schematic overview of the sequences is provided in Figure 7.7. The second manipulation we added to the head movements was the *range of the turn angle* of the robot, varying from narrow to wide range. We operationalized this in three different turning angles; no turn (0 degree), narrow- (15 degree), medium- (30 degree) and wide turn (45 degree), as can be seen in Figure 7.7. The 0-degree turn range was only used for sequence D, which would be a 180-degree turn. Therefore, a total of 16 videos were created¹.

Videos of each condition were created using Unity². Each video had a length between 10 and 12 seconds, and featured a first-person view of the person and two other people following the robot (Figure 7.8). Two seconds into the video, the robot would start to turn its head. In the video the robot and the guided people moved past static objects (colored basic shapes), so as to strengthen the feeling of moving through space with the robot.

Figure 7.8: Participants were shown a short video sequences in which they were part of a 3-person group following a robot.

7.3.1.2 Procedure

Participants were recruited through the CrowdFlower platform³, and were invited to complete the online questionnaire. Participation was limited to people from European countries⁴. The questionnaire consisted of general information about the experiment and procedure. After that, participants were shown each of the 16 videos in a randomized order. For each video participants were asked to rate on a 7-point semantic differential scale how natural they found the head turn of the robot. Six control

¹YouTube playlist available at <http://goo.gl/YOMLT0>

²<http://www.unity3d.com>

³<http://www.crowdflower.com>

⁴Participants from Austria, Belgium, Denmark, France, Germany, Ireland, Luxembourg, the Netherlands, Switzerland and the United Kingdom were allowed to participate.

questions were added to measure recollection of the video, asking for specific details such as the colors of the various objects. Four demographic questions (gender, age, nationality and prior experience with the robot) were included at the end of the questionnaire. After completing the questionnaire participants were paid €0.40 through CrowdFlower. 39 participants who failed to correctly answer at least five of the six control questions were excluded from the sample.

7.3.1.3 Participants

A total of 29 participants completed the survey, and took on average 8 minutes to complete the survey. The sample consisted of 19 males and 10 females, aged between 21 and 71 ($M=38.1$ years, $SD=11.90$ years). Participants had the German (34.5%), British (17.2%), Austrian (17.2%) and Dutch (13.8%) nationality.

Post-hoc Kruskal-Wallis tests revealed no significant differences between the different nationalities (all p 's $\geq .08$).

7.3.1.4 Results

A Kruskal-Wallis test revealed a significant difference between the 16 different behavior sequences $H(15)=34.82$, $p<.005$ for naturalness. Mean values of naturalness for all videos are provided and visualized in Figure 7.9. As can be seen there is a preference for turn type C with a medium turn range. Based upon these results we decided to implement this sequence on the robot. In this sequence the robot first turned 210 degrees counterclockwise, waited for 2 seconds, turned 60 degrees clockwise, waited for another two seconds before returning to the original forward position (Figure 7.10).

		Turn type				
		A	B	C	D	E
Range of turn angle	0°				4,48 (1.99)	
	Narrow	5,55 (1.86)	4,90 (2.09)	5,59 (1.92)	5,38 (1.82)	4,79 (1.61)
	15°					
	Medium	5,31 (2.09)	5,24 (2.44)	5,86 (1.81)	5,52 (2.01)	4,76 (1.86)
	30°					
	Wide	4,72 (2.1)	4,55 (2.23)	4,07 (2.09)	5,52 (2.34)	3,76 (1.94)
	45°					

Figure 7.9: Heatmap showing mean values and standard deviation (between brackets) for the 16 tested head turn sequences; sequence C30 was deemed most natural by participants.

7.3.2 Main study: Method

In Section 7.3.1 we reported a pre-study where we investigated which head turn was perceived as most natural out of a total of 16 head turns. In the current study we investigated in a real world context people's preferences for this dynamic head

behavior compared to a constant head behavior where the robot continuously faced the participants.

7.3.2.1 Hypotheses

As far as we know, this is the first study to compare people's perceptions of a robot's constant and dynamic head turn behaviors in the context of guiding people. Because the dynamic head turn behavior was rated as most natural, we expected that this behavior, when implemented on an actual robot, would be considered as more human-like and socially competent.

Therefore, our first hypothesis is as follows:

H1. People perceive a robot with dynamic head turn behavior as more socially competent as compared to a robot with constant head behavior.

At the same time, when people feel a robot is more socially competent their attitudes towards the robot could also be more positive.

H2. People have a more positive attitude towards as their perception of its social competency increases.

7.3.2.2 Manipulation and apparatus

We manipulated the head behavior of the robot to be either *constant* or *dynamic* (Figure 7.10).

- In the *dynamic head* condition the robot turned its head, and faced the participants at specific intervals. Based on the pre-study, we implemented the turning sequence of the robot head, which consisted of the robot facing forwards, and turning its head backwards every 25 seconds.
- In the *constant head* condition, the head faced the participants, but turned left and right every 25 seconds, just like the dynamic condition the head turned left and right every 25 seconds

For this experiment we used the SPENCER robot platform, which is equipped with a 2-DoF turning head. SPENCER uses four RGD-B cameras and two SICK LMS 500 laser scanners for navigation and motion planning. In the experiment the robot drove autonomously to five way-points while avoiding obstacles and people. The positions of the people following and moving around the robot were generated using a laser-based people tracker.

7.3.2.3 Task and procedure

The experiment was conducted in a public hallway of a university building. Recruited participants were informed of the goal of the robot (guiding passengers at an airport), and that we were in the final phase of prototype development. We further told participants that we were interested in their evaluation and feedback on some of the behaviors of the robot. To do so, participants were guided by the robot while it drove two laps around the hallway and filled out a questionnaire afterwards. Participants were guided by the robot for about four minutes ($M=273$ seconds, $SD=52$ seconds). During this time the robot drove between 174 and 184 meters ($M=178.33\text{m}$, $SD=2.28\text{m}$)

Figure 7.10: Two 25-second head behavior sequences were programmed on the robot: the head turning condition, based upon the insights gained in the pre-study, and the head facing condition in which the robot faced the participants.

at an average speed of 0.67 m/s ($SD=0.10$ m/s). The difference in distance was caused by people hanging out in the hallway, which forced the robot to drive around them to the waypoints.

Following the briefing, the experiment started with the robot navigating through the hallway, and participants following the robot (Figure 7.11). Having followed the robot through the hallway, participants were asked to complete a pen-and-paper based questionnaire consisting of open and closed questions (Section 7.3.2.4).

After completing the questionnaire, participants were debriefed, and any remaining questions of the participants were answered. The total duration of the experiment was about 15 minutes. Participants were paid €5 for participating in the experiment.

7.3.2.4 Measures

Both objective and self-reported subjective measures were collected during the experiment.

We collected video data through the rear-facing RGD-B cameras and through two

Figure 7.11: 29 groups of 2-3 persons followed the robot as it guided them in a hallway.

cameras mounted on tripods. The data from the laser-based people tracker were used to calculate the distance between the robot and participants.

Participants' subjective feedback was collected through a self-reported questionnaire. A 49-item questionnaire consisting of both open and closed questions was designed to measure participants' user experience (the complete questionnaire is available in Appendix I). Trust in the robot was measured through the competence (6 items, $\alpha=0.66$) and goodwill (6 items, $\alpha=0.69$) constructs of the Source Credibility Scale (McCroskey and Teven, 1999). Perceived anthropomorphism (5 items, $\alpha=0.58$) and likeability (5 items, $\alpha=0.81$) scales were used to assess participants' general perception of the robot's design (Bartneck, Kulić, Croft, and Zoghbi, 2009).

Attitude towards robots was measured through the NARS (Nomura et al., 2008) interaction with robots (6 items, $\alpha=0.66$) and social influence of robots (4 items, $\alpha=0.68$)⁵ subscales. Experience with robots was measured through one multiple-choice item, and engagement with the robot through one 7-point Likert scaled item. Satisfaction with the robot was measured through 3 7-point Likert scaled items ($\alpha=0.80$) based on Lee et al. (2006b).

Eight demographic questions provided us with information about participants' gender, age, study program and flying experience. Finally, three open questions were included in order to get a richer understanding of the positive and negative aspects of the robot:

1. What did you like about the robot's behavior?
2. Do you have any comments and/or suggestions regarding the behavior of the robot if it were to guide you at an airport?
3. Which aspect of the guiding do you think should be improved before deploying this robot at an airport?

7.3.2.5 Participants

Small groups of 2 or 3 people were recruited at the Computer Science faculty of the University of Freiburg, Germany. 72 participants participated in the experiment, due to technical issues with the robot data of two groups had to be excluded from the sample, which therefore consisted of 68 participants, divided over 29 groups (Table 7.2).

Participants were aged between 18 and 35 ($M=24.56$, $SD= 3.86$); 73.5% of the sample was male, 26.5% female. Most participants had the German nationality (64.7%), followed by Russian (5.9%), Chinese (4.4%) and Indian (4.4%) nationalities. 48.5% of the participants had a background in Computer Science, other participants generally had a technical background as well, such as from the field of Embedded Systems Engineering (14.7%) and Microsystems Engineering (16.2%).

Participants' experience with robots was high: 23.5% indicated they had seen robots before, 19.1% of the participants previously worked with robots, and 50% indicated they had previously built or programmed robots themselves.

⁵Item "I feel that in the future, society will be dominated by robots" removed due to low internal reliability

Condition	2p. group	3p. group	Total
Constant head	16 (8)	15 (5)	31 (13)
Dynamic head	22 (11)	15 (5)	37 (16)

Table 7.2: 68 participants, divided over 29 groups (between brackets) participated in a between-groups experiment

7.3.2.6 Data analysis

Internal reliability for the subjective scales were deemed acceptable (see Section 7.3.2.4). Due to the questionnaire being administered on paper, the sample contained 12 missing values. In order to carry-out list-wise comparisons, we replaced these with the means (of the entire sample) of the respective variable.

A Kolmogorov-Smirnov test was conducted to test for normality. All dependent variables were normally distributed, with the exception of the NARS interaction subscale. Therefore, in general, we conducted *t*-tests and ANOVA's to test our hypotheses. For the NARS interaction subscale we conducted Kruskal-Wallis and Mann-Whitney tests instead.

We grouped all answers to the open questions into different categories. Some participants chose not to answer one or more open questions, while others provided multiple answers to a single question. Therefore we use percentages to show differences of the frequencies of categories between the two conditions (Tables 7.3 - 7.5).

Data from the SICK laser scanners was used to calculate the mean distance between the participants and robot. During the experiment we observed that participants did not always understand the intention of the robot when it turned on the spot at the end of the hallway, and when making sharp turns (e.g. 90 degrees or smaller). Therefore we excluded the areas where the robot turned when calculating the distances between the robot and participants (Figure 7.12).

Figure 7.12: Tracks showing positions of a robot (blue track) guiding a 3-person group (purple, red and yellow tracks). Highlighted area shows area taken into consideration when calculating features (such as distance).

7.3.3 Results

In this section we report the results for each of the hypotheses. We will first present some general results.

Participants' general subjective assessment of the robot was rather positive: mean ratings for competence ($M=4.90$, $SD=0.76$), goodwill ($M=4.53$, $SD=0.89$), and likeability ($M=5.61$, $SD=0.80$) were above the midpoint of the 7-point Likert scale. As expected, the negative attitude towards robot's subscales interaction ($M=2.27$, $SD=0.89$) and social influence ($M=3.66$, $SD=1.27$) received relatively low scores. Self-reported engagement was also high ($M=5.66$, $SD=0.93$). This leads us to conclude that the participants had an overall positive attitude towards the robot. This conclusion is supported by the open questions, where participants indicated that appreciated the autonomous navigation, including the motion planning system (Table 7.3).

In the constant head condition, participants kept more distance from the robot ($M=1.91\text{m}$, $SD=0.59\text{m}$) compared with participants in the dynamic head condition ($M=1.84\text{m}$, $SD=0.79\text{m}$), $t(66)=-0.44$, $p=.66$.

From the answers to the first question; *"What did you like about the robot's behavior?"*, we gather that participants in particular liked the motion planning capabilities and the speed of the robot (Table 7.3). In the dynamic head condition participants mentioned appreciation for the head turning behavior more often as compared with the constant head condition.

7.3.3.1 Head turning behavior did not influence perceived social competence

A t -test did not reveal a significant difference between the two conditions for social competence. We thus failed to find support for H1 which stated that people would perceive a robot with dynamic head movement as more socially competent.

"That he turned his face to us sometimes, which makes me feel like cares for his customers."

Category / condition	Constant head	Dynamic head
(no comment)	2 (6.45%)	3 (8.11%)
Motion planning	9 (29.03%)	13 (35.14%)
Head appearance	5 (16.13%)	7 (18.92%)
Head checking behavior	4 (12.9%)	14 (37.84%)
Speed	8 (25.81%)	9 (24.32%)
Robot appearance	1 (3.23%)	1 (2.70%)
(Smooth) movement	2 (6.45%)	5 (13.51%)
Get the job done	3 (9.68%)	0 (0%)
Quiet mechatronics / computers	3 (9.68%)	1 (2.70%)
Total	50	35

Table 7.3: Participants in particular liked the motion planning capabilities of the robot; in the dynamic head condition participants showed appreciation for the head turning behavior. Table reports number of comments and percentage of 68 participants.

Figure 7.13: The only significant difference between the two constant head direction (white bars) and dynamic head direction (gray bars) was on the anthropomorphism scale (** indicates a significant difference at $p < .01$).

The answers of the open questions provides additional information not measured through the questionnaire. This is illustrated by the quote above. In the dynamic head condition, 37.8% of the participants found the behavior helpful, as it somehow indicated the robot checked on them. As only 12.9% of the participants in the constant head condition showed positive appreciation for this feature of the robot, participants in the two conditions seem to have noticed a difference between the head turn behavior. We will reflect on this, and the methodological implication, in the discussion.

7.3.3.2 Social competence

The second hypothesis stated that people have a more positive attitude towards the robot as their perception of its social competency increases. We found that perceived competence was significantly negatively correlated with “negative attitudes towards interaction with robots” on perceived competence ($r = -.38$, $p < .05$), and positively

Category / condition	Constant head	Dynamic head
(no comment)	4 (12.9%)	3 (8.10%)
More acceleration/deceleration	2 (6.45%)	2 (5.41%)
Communicate (motion) intent	2 (6.45%)	2 (5.41%)
Faster turning arcs	3 (9.68%)	1 (2.70%)
Feedback: map	5 (16.13%)	9 (24.32%)
Feedback: progress	8 (25.81%)	7 (18.92%)
Feedback: auditory	3 (9.68%)	7 (18.92%)
More fluent movement	2 (6.45%)	2 (5.41%)
More interaction	4 (12.9%)	7 (18.92%)
Smooth path planning	1 (3.23%)	0 (0%)
Speed adaptation	2 (6.45%)	3 (8.11%)
Total	32	40

Table 7.4: As improvement the robot could have provided more feedback, for example route information through a map.

correlated with likeability ($r=.49, p<.001$) and anthropomorphism ($r=.30, p<.05$). Therefore, we accept H2 as participants' had a more positive attitude towards the robot as their perception of the social competency of the robot increased.

“He really took care of our group and searched the right way through the hall. He was also kind of cute with his turning head.”

7.3.3.3 Autonomous driving is not enough

When asked about improvements for the robot (Table 7.4), participants indicated that they would have liked to receive more feedback, for example route information through a map and the progress of the tour. These two categories, making up 43% of the comments, had previously been implemented on the user interface of the robot, but were not used in the experiment.

The final open questions inquired about capabilities of the robot which participants felt should be improved before being deployed in an actual airport environment. Answers were quite diverse, as can be seen in Table 7.5. The feature mentioned most was speed adaptation (by 15 participants), followed by more feedback to the passengers while they're being guided, including a map and/or time to arrival.

This indicates that the robot's autonomous driving and motion planning capabilities were appreciated, but at the same time more feedback should be provided in order to create a more positive use experience.

7.3.4 Discussion

The work presented and the results reported in this paper provide several topics for discussion.

7.3.4.1 Implications for design

An interesting insight we gained was that participants were positive about the autonomous navigation and motion planning capabilities of the robot, but at the same time

Category / condition	Constant head	Dynamic head
(no comment)	11 (33.33%)	11 (26.83%)
Speed adaptation	4 (12.12%)	11 (26.83%)
Feedback route / flight info	5 (15.15%)	5 (12.20%)
Feedback intention	3 (9.09%)	1 (2.44%)
Feedback shops	1 (3.03%)	2 (4.88%)
Engage conversation.	1 (3.03%)	3 (7.32%)
Carry luggage	1 (3.03%)	2 (4.88%)
Misc	4 (12.12%)	6 (14.63%)
Feedback - auditory	3 (9.09%)	0 (0%)
Total	33	41

Table 7.5: Most participants suggested speed adaptation should be implemented, followed by more feedback to the passengers while they are being guided, including a map and/or time to arrival.

more critical about other, mostly interaction, capabilities. The robot should have provided more feedback on its progress, for example through a map or through auditory feedback. This could argue for a more “all singing, all dancing” robot, rather than a robot where individual features - here head turn behavior - were evaluated.

At the same time, the environment we used was a semi-public space with variances in the level of crowdedness. Different from a controlled lab study, the quality of video data was more noisy and therefore less useful. The design of our measures (heavily relying on closed questions) might therefore not have been the best choice, especially given that we gathered most interesting results from open-ended questions. In future research a setting like this could warrant interviews as the main source of data, such as employed in Rothenbücher, Li, Sirkin, Mok, and Ju (2015).

Using a robot with limited functionality has the advantage that you receive feedback on a variety of topics. Most of the suggestions for improvement addressed by the participants are not shocking; they could have been implemented. An example of such is the request for more interaction and feedback through a map. What was particularly noteworthy was that a group of 15 participants (20.8%) indicated the robot should be equipped with some form of speed adaptation, a feature not only related to motion planning, but requiring the integration of various components (such as people tracking and potentially re-identification). For future project consortia this could be an interesting issue to address in a multi-disciplinary project team. We do want to stress that participants were fully briefed about the future goal of the robot in the context of an airport, e.g. for guide robots in other environments (zoos, museums), this requirement could be less important.

For designers of robot behaviors our findings imply that when designing behaviors for robots in public space imitating human behavior, especially the “checking back”, can provide improve interaction between user and robot.

7.3.4.2 Limitations

This work is limited in several ways. For example, the scenario employed was quite artificial as participants only had to follow the robot while it drove two laps through a hallway corridor; and as such the participants were not dependent on the performance of the robot. In previous works, such as Study 4 we mitigated this effect of a lab study by introducing a context where it was natural for a robot to (in that case) approach. An example in the current situation could have been to quiz participants on certain elements they encountered while following the robot and coupling the reward to their performance.

A second limitation was the pre-defined intervals at which the head of the robot turned. Research in HRI investigating the applicability of gaze found that when providing information about exhibits in museums. It was found that gazing at the audience at interactionally significant points, or at the end of a sentence increased participants’ attention to the robot (Kuno et al., 2007a; Yamazaki et al., 2009). In retrospect, it could have been more effective if the robot would have turned around to gaze at the participants when approaching turns.

The third limitation relating to the setup is the comparability between the two conditions. A logical addition would have been a third condition, which constantly

faced forward, as this would in a way have been a logical consequence of the results found in Study 7. During the design of this experiment we were faced with limited resources, primarily in time and available number of participants, and therefore we made the choice of using the degrees of freedom available in the head of the robot, and thus using a condition in which the head turned back to face the participants, as opposed to a second condition where the head of the robot would constantly face forward.

The final limitation was the sample, which had a mostly technical background, and differed from the sample of the pre-study, primarily in terms of age. When deployed, it is likely that robots in general will interact with people not accustomed working with technology and who thereby could potentially have different expectations of the capabilities of robots. Not only was the sample quite homogeneous in terms of age and field of employment (students), the sample also primarily included people with a technical background. Therefore, a logical follow-up could be to run this type of studies in (semi-)public spaces which audience is more representative of the general population. A shopping mall or store could be an example. The study reported in Chapter 8 addresses some of the limitations of context and sample by evaluation the same robot in a guidance scenario at an international airport.

7.3.5 Conclusion

In this study we set out to investigate participants' expectations and evaluation of a guide robot when being guided around. We have first conducted a pre-study to investigate the most appropriate way to "look back" for a robot. This behavior, and one other behavior where the robot continuously faced participants was then implemented on a robot. 72 participants took part in the experiment where they followed the robot and afterwards completed a questionnaire.

The contribution is two-fold to the HRI field. One the one hand we found that participants seemingly interpret a robot's simple head turning behavior as "looking at them", and caring. The second contribution is this study shows that validated questionnaires may not always be the best way to go in semi-controlled settings when conducting first tests with robots. It is clear that more qualitative methods (interviews, questionnaires with open questions) are necessary to complement quantitative research methods for field trials.

7.4 Chapter conclusion

In this chapter we have presented two studies, both related to the behavior of a robot's head turn while guiding a small group of participants in fictional guiding scenarios. The goal was to get a deeper understanding of people's preferences for and perceptions of a robot's head turn behavior, summarized in RQ-IV: "What head turn behaviors are appropriate for a robot guiding small groups?".

In the first experiment we manipulated the behavior of the head in a static way, and found that participants had a preference for a forward-facing head if it displayed eyes. In the second experiment we implemented a more dynamic head behavior on

the SPENCER robot. Despite our expectations relating to the effectiveness of “looking back” to the participants we could not accept this hypothesis based on the self-reported questionnaire. The open questions however did show participants appreciation for the head behavior. One possible limitation of this experiment was that the implementation was too artificial, in that the robot turned its head at a pre-specified interval, rather than at pre-specified locations, such as when turning.

Based on the results of these studies, robots that guide small groups of people should look towards the driving direction. The results of especially Study 8 hint towards a more complex answer to the research question: looking back was appreciated by the participant, or at least, intentions relating to “caring for the group” were attributed. It was beyond the scope of the experiment to investigate at which point a robot should turn its head or if these findings would also apply to robots featuring a less anthropomorphic design. These are all potential avenues for future research, and may be especially relevant in case future research also shows that people would be less inclined to follow a robot if it were not to make eye contact at times.

In Chapter 8 we will present a case study of the in-the-wild deployment of the SPENCER robot where we evaluated airport passengers’ responses to the SPENCER robot in a guiding scenario.

8

Moments of Truth: A Case Study in the Wild

"I'm no writer; but if I were, it seems to me I'd wanna poke my head up every once in a while and take a look around, see what's going on. It's life, Jake. You can miss it if you don't open your eyes."

*Benjamin Sisko.
Star Trek Deep Space Nine: The Visitor*

The work reported in this chapter has partly been presented as: Joosse, M.P. & Evers, V. (2017) A Guide Robot at the Airport: First Impressions. In Proceedings of the HRI'17 Companion, pp. 149-150.

In this chapter we will present a user evaluation of the SPENCER robot's field deployment at Schiphol airport. We will first provide details on the context of the study and technical specifications of the SPENCER robot platform. Following this introduction we will present the methodology used to collect airport passengers' feedback.

8.1 Context of the study

In the previous chapters we have described user studies investigating appropriate non-verbal behaviors for a wayfinding robot, from settings without specific contexts to settings with (fictional) people, obstacles and an autonomous driving robot. These studies were mostly conducted in laboratories or other controlled settings, which had advantages as it allowed us to compare different conditions. A disadvantage of these studies is that there was a lack in realism (Oulasvirta, 2009, p.92). It is in particular difficult with lab studies to understand how contextual factors will influence the future interactions between a robot and its users (Šabanović, Reeder, and Kechavarzi, 2014). To mitigate some of the disadvantages of lab studies, we have conducted a study outside of the laboratory in an airport. In this chapter we will present a case study investigating airport passengers' evaluation of the capabilities of the SPENCER platform.

During the project lifetime the robot was deployed at Schiphol airport for technical tests at two occasions. These tests included mapping the environment, testing and fine-tuning technical components such as the robot's ability to autonomously navigate and localize and track people. One of the moments of truth of the SPENCER project came with the user experience evaluations conducted immediately following the technical tests, to show that besides being technically capable, the robot is also deemed acceptable and useful by representative users, in this case airport passengers. The study presented in this chapter aims at addressing research question RQ-V "Under which circumstances would a wayfinding robot provide added value to airport passengers?".

Before presenting the methodology and results of the use case, we will first provide background information on the technical specifications of the SPENCER robot and its operating environment.

8.1.1 The SPENCER platform

The design of the SPENCER platform called for an aesthetically-appealing platform, equipped with hardware that allowed the robot to autonomously navigate for an extended period of time. The SPENCER platform measures 81 x 80 cm and has a height of 192 cm. The maximum tested velocity of the platform was 1.6 m/s, though the maximum operating speed is 1.3 m/s for safety reasons. When moving forward, the robot can accelerate at 0.9 m/s^2 . The number of batteries allows for a full day's operation without charging. During operation, safety for people, the robot and environment was provided through various means including two emergency stop buttons, a wireless emergency stop button operated by a trained operator and two bumpers on the robot.

The sensors of SPENCER consist of two SICK LMS 500 2D laser scanners, providing 360° coverage and mounted at 65 cm height. Two front and two rear RGD-B Kinect sensors and one forward-facing stereo camera complement the sensor suite. In order to navigate to a destination, robots require similar information to people; challenges which could be described as "where am I?", "where do I need to go?" and "are there obstacles (including people) along the way which I need to avoid?". To deal with these challenges various software modules have been developed and tested. SPENCER uses a Normal Distributions Transform Occupancy Map (NDT-OM) to store a map of Schiphol, which is augmented with a Conditional Transition Map in order for the

Figure 8.1: The study and preceding technical tests were conducted at Schiphol airport.

robot to learn the patterns of motion (Kucner, Saarinen, Magnusson, and Lilienthal, 2013). These patterns can be used to predict the direction of motion of other actors, which is an important feature for socially aware navigation (Triebel et al., 2016). SPENCER's motion planner uses a Theta*-RRT algorithm, and has been designed in a way that it generates smooth and legible paths, at a minimal planning time (Palmieri, Koenig, and Arras, 2016) and SPENCER is capable of learning socially normative behaviors, such as social compliant group avoidance (Okal and Arras, 2016) and adapting its motion behavior to the group (Fiore, Khambhaita, Milliez, and Alami, 2015).

SPENCER combines the advantages of laser range scanner (large field of view, robots against illumination) and RGD-B sensors (appearance information) in a real-time multi-modal people tracking framework (Linder, Breuers, Leibe, and Arras, 2016). To understand social situations, it is important to detect human attributes, such as gender (Linder and Arras, 2015) and the estimation of head pose (Beyer, Hermans, and Leibe, 2015).

SPENCER is furthermore equipped with a boarding card reader and touchscreen to interact with passengers. To strengthen the humanoid appearance, SPENCER was also equipped with a 2-DoF head.

8.1.2 Operating environment

SPENCER has been deployed at Schiphol, the main international airport of the Netherlands. Schiphol has one terminal, separated into four halls, which serves on an annual bases 63 million passengers, of which 37.8% transfer passengers¹. The deployment of the SPENCER robot took place in the Schengen-part of the terminal: this area encompasses Lounge-1, Transfer-2, the B- and C-gates and the upper level of the D-gates. As SPENCER would have to take an elevator to reach the Schengen D-gates this area was deemed beyond the scope of the current tests, and therefore not mapped. In its totality, the deployment area featured lounges and corridors with a variety of surfaces and lighting conditions (Figures 8.2 and 8.3). Additional challenges which characterize the operating environment include conveyor belts and obstacles such as benches, information signs and artworks as previously identified in Section 4.2.

Figure 8.2: Participants were guided from Lounge-1 towards the B-gates or vice-versa.

¹Source: Monthly Traffic & Transportation figures (in Dutch), January-December 2016. Retrieved from: <https://www.schiphol.nl/nl/schiphol-group/pagina/verkeer-en-vervoer-cijfers/>, last accessed 18 April 2017.

8.2 Methodology

Early on we realized that user tests in a public space like Schiphol airport would introduce factors we would be unable to control for, such as crowdedness, passengers' level of travel experience, and the behavior of other passengers around the robot. We therefore decided to conduct a (single)case study with 16 airport passengers. A case study focuses on the observation of a phenomena in a context that is beyond the control of the investigator (Lazar, Feng, and Hochheiser, 2010, p. 148). This method is particularly suitable when answering questions relating to the "how" and "why" of an event, and when the researcher does not have full control over the events within context (Gray, 2014, p. 247). In this study the phenomenon we observed was passengers experience with a wayfinding robot at an airport - a context in which we cannot control events. The goal is to come to a deeper understanding of the circumstances under which a wayfinding robot provides added value for airport passengers. We defined the following three research questions:

RQ1 What aspects do people appreciate during their experience with SPENCER?

RQ2 What features should be added to SPENCER?

RQ3 How does a service, like SPENCER, contribute to passengers' customer satisfaction?

We conducted a case study in which 16 (uninformed) passengers participated, divided into 9 groups which consisted of anywhere between 1 and 3 passengers. Participants were invited to follow the robot while it guided them to a location at the airport. To answer the research questions we collected data through a self-reported questionnaire and a group interview. Additional data was collected through observations from the researchers (Section 8.2.2), of which we will summarize the results (Section 8.3). The case study was conducted during two days. In between these days we made minor adjustments to the measurement instruments based upon the experience we gathered (see Section 8.2.2).

8.2.1 Use case and test scenarios

The use case for the SPENCER project was defined as follows: the robot provides guidance to passengers with a minimum connecting time. This target group does not include those passengers who would have to run in order to catch their flight, as the robot would be unable to safely reach the speeds required to guide them. We realized that it would technically not be possible for the robot to pass the Schengen filter and use elevators - at least not without the assistance of people. Therefore, the use case was scaled down, by collecting passengers after the Schengen filter, and thereafter guiding them to their departure gate.

We designed a study in which SPENCER would guide passengers through several parts of the terminal, varying in (passenger) crowdedness and availability of shops. Together, these situations made for a representative selection of the situations a transfer passenger (and a robot guiding him/her) could encounter at Schiphol. This study

Figure 8.3: Map of Schiphol: B- and C-gates and Lounge-1. The robot guided passengers from the lounge-1 (T) or Starbucks (S) to the gate B-18 (G) - red dashed line, or from gate B-18 to Lounge-1 (L) - black dashed line. Legend: ■shops and facilities; ■Schengen area; data recorded; ■Schengen area; no data recorded; ■no data recorded.

featured two similar test scenarios; the first test scenario featured passengers being guided to their departure gate, the second scenario passengers being guided from their arrival gate to the lounge area, from which passengers could either go to the baggage claim area or the intercontinental flights. In both scenarios SPENCER guided through the more crowded Lounge-1, the junction near Transfer-2 and the corridor between B-gates and Lounge-1, which can be characterized as a long hallway, featur-

ing moving walkways and static objects such as columns, benches and artworks.

In the first test scenario SPENCER would pick up passengers in Lounge-1 or at the Starbucks (denoted by S in Figure 8.3), and guide them to gate B18. This scenario required passengers to interact with the robot using a test boarding card we provided to them for the purpose of the study (Figure 8.4).

Figure 8.4: The first scenario required passengers to identify themselves using their test boarding card (left), after which they were guided to gate B18 (right).

The second scenario was one where SPENCER picked up passengers at gate B18 and guided them to Lounge-1, specifically to the stairs which people had to take in order to get to the baggage hall. In this scenario passengers did not have to use a boarding card to interact with the robot, they used the touchscreen display on the robot (Figure 8.5). For both test scenario's we collected the same data from the passengers.

Figure 8.5: In the second scenario the robot guided passengers from the B-gates towards the baggage hall in Lounge-1.

8.2.2 Data collection methods & measures

We used different data collection methods during the case study - each with their own advantages and disadvantages (Gray, 2014). The data was collected on two separate days; on the second day we added a fourth data collection method. Given that the goal was to better understand a phenomenon this practice, dubbed controlled opportunism, was applicable (Eisenhardt, 1989). Similarly we added two questions to the questionnaire on the second data collection day. We made use of the following data collection methods, of which we will describe the measures in further detail below.

1. Feedback questionnaire (individual, self-reported)

2. Interview (group)
3. Video recordings of the interaction
4. Notes taken by one of the researchers during guiding²

The feedback questionnaire consisted of questions related to participants' journeys, the group size and participants' gender and age. Furthermore, the questionnaire contained eight 7-point Likert scaled statements about the appearance and behavior of the robot. The topics of these statements were based upon previous work in HRI (e.g. Weiss et al., 2015; Heerink et al., 2006) and assessed participants' satisfaction with the SPENCER's *speed*, *driving behavior* and *feedback*. After the first test day we included two additional questions about the robot, as well as a question on participants' opinion of robots in general. Both questionnaires can be found in Appendix J.

Semi-structured interviews (Appendix J) were conducted in order to clarify and augment quantitative results from the questionnaire, and to give the participants the opportunity to mention topics not addressed in the questionnaire. The interview always started with asking participants about their first impression of the experience. The middle block of questions was conducted in a semi-structured manner and addressed topics such as positive & negative experiences and suggestions for improvement. The final block inquired about customer experience and situations at the airport where a robot would be of added value.

Finally, we asked participants in which situations they believed a robot providing services like SPENCER would be interesting for them, and if a wayfinding robot would improve their customer satisfaction.

While participants were following the robot, one of the researchers made a video recording of the events using a handheld camera. This ensured we could re-watch the events that occurred in order to better interpret statements made by participants during the interviews.

On the second test day one of the experimenters accompanied the group and made notes about events participants mentioned they found noteworthy. Before the experiment, participants were encouraged to think aloud about their experiences.

8.2.3 Participants

Seventeen passengers participated in the nine trials. One participant of group #9 followed half of the tour and did not complete the questionnaire, though this participant did participate in the interview afterwards. This participant's interview data has been included in the analysis. As such, the sample consisted of 16 participants: 10 males and 6 females, aged between 26 and 54 ($M=36.81$, $SD=9.01$). The group composition was as follows: five single participants, two dyads and two triads. Nine participants indicated the purpose of their journey was business, the other seven participants traveled for other reasons. Information about the group distribution is provided in Table 8.1.

²We added this measure for the user studies on the second test day as we found out on the first test day that because the distance traveled was quite long, participants forgot specific positive or negative events which occurred during the guiding.

#	Test day	Time	Group size	Scenario	Track	Distance (m)	Duration (mm:ss)
1	1 (Sunday)	12:00	1	1	Transfer 2 - Gate B18	461.6	07:36
2	1 (Sunday)	12:40	2	1	Transfer 2 - Gate B18	393.7	08:05
3	1 (Sunday)	13:50	1	1	Transfer 2 - Gate B18	431.2	14:41
4	2 (Wednesday)	13:00	3	1	Starbucks - Gate B18	377.4	06:35
5	2 (Wednesday)	13:20	3	2	Gate B18 - Lounge-1	473.6	10:17
6	2 (Wednesday)	13:55	2	1	Lounge-1 - Gate B18	439.5	16:19
7	2 (Wednesday)	15:20	2	1	Starbucks - Gate B18	407.1	06:44
8	2 (Wednesday)	15:55	1	2	Gate B18 - Lounge-1	467.9	08:09
9	2 (Wednesday)	16:45	1 ³	1	Starbucks - Gate B18	413.5	06:33

Table 8.1: 16 participants divided over 9 groups, participated in the user evaluation at Schiphol.

8.2.4 Procedure

Airport passengers were approached by both the lead experimenter and a KLM staff member, and invited to participate in a trial with the SPENCER robot. The experimenter informed passengers that SPENCER was developed by the airline and a number of university partners, and funded by the European Commission. We further told them that the goal of the robot was to guide them to a gate (or to the lounge, in the case of the second test scenario). We emphasized that we were interested in the passengers' experience when using the service. Participants were asked to follow the robot, and afterwards complete a brief questionnaire and participate in an interview. The expected duration of a session was between 30 and 45 minutes. As a consequence, only participants who had ample time to fill at Schiphol participated.

Having given consent to participate, detailed instructions about the specific test scenario were provided to the participants. Participants participating in the first test scenario were provided a fake boarding card, and an explanation that the robot was equipped with a boarding card reader, and instructed to use the fake boarding card in order to have the robot guide them to gate B18. Participants in the second test scenario were not provided with a boarding card but informed that they could use the touch screen interface to get the robot to guide them to the baggage hall.

While being guided by the robot, both robot and participants were followed by two experimenters who stayed in close proximity behind the field of view of the participants. One of these experimenters recorded the interaction with a video camera. In addition 3-4 researchers who monitored the technical status of various robot components joined but followed from a further distance from the robot so as not to influence the participants. One of these researchers was always present within a few meters of

the robot with an emergency stop button for the safety of the robot, participants and other passengers. No situations arose during the evaluations which prompted the use of the emergency stop button.

After having arrived at the target destination, participants were led to a quieter area where they could sit down. Here participants first individually completed the self-reported questionnaire, after which the experimenter conducted the interview with the group. This interview was recorded, and took place in the form of a semi-structured interview. During the interview the experimenter provided the participants with ample opportunity to speak their mind about both positive and negative experiences with, or comments related to the robot.

8.2.5 Data analysis

We will report the results of the feedback questionnaire and interviews. Results have been augmented with the researchers' observations, in the form of "we noted that / we observed that [...]". The interviews with the recruited participants were transcribed and coded. In total the participants provided us with 167 statements. A coding scheme was developed by the lead investigator to code (1) the nature of the statement and (2) the topic of the statement. We divided the interview into four parts; each part was coded slightly differently. During the interview we distinguished between the following parts: (a) participants' first impressions, (b) points for improvement, positive and negative aspects, (c) in which circumstances would a service like this be attractive to use, and (d) would a service like this influence customer satisfaction. The full coding scheme can be found in Appendix K. All data was coded by the lead investigator, and three additional coders unaware of the goal of the experiment. In total 70.1% of the data was coded by more than 1 coder, reaching acceptable intercoder agreement for the nature of the statement ($\kappa=0.69$) and topic of the statement ($\kappa=0.73$).

8.3 Results

In this section we will present the results of the case study. We will first present general findings, as gathered through the feedback questionnaire. In the consecutive subsections we will discuss particular topics discussed during the interviews. Where applicable we will illustrate our findings with quotes and photos.

Figure 8.6 shows the results of the feedback questionnaire. Three questions were reformulated during the experiment, therefore we present means for both question formulations, of which details are provided in Appendix J. As can be seen in Figure 8.6 self-reported scores are generally high for all measures; this warrants the first conclusion that participants in general had a positive impression of the robot. There are other reasons for why these ratings are so high; we will discuss these in the discussion (Section 8.4). The items "reliability" and "acted appropriately" scored relatively low; the interview data allows us to interpret this result, which we discuss further on in this section.

Though the sample size is small - the findings are therefore not generalizable - we have conducted *t*-tests to test for differences on the dependent variables. Specifically we have tested for differences in *gender* and *journey purpose*. We found a significant

Figure 8.6: Results of the self-reported questionnaire indicated participants were in general satisfied with SPENCER's performance. Legend: ■ Complete sample (N=16), ■ Test day 1 (N=4), ■ Test day 2 (N=12).

difference of gender on likeability and appearance: females thought the robot was more likeable ($M=7.00$, $SD=0$) than men ($M=5.45$, $SD=1.51$), $t(16)=-2.68$, $p<.05$. In a similar fashion females thought the robot's appearance was significantly more appealing ($M=7.00$, $SD=0$) compared with males ($M=5.91$, $SD=0.83$), $t(16)=-3.43$, $p<.01$.

The interview data has been analyzed using the approaching outlined in Section 8.2.5.

8.3.1 Participants' first impressions

Participants had a positive first impression. Nine out of thirteen participants (69.2%) whom we asked about their first impression of the robot indicated this was a positive impression. Participants indicated SPENCER was a nice, well-designed robot which provided a different experience to the airport. The remaining four participants were neutral. One participant reported to be shocked to see a robot at the airport. Two other participants indicated they did not understand the purpose of the robot immediately, and that it could for example be a security robot as well. The fourth participant indicated he was surprised at first, though he indicated that he was satisfied with the experience after having been guided. Through follow-up questions we gathered information on which specific aspects were positive and negative; we will discuss these in the next section.

8.3.2 Positive and negative experiences while being guided

We discussed the general ease of use of the robot with ten out of sixteen participants. Nine participants were positive about the ease of use, in particular because they only had to follow the robot. For the second test scenario, one participant illustrated this by saying "just push the button". One participant who participated in the first test scenario mentioned that especially in the beginning it was, a little problematic to understand how to use the robot. This turned out to be mainly due to the part where

participants had to use the test boarding card to identify themselves to the robot.

8.3.2.1 Input & output modalities

Three out of twelve participants who participated in test scenario 1 explicitly referred to problems when scanning the test boarding card. From the videos we observed that all 12 participants had issues with holding the card at the correct angle in front of the scanner. In retrospect the participants would have needed more instructions on where to hold the boarding card. Not all passengers thought the issues they encountered led to a negative experience. One of the participants described the action as *“you just have to keep following, you hold up your boarding pass, it reads it, and then it says just follow me. Three steps. Simple.”*. One explanation could be that due to the novelty factor the participants in the study were more forgiving towards the robot than they would otherwise have been, if this had been an operational service.

The feedback provided to the participants by the robot was addressed during interviews with 12 participants. When asked if the amount of feedback provided was adequate, five participants stated that they were satisfied with the amount of feedback the robot provided: the name of the destination, and the time and distance the robot estimated it would still take to reach that destination. The majority of the participants (10) indicated auditory feedback- and communication could, or should, be improved. Three participants mentioned that the robot could use auditory feedback, to warn bystanders to make room; as the robot already used audio when encountering this situation it could be that the level of the audio was not high enough in those situations (crowded situations). One participant who noticed this feature mentioned she believed it was very polite of the robot. Next to audio as a warning signal, participants mentioned that additional auditory cues (speaking) would be an improvement, either for general politeness behaviors (saying good morning), or just because it would be a nice asset as this would be more user friendly.

Participants also indicated that the interaction could be improved through additional visual feedback. Two examples were mentioned: showing the progress on a map displayed on a touchscreen, and the possibility to select an interface with languages other than English. Both were mentioned by one participant as improvement.

8.3.2.2 Speed & driving behavior

As the robot drove in front of the participants to a destination, feedback about the driving behavior was a major topic addressed during the interviews. Participants perception of the speed varied, and ranged from too fast (2) to too slow (1), though most people indicated the speed was good (4), or good for me but too fast for families (2). The variation of these comments can be explained by the different levels of crowdedness in the lounge area. In general participants who followed the robot through the lounge when it was more crowded indicated the robot was too slow in those areas.

When asked about the driving behavior of the robot, participants commented that the robot stopped too often in the lounge area (or any other area where it happened to be crowded). One participant formulated it as follows: *“I can appreciate how Spencer, when there’s someone walks in front of them, stops. In the safety aspect I can appreciate*

that.". The participant indicated that this, coupled with the fact that re-planning took a few seconds, made the robot less suited to operate in crowded areas. Three participants also indicated they would have liked to get feedback each time on why the had robot stopped.

Three topics we addressed to a lesser extent in the interviews were the appearance, trustworthiness and reliability of the robot. The four participants who commented on the appearance were all positive, in that they appreciated the design of the anthropomorphic shell (described as cute). There were concerns about the height of the robot, which one of the participants indicated was too tall. We asked eight participants if they thought SPENCER was reliable; the answers were moderately positive. On the one hand two participants indicated the reliability was "just neutral", on the other hand participants indicated the robot was robust, and appreciated that the robot continued driving as soon as it noticed that it was safe to move forward. Therefore we conclude that participants were moderately positive about the reliability and trustworthiness of the robot.

We addressed the ease of use of the robot in interviews with eight participants, and the answers we received were in general positive. This is interesting because these also included participants who had issues with the boarding card reader. Reasons for why participants indicated it was easy to use the robot were that you "just push the button", "you don't have to look" because you can just follow the robot and "you don't have too much cognitive load". It does therefore seem that participants appreciate a robot which they can follow, and that they are quite tolerant of small failures earlier on in the interaction.

8.3.3 Situations where a robot would be useful

We asked 10 of the participants in which situations a wayfinding robot such as SPENCER would be useful at the airport. All participants agreed that they would not require the guiding services as provided by the SPENCER robot at this time. Four participants indicated they might consider using the robot if they were at an airport they were unfamiliar with, either because they would not visit it often or in the case of fewer signs and directions at that airport. Another participant indicated the robot would be especially useful if it could provide extra information in cases where normal operations are disturbed, for example due to bad weather. Three participants indicated they though the robot would be useful for other people, such as elderly people, or perhaps families with young children. One of the participants could also imagine he might find a wayfinding robot useful when he were older.

Three participants indicated that using the robot would, for them, be more attractive if the robot were capable of offering additional services. Examples of the services mentioned included pointing passengers to restaurants and providing discounts for (e.g.) coffee. There are therefore various moments when a wayfinding robot could be useful: disturbance of normal operations, unfamiliarity at the airport and additional services being provided by the robot to passengers. The participants we recruited for this study did however, with the exception of a novel, new experience, not see any major added value to a wayfinding robot. We will further discuss this in the discussion (Section 8.4).

8.3.4 Customer satisfaction

We received input for future developments of wayfinding robots, and we learned about situations where people believed a robot guide could be beneficial. The final question we asked fourteen participants was whether using a robot would increase their customer satisfaction with the airline.

Twelve out of fourteen participants indicated that a robot like SPENCER would improve their customer satisfaction. Among other things, participants indicated that the robot showed that the airline is not afraid of technology and working on innovation. The two participants who indicated that the robot would not directly improve their customer satisfaction were more or less indifferent, both indicated it would be a nice optional service but nothing more than that. We will discuss the implications of this feedback in Section 8.4.2.

8.4 Discussion

In this chapter we have described the evaluation of SPENCER's real-world deployment at Schiphol airport with respect to technical performance and user experience. The platform was technically safe and the technical components all worked well enough to conduct user tests with real passengers in a guiding scenario. Crowdedness was a challenge for the deployment as it triggered safety measures that made for abrupt stopping and re-planning. This re-planning took time and caused delays, which was reflected in the results gathered from the interviews and feedback questionnaire. During this part of the deployment, SPENCER drove in total 3865 meters autonomously without any critical failures that resulted in aborting trials or having to trigger the emergency stop button. Therefore, we feel that we can conclude that the robot was technically safe.

In this section we will discuss two topics. First, we will discuss the highly positive ratings reported on the feedback questionnaire and reasons why this may have occurred. Secondly, we will discuss the use case we tested and the implications of participants' feedback on the requirements for and added values of robotic technology in the transfer process.

8.4.1 Participant bias

Participants were moderately-positive about the performance of the robot; though the robot stopped often it was deemed reliable, the perception of the driving speed was in general satisfactory, and participants indicated that the robot would improve customer satisfaction. On the other hand, the participants we interviewed indicated that the robot as presented did not provide added value for them: the main reason was because they were familiar with Schiphol airport and did not require the guiding service the robot offered. The positive results of the self-reported questionnaire could be biased because of two reasons: a novelty effect or the Hawthorne effect.

A novelty effect is used to describe a situation where an increased performance is due to the increased interest in the novel technology and not so much due to actual increased performance. Given that this particular technology was introduced for the

first time at an airport, this is likely to have occurred. In HRI the introduction of a robot has led to novelty effects, such as increased attention and interactions with a receptionist robot (Gockley, Bruce, Forlizzi, Michalowski, Mundell, Rosenthal, Sellner, Simmons, Snipes, Schultz, et al., 2005) and a robot tutor in a school (Kanda, Hirano, Eaton, and Ishiguro, 2004). Novelty effects do not occur by definition in HRI; in the studies reported by Šabanović, Bennett, Chang, and Huber (2013) and Hayashi et al. (2007) usage statistics did not show a drop in the number of interactions after the start of use of the robot. What these studies have in common is that the people interacting with the robot did so over a longer period of time - Hayashi et al. (2007) explicitly mentioned a novelty effect could be expected as people used the deployment location (a train station) daily to commute to work. To investigate whether a novelty effect did occur for the SPENCER robot we would have needed to deploy this robot for a more extended duration and collect additional data on for example the actual usage of the robot as opposed to the user experience data we currently collected.

The second effect which could have caused a bias of participants to respond (overly) positive is the observer effect. An observer effect refers to the participant changing his or her behavior - in this case responses - because he or she is conscious of being observed (Saunders, Lewis, and Thornhill, 2007, p. 301). As researchers we have tried to introduce ourselves as a neutral party in the development of the project; we did not indicate that we were actively involved in the development, rather that we were looking for the passengers' insights to further develop the robot. Regardless of our position within the research project, during the experiment, robot and participants were followed by at least 5-6 researchers, and interactions were video recorded. During the interviews we took care to ask open, non-leading questions so as not to influence participants into giving a socially acceptable answer. A solution to the novelty effect is habituation: recruiting participants who are familiar with the object of study, mainly due to their exposure over a longer period of time. This could not be achieved due to two reasons: (1) the use case, system and environment were not designed with repeated interactions in mind, and (2) the state of the system was not yet at a level where the system could function without any monitoring.

8.4.2 A niche for wayfinding robots in the airport transfer process

All participants whom we asked about situations where a service as provided by SPENCER would be useful indicated that the current service would not be useful for them. On first sight this could warrant the conclusion that the current state-of-the-art in robotics technology is not yet at a state where a wayfinding robot at a busy international airport provides services that improve passenger way-finding or reduce transfer time. We do feel that there are two arguments to counteract this claim.

The first reason is that participants who accepted our invitation to participate were all willing to participate and mentioned that they were all familiar with the airport. Furthermore, none of the participants indicated during the trial that he/she had limited experience with flying. While each participant was different in that he/she had different reasons for flying and was at the beginning, middle or end of their journey they were all experienced with air travel in general. Furthermore, none of the participants were flying with a family (e.g. with children). The participants were

thus part of a group that could be described as experienced travelers in their early and middle adulthood (Levinson et al., 1986). Participants with other levels of familiarity with flying, a different cultural background and different demographic characteristics (in terms of age and group composition) could have had different perceptions of the value of the current robot.

There is a second reason to counteract the claim that current robotic technology is not yet ready to provide services that improve passenger wayfinding. Three participants indicated that the robot could mitigate the frequent stopping by providing additional services, such as directions to restaurants and bathrooms. Other situations in which a robot could be useful would include those situations where disruptions of normal operations occur.

Therefore we conclude that the passengers who we interviewed were all positive about the general experience with the robot, though under normal circumstances they would not find the robot useful for themselves. The passengers did envision the robot to be useful for instance elderly people, or when they were at airports they are unfamiliar with.

8.5 Chapter conclusion

We have presented a case study in order to come to a better understanding of the circumstances under which a wayfinding robot could provide added value to airport passengers. Sixteen participants, divided into nine groups, participated in the case study where the robot guided them to a location, and afterwards participants completed a questionnaire and participated in an interview.

Our results indicate that a wayfinding robot like SPENCER would improve participants' customer satisfaction. The situation we evaluated however was not optimal for the target user group who participated as they were quite familiar with Schiphol airport. The participants indicated that for them, the robot would not add very much value; however they did indicate they might be inclined to make use of a wayfinding robot if they were at an airport they are unfamiliar with or during disruptions of regular airport services. Likewise they indicated the robot might be useful for people who were unfamiliar at the airport. Providing additional services, such as providing directions to restaurants and bathrooms, and discounts in shops, was also considered a point for improvement.

In this chapter we set out to investigate RQ-V: "Under which circumstances would a wayfinding robot provide added value to airport passengers?". From the real-world deployment at Schiphol airport we gathered interesting insights for the design of robots which are deployed in public spaces. For example, a wayfinding robot such as SPENCER could provide value to airport passengers unfamiliar at the current airport, or not experienced with air travel in general. As technology advances and robots in semi-public spaces become capable of traversing crowded spaces in a more time-efficient way we expect the potential value for robots in public spaces to increase. This concludes the final chapter in which we have presented user studies. In the next chapter we will discuss the contributions and limitations of this research.

9

Discussion & Future Work

"If I let you assimilate technology from the future
there's no telling how events would be altered."
"You're willing to alter the future by getting
Voyager home now."
"Oh, yes, but there's a difference."
"Do what all good pragmatists do, admiral.
Compromise."

*Admiral Janeway, the Borg Queen.
Star Trek Voyager: Endgame, part II*

In this chapter we will summarize the findings and contributions of this thesis, and discuss some of the limitations of the work presented. We will conclude this chapter by presenting avenues for future work.

As posited at the beginning of this thesis, the development of social robots, and the study of human-robot interactions, can be approached from different disciplines. These disciplines include robotics, computer science, psychology and social sciences (Dautenhahn, 2007). In this thesis we set out to investigate two non-verbal behaviors for wayfinding robots: interpersonal distance and gaze. We approached this through the investigation of people's preferences, perceptions and actual behaviors. The work reported in this thesis presents several methodological and empirical contributions to HRI. These contributions were attained by conducting nine user studies exposing people to, and asking them about their perceptions of different behaviors of robots. Over the years leading to this thesis, we have been pragmatic when choosing the type of study: where possible we have conducted field studies, when necessary we conducted lab studies. We also employed research methods which could be considered novel, or perhaps controversial for a number of reasons, such as the use of crowdsourcing, or projective measures. Therefore we will attempt to weigh the contributions against the limitations of the studies.

We will first discuss the main results in relation to the research questions (Section 9.1), followed by a discussion of the limitations of the work presented (Section 9.2).

In the second part of the chapter, we will discuss the implications of these results on theory (Section 9.3) and (the study of) social robots (Section 9.4). We will conclude this chapter with suggestions for future research in the field of Human-Robot Interaction (Section 9.5).

9.1 Discussion of the main findings and contributions

In the first part of this chapter we will discuss the main findings and empirical and methodological contributions of this thesis to the HRI field. Each of chapters 4-8 was associated with one research question (RQ-I - RQ-V).

9.1.1 Dynamics of people and their expectations of robots at airports

The first empirical contribution of this thesis is the insight we gained into the dynamics of the semi-public space investigated: an international airport (Chapter 4). With RQ-I we set out to investigate how people expect a robot to respond when it encounters social situations while guiding a small group of people. In order to answer this question we conducted three user studies, in particular because we first needed to understand the dynamics of an international airport before we could identify (conflicting) social situations that may arise when a robot guides a small group. The contribution of the work reported is two-fold; the insights we gained in both Studies 1 and 3 have separate contributions to the field.

We performed a contextual analysis (Study 1) in which we conducted a systematic observation of the future operating environment of the SPENCER robot - an international airport. Unlike other contextual analysis in HRI (such as Weiss et al., 2011) the analysis we conducted was non-participatory, and therefore the conclusions we drew are solely based on our (the researcher team's) interpretation of the events. The results of this analysis provided insights into passengers behaviors which guided the remainder research. Examples of these insights include passengers behavior when waiting in the lounge area (used in Study 3) and insights into the general transfer process which guided the development of the use case in Study 9. Furthermore, we gained insights for the technical capabilities of the SPENCER robot platform: we observed people (seemingly) organizing themselves in groups of 2-3. This is relevant for the robot's people perception and tracking capabilities as it requires either that the robot platform should have the capability to detect people temporarily occluded behind other people, or it limits the size of groups that can be guided by the robot. This contextual analysis thus provides insight into people's behavior in crowded environments, which may be relevant for researchers investigating (or developing) robots in similar semi-public spaces. As well as these insights, the research is relevant for the broader HRI community as it shows why the study of a robot's operating context is not only relevant for the design of use cases and user studies, but rather also for the development of technical components. In SPENCER in particular we conducted a data collection session parallel to the contextual analysis, which data was subsequently used to train and test algorithms in simulation before testing these algorithms at Schiphol airport with the SPENCER robot.

The second contribution to the field is more closely related to the research question, and concerns Study 3 in which we conducted an exploratory study to investigate how people expected a wayfinding robot to behave when encountering conflicting social situations at an airport. As in other studies we asked people how they believed the robot should behave. Unlike Lichtenthäler, Peters, Griffiths, and Kirsch (2013) and Breazeal, DePalma, Orkin, Chernova, and Jung (2013) we did not ask the participants control a real robot platform. We collected free-text responses as we did not want to bias participants' answers and were interested in their spontaneous response (Reja, Manfreda, Hlebec, and Vehovar, 2003). The contribution of this study to the field of HRI is that we have gained a better understanding of people's expectations of a robot's behavior. As far as we have attempted in this research to elicit which social norms should govern the behavior of social robots, and how these might differ from social norms people adhere to, this study comes closest to achieving that by investigating what people believe is acceptable behavior in different situations.

The design choices we made in this study in terms of type of data collected (qualitative data, free-text responses) and study design (number of variables manipulated) are among the reasons why this study was quite costly in terms of resources. We conducted this study as we were unaware of any studies investigating people's preferences for robot actions in such a way. Despite the large amount of (detailed) data, for now we have to conclude that we cannot provide conclusive answers to RQ-I ("How do people expect a robot to respond when it encounters social situations while guiding a small group of people?"). This is in large part due to the large variety in the answers. A solution which would have allowed us to do so, would have been to conduct a follow-up study, and provide participants with a number of situations and responses. Similar to the research of Brauer and Chaurand (2010) we could have asked participants to indicate (on a scale) how deviant a response of a robot would be, and perhaps also how likely they would be to express their disapproval of the behavior of the robot in a particular situation.

In answering RQ-I, we found that people expected a wayfinding robot to adhere to the needs of the majority of the group, e.g. when the situation applies to a larger part of the group, the robot should wait.

9.1.2 Preferences for human-robot approach distance and direction

The second empirical contribution of this thesis concerns the research presented in Chapter 5, which investigated (RQ-IIa) how a robot should approach a small group of people, and (RQ-IIb) whether people from different countries would prefer a robot to approach them differently.

The research on personal space in HRI is primarily concerned with robots approaching individual people as we identified in the literature review in Chapter 2. As 70% of the people in crowded environments in public spaces are part of groups (Moussaïd et al., 2010), we set out to investigate how robots should approach small groups of people. Previous research in psychology found that there are differences in the personal space of people, which could be especially interesting considering the culturally diverse population one can expect to encounter at an international airport. Since this factor has not been investigated in HRI beyond the work of Eresha et al.

(2013), who found that Arabs and Germans have different expectations of the distance between robots and people, we set out to further investigate this factor. The factors groups and culture have been addressed with Studies 4 and 5 reported in Chapter 5, and contribute to respectively RQ-IIa and RQ-IIb.

In Study 4 we found that a robot should enter a group at the earliest opportunity, with which we refer to the fact that a robot should not drive around a group in order to enter at the biggest opening. We observed that the group made room for the robot, and adapted their spatial arrangement in order to include the robot. This behavior could have been an artifact of the context in which this study took place: the participant operating the telepresence robot and the other participants were part of a team. Furthermore, the group was dependent on the participant operating the telepresence robot as he/she carried information necessary to complete the task. Another reason could be that a robot may be allowed to refrain from conforming to the social norms people adhere to when they wish to join a group. Kendon (1990, p. 230-231) describes the complex procedure of joining an existing F-formation; in particular that a person who wishes to join a group stops outside the O-space, and waits until one of the group members makes contact before he/she is allowed to enter the O-space. It is possible that this highly dynamic process exists for robots as well. However, in the current study we observed that the robot operator could approach the group and did not require an invitation to join, rather the group immediately accepted this member and re-arranged their spatial arrangement. For now, we can conclude for RQ-IIa that a robot should join a group at the earliest opportunity, and that rather than a process in which both parties adapt their spatial arrangement, in HRI the people adapt their spatial arrangement to include the robot.

In Study 5 we investigated cultural differences in the perception of human-robot approach distance and direction. In our analysis of preferences for approach behaviors, we found that Chinese participants had a preference for closer approaches over farther approaches, and that preferences from people from the United States were the opposite. Our results are partially in line with the hypothesis of Hall (1966) that contact cultures (e.g. South European, Latin American and Arab cultures) maintain less distance than non-contact cultures (e.g. North European and North American cultures). The method we used is comparable to the work of Eresha et al. (2013) in that we also investigated cultural differences in a static setting. The main limitations of this work is the limited ecological validity as we used a static setting to investigate a dynamic concept. The fact that participants were provided with a view from above removed any information about the height of the actors in the scenes. To summarize, in support of RQ-IIb we found that people from different (national) cultures seem to prefer different types of robot approaches. However, the differences we found were small and therefore future research, mitigating some of the limitations of this study would be needed to provide a more definite conclusion. Provided that there are meaningful results, in this case different preferences for a robot's approach distance and direction. A second question would be how to implement these features on a robot platform. Rather than implementing different behaviors, and going through the technical (and ethical) challenge of detecting a person's cultural background it might be better for a robot to err on the side of caution and come to a halt at the fur-

the most comfortable distance. It could be that people who have a preference for a closer distance would consider the robot as being “cold” (Glick et al., 1988). On the other hand, when a robot does violate the norm of personal space people do compensate for this intrusion (Sardar et al., 2012).

Rather than the design of full body behavior, the third empirical contribution relates to design of a robot’s head turn behavior.

9.1.3 Design of a robot’s head turn behavior

The third empirical contribution of this thesis is the insight we gained into the effects of anthropomorphic robot head directions- and turn behaviors while guiding a small group of people (Chapter 7).

Social robots are designed in different ways, though these designs often include humanoid design elements which are reflected in the design of the robot shell and at times include a head and/or eyes. Examples of guide robots with these elements include the Robovie, FROG, and Nao platforms. The goal of Studies 7 and 8, described in Chapter 7 was to get a deeper understanding of people’s preferences for- and perceptions of a guide robot’s head turn behavior, specifically through RQ-IV (“What head turn behaviors are appropriate for a robot guiding small groups?”).

In Study 7 we found that a robot with an anthropomorphic head and eyes should face towards the driving direction while guiding a small group. In Study 8 we tested the perception of different head turn behaviors. Similar to Srinivasan and Takayama (2016) we crowdsourced people’s perception of different head turn behaviors in order to explore different types of head turn behaviors. We then implemented the highest-rated turn behavior on a robot. The results of Study 8 hint towards a more complex answer, and therefore instead of a partial answer to RQ-IV this answer could better be interpreted as a direction for future work. Looking back towards the participants was being appreciated by the participants, or at least, intentions relating to “caring for the group” were attributed. We believe that this result complements the results of Shiomi et al. (2010); robots should gaze at people in order to attract their attention, and gaze towards the driving direction when guiding, just like people do when walking (f.e. Hicheur et al., 2005; Hollands et al., 2002). What is currently unknown yet is whether this head turn behavior could replace subtle unconscious head turn behaviors people use when they communicate their future locomotion intentions to other people (Hollands et al., 2002), or alternatively that a robot should use this head turn behavior only on straight walking paths. When a robot does look back occasionally, participants interpreted this behavior as caring, thus in a positive way.

9.1.4 Added value of wayfinding robots at airports

The fourth and final empirical contribution of this thesis is the insights we gained into the circumstances under which current robotic technology could provide additional value to airport passengers (Chapter 8).

In Study 9 we set out to investigate under which circumstances a wayfinding robot could provide added value to airport passengers. To that end we conducted a case study in which participants used the SPENCER robot to get to a location at

Schiphol airport. The feedback provided by the participants provided guidelines for the design of future wayfinding robots in airports, and general semi-public spaces. For example, participants believed that a wayfinding robot could be useful for airport passengers unfamiliar with the current airport, or not experienced with air travel in general. As technology advances and robots in semi-public spaces become capable of traversing crowded spaces in a more time-efficient way we expect the potential value for robots in public spaces to increase. A limitation of the study was that the participants who agreed to participate in the study were all familiar with air travel and Schiphol airport. Despite this sample bias we believe the results provide valuable insights for the (technical and behavioral) design of future social robots in public space, such as the circumstances under which people expect a robot to be useful.

9.1.5 Preferences for personal space in HRI: behavioral and projective measures

The final contribution of this thesis is a methodological contribution. Specifically, we gained insight into the applicability and limitations of projective measures to assess patterns in preferences for personal space in HRI (Chapter 6).

In Study 6 we investigated the applicability of projective measures to investigate trends in personal space for robots; we compared three measurement methods and found high correlations between the three methods (two projective measures and one behavioral measure). The two projective measures did not yield exactly the same results as the behavioral (real-life) measure, though the direction of the result (participants in the in-group position themselves slightly closer to the robot as compared to the those in the out-group, *n.s.*) was similar. As such, we (carefully) conclude from these results that projective measures could be used in HRI as a tool to initially explore changes in personal space when manipulating a variable - here we manipulated group membership. It is possible that the between-subjects manipulation (group membership) was not manipulated enough; for example in the work of Nass et al. (1994) participants worked on a collaborative task with the robot. In our experiment there was no task involved - except for approaching the robot. We made the decision to not include a task as the main goal was to compare the three measurement methods which were presented on different media (e.g. two in real life, one on a laptop).

This research contributes to a better understanding of the limitations and applicability of projective measures to assess personal space in HRI. The results indicate that projective measures could be used to explore changes in personal space when an independent variable was manipulated. However, we cannot completely support RQ-III as there were large differences between the three measures (on average 15 centimeters). We therefore conclude that projective measures can be used to investigate the direction of effects of independent variables on personal space.

In Section 9.1 we have discussed the main findings and contributions of the research reported in this thesis. We have answered the research questions. In the remainder of this chapter we will first discuss some of the broader limitations of this work, after which we will discuss the implications of this work for theory (including research into HRI (interpersonal) distance and gaze) and the study of social robots.

9.2 Limitations

The results of the studies presented in this thesis provide the field with insights into the use of non-verbal behaviors for wayfinding robots. There are limitations to this work which go beyond the individual studies. In this section we will reflect on three of those limitations.

9.2.1 Sample bias

One of the points we mentioned in Chapter 3 was a tendency of research conducted in Western countries to recruit participants who belong to Western, Educated, Industrialized, Rich, and Democratic (WEIRD) countries (Henrich et al., 2010). We conducted six studies in which we recruited participants - with the exception of Study 2 - to interact with robots. Basic demographic information has been summarized in Table 9.1¹. As can be seen, the samples can in general be classified as student populations. This is therefore a limitation in that the results may apply to a small percentage of the world population.

Because of this the studies reported in this thesis have limited external validity. This may be similar to most lab experiments, such as studies in social psychology, which have relied on students as well (Henry, 2008). We think there are two dimensions to this convenience sample. On the one hand there is age, and on the other hand the primarily technical background of the participants. The question of the relevance of age of the participants is interesting because Ezer, Fisk, and Rogers (2009) found that younger and older adults have similar expectations of robots as long as they have comparable experience with technology. In order to address the challenge of sample bias we have crowdsourced data on several occasions to gather data from a diverse sample.

Study	N	Age range	M (SD)	Male / Female	Western nationality	Background
2	39	19-59	27.50 (8.80)	21/18	-	-
4	56	18-32	20.09 (2.20)	43/13	92.9%	100% students
6	68	19-59	24.81 (6.41)	46 / 22	84.3%	79.4% students
7	25	17-40	23.76 (5.93)	14/11	88.0%	-
8	68	18-35	24.56 (3.81)	50 / 18	70.0%	95.6% students
9	16	26-54	36.81 (9.01)	10 / 6	-	-

Table 9.1: Basic demographic information of the samples of studies 2, 4, 6, 7, 8 and 9. Studies 2-8 involved participants in lab settings. Study 9 was conducted at Schiphol airport.

¹We define “Western countries” as those countries belonging to the Anglosphere (a set of countries from the British Isles which maintain close political cooperation) and/or the European Single Market. These countries include the United Kingdom, Ireland, the United States, Canada, Australia, New Zealand, Iceland, Norway, Sweden, Finland, Denmark, Belgium, the Netherlands, Germany, France, Austria, Hungary, Switzerland, Spain, Portugal, Italy, Greece, Estonia, Latvia, Lithuania, Poland, Slovenia, Croatia, Czech Republic, Romania, Bulgaria, Malta, and Cyprus. Source: Wikipedia: Western World. Retrieved from: https://en.wikipedia.org/wiki/Western_world. Last accessed 24 April 2017.

9.2.2 Crowdsourcing

We chose to use crowdsourcing in Studies 3 and 5, as it would allow us to gauge the attitudes of people from different countries in a resource-efficient way, without actually compromising ecological validity (Baldassare and Feller, 1975). In retrospect there were several advantages and disadvantages, which we have previously discussed in Joosse, Lohse, and Evers (2015b). Crowdsourcing allowed us to collect data from diverse samples, and the samples were indeed more varied in terms of basic demographics (Table 9.2) than the samples of our lab studies. In that sense one could argue that these two studies would have higher external validity as compared to typical lab studies.

Study	N	Age range	M (SD)	Male / Female	Western nationality	Background
3	118	17-66	32.33 (11.99)	72/45	55.9%	16.9% students
5	181	16-73	37.50 (12.54)	93/88	47.5%	14.4% students

Table 9.2: Basic demographic information of the samples of studies 3 and 5.

However, crowdsourcing did not turn out to be the holy grail. Collecting data from the United States was very straightforward and took a few hours. Collecting data from China took significantly longer. Quality control was a laborious task, which turned out to go beyond standard control questions in a post-experiment survey. We were forced to exclude up to 47% of the data, and this was easier in Study 3 as compared to Study 5, as we could look at the qualitative data in Study 3, and identify cases where participants provided the same answer of a meaningless sequence of alphanumerical characters. On the other hand, in Study 5 the data consisted of quantitative data (numbers) and it was more challenging to identify outliers (and fake data).

Therefore crowdsourcing does have advantages; it allows for a diverse sample, and the data can be collected in a short time. The disadvantage we encountered was that quality control turned out to be a very laborious task. The final limitation of this thesis concern the self-reported measures.

9.2.3 Self-reported measures

The third limitation, or challenge we encountered, concerns the self-reported measures we administered². HRI is still a young field, and at the time of data collection there were a limited number of measures specifically designed for HRI. Examples of the questionnaires we used in the studies include the Godspeed questionnaire (Bartneck et al., 2009) and Negative Attitude Towards Robots (NARS) questionnaire (Nomura et al., 2008). When including different measures in a questionnaire, questionnaires can easily become quite long. This is a challenge as long questionnaires reduce response rates and validity (Dillman, Sinclair, and Clark, 1993; Teng, Tseng, Chen, and Wu, 2012). We do not know if this is a bigger challenge when conducting research in public spaces as compared to lab research where participants may perhaps be more accustomed to completing long questionnaires.

²See Appendices A-K.

In order to conduct research in public spaces we feel it is necessary to develop shorter versions of questionnaires. An example could be the mini-IPIP questionnaire which measures 5 personality constructs with 20 items (Donnellan et al., 2006).

In addition to measures developed for HRI we used other validated, or at least previously-used self-reported measures to measure constructs such as personality. The internal validity of these constructs was often weak. We used ?, p. 367's interpretation of Cronbach's alpha (α) to assess whether internal reliability of a scale was high enough for it to be used in the data analysis; in particular scales with a reliability of $\alpha \leq 0.5$ were deemed unacceptable to be used in the data analysis. This limitation especially concerns personality measures. At times this forced us to abandon the specific measure during the data analysis.

In the next section we will discuss the theoretical implications of contributions of this thesis.

9.3 Implications for theory

The user studies in this thesis provide various implications for theory. In Chapter 2 we provided the theoretical background and related works related to the user studies presented in this thesis. One of the conclusions we drew was that people anthropomorphize a robot (attribute human-like characteristics) and respond to the actions in a social way. The studies conducted showed support for this hypothesis, both in participants' actions and mental models about the robots.

In Studies 4, 5, and 6 we investigated people's behaviors and expectations of approach distance and direction from (virtual representations of) robots. In line with previous research investigating personal space in HRI we found in Study 6 that most participants would assume a distance to be expected from research in psychology, in the personal space zone (between 40 and 120 centimeters from the robot) when approaching the robot. The results of Study 5 showed that people from different (national) cultures have different preferences for (or opinions of) human-robot approach distances.

In Studies 7 and 8 we investigated participants' perceptions of and preferences for head turn behaviors of a guide robot, while the participants were following the robot. We started with a static head behavior, and eventually tested a more dynamic head turn behavior. In Study 7 we found that people had a strong preference for a forward-facing head with eyes. In Study 8, we found support for the anthropomorphism- and media equation theories: people attributed emotions and intentions to the turning behavior of the head. This implies that when designing the behavior of robots, it is important to mimic human-like behavior, in a way supporting the media equation.

As well as theoretical implications, there are also more practical implications for the study of social robots.

9.4 Implications for the study of social robots

In Chapter 3 we outlined the challenge of realism versus control; whether to conduct lab or field studies. Over the course of this research, we conducted field studies where

possible, and lab studies where necessary, or simply more appropriate.

In Chapter 3 we outlined the studies documented in this thesis, the different approaches (in terms of lab and field studies), and the data we collected. There is a trade-off trade-off between realism and control in studies (Oulasvirta, 2009, p. 92), which became apparent in the studies documented. In general most of the studies documented were conducted in semi-controlled environments.

In Study 9 we conducted a totally different type of study, as we asked people to follow a real robot at the airport. We did not find that people believed a guide robot at the airport was a bad idea. On the contrary. Just not for them, and not now. Participants were satisfied with the performance of the robot, though they did not foresee that the robot would have any benefit for them personally. At least not if it could not offer additional services, such as directions to bathrooms, and suggestions for shops and restaurants. Therefore, the current robot was perceived more as a driving information kiosk, and consequently an interesting question would be to investigate how different user groups would consider a guide robot compared to existing solutions such as a smartphone application. The implication of Study 9 would be a call for more field studies. One of the major challenges when conducting Study 9 was that these types of studies are resource intensive, as the study required quite a large project staff.

A call for more field studies in HRI could go hand in hand with the interest of media and general public in robots. We believe that robotic technology will have a place in our society. The way we envision it would be more or less like the SPENCER platform: driving information kiosks capable of autonomous navigation. Semi-public spaces would be an ideal place to test future robots, as they have well-defined boundaries and have specific purposes which allows for the definition of clear use cases.

This makes it impossible not to start doing research & development in the confinements of a lab. Lab studies have advantages over field studies in terms of control that can be exercised by the experimenter on variables. We do believe that it is important to “get out there”. By “going out there”, thus leaving the comforts and safety of the lab, one will typically encounter more diverse problems, and it will hopefully also help in getting an honest story about the future developments of robots and other technology, including drones and autonomous cars, to society.

9.5 Future work

There is a tendency to always want more, such as more data (usually in the form of experiment participants) and more experiments in general. In this thesis we have presented a number of user studies on a variety of topics, and due to the number of limitations and calls for future work already presented in Chapters 4-8, perhaps this thesis leads to more new questions than it answers. All five contributions outlined in Section 9.1 deserve further research. We will specifically provide two suggestions for (directions of) future work.

The first line of research which would be interesting to further explore is cultural differences in HRI. In this thesis we have primarily investigated national cultural differences in people's attitudes towards robots. It would be of interest to investigate

whether there are cultural differences in people's behavioral response to robots, for example in terms of personal space. Currently the only work which has investigated this with a physically embodied robot is the study conducted by Eresha et al. (2013). Especially considering that cultural differences in human-human personal space exist (such as Sorokowska et al., 2017) it would be of interest to investigate in what way people from different cultures would respond to a larger robot platform.

The second issue to explore further is not so much a line of research but rather a research method. The case study we conducted had a relatively small, homogeneous sample. By conducting studies in the wild over a longer period of time it would be possible to recruit a more diverse sample which would provide deeper insights into future applications (and the technical capabilities required) of social robots in public spaces. Furthermore, it could be that a novelty effect was present in these interactions, as this effect occurred with other robot deployments as well (Gockley et al., 2005; Kanda et al., 2004). A more longitudinal study could provide insights into the long-term usage of SPENCER beyond these first responses to new technology.

In Chapter 10 we will conclude this thesis by summarizing the findings and implications of the work that has been presented.

10

Conclusion

“That’s the future. What a fascinating modern age we live in.”

Jack Aubrey.

Master and Commander: The Far Side of the World
(2003)

In this thesis we have studied people’s expectations of, preferences for and responses to social robots with a wayfinding task at an airport. In particular, we investigated the effects of two non-verbal behaviors on people’s evaluation of robots; namely interpersonal distance and gaze. The research was guided by five research questions. In this chapter we will present the conclusions of this thesis. To do so, we will summarize the results of the research questions.

10.1 Findings

We investigated people’s expectations of a robot’s behavior when encountering conflicting social situations, and found that in general, people expect a robot to continue unless the situation applies to the majority of group members (RQ-I).

When approaching a group of people, for example to join a conversation, a robot can suffice by entering an opening in the P-space of a spatial arrangement, after which the group re-arranges its spatial arrangement (RQ-IIa). In the subsequent study, we found that (national) cultural differences in preferences for human-robot approach distance seem to exist (RQ-IIb), in particular we found that Chinese participants were more comfortable with closer approaches than participants from Argentina and the United States. We acquired these results through a projective measure, where participants were not actually interacting with a physically embodied robot. In a comparison of different measurement methods, we found that projective measures could be used to explore changes when manipulating an independent variable (RQ-III).

For anthropomorphic robots equipped with a head, we investigated in what direction the head should face (and turn) when guiding a small group of people. We

found that a robot's head should face toward the driving direction when guiding a small group of people, though, when a robot occasionally turns back this seemed to be understood as caring (RQ-IV).

Finally we investigated how people respond to and behave when being guided by a wayfinding robot in an actual airport environment (RQ-V). Our results show that participants were positive about the robot's appearance and general behavior. Due to crowding at the airport the robot often stopped. The additional time required to arrive at the target location resulted in participants indicating the robot would be a useful service for in particular for people inexperienced with air travel, or to provide information in situations where disruptions of normal operations occur.

Together these results contribute to a deeper understanding of people's expectations of, preferences for, and responses to non-verbal behaviors of social robots in public spaces, specifically airports. We have extended previous work on personal space in HRI to interactions with (small) groups, and perception of a robot's head movement. The results show that non-verbal behaviors of social robots (in public spaces) should be designed in a human-like way, specifically in terms of which distance to keep from groups of people, and in which direction to gaze when guiding (small) groups. Furthermore, this thesis provides insights into the impact of social robots in society - in particular the tasks a robot could fulfill at airports in terms of technical capability and societal desirability. Finally, this thesis shows that in order to design behaviors for future social robots it is necessary for HRI researchers to study people's interactions with robots in the context of use during the development cycle of a social robot.

10.2 Implications

The work reported in this thesis has several implications for the development and study of robots in public spaces.

For designers and researchers working in HRI, this thesis provides several contributions. The work reported in this thesis contributes to existing work in human-robot personal space, specifically by investigating human-robot approach distances and directions in the context of a robot guiding a small group. Our results indicate that there are (national) cultural differences in people's perceptions of human-robot approach distances. Secondly, our work contributes to the understanding of people's perceptions of an anthropomorphic robot's head turn behavior. In particular an anthropomorphic robot's head should face towards the driving direction while guiding a small group of people, while at the same time turning backwards towards the participants was interpreted by participants as "caring for them". Thirdly, our research implies that projective measures could be used to approximate people's responses to behavioral measures, at the same time that these measures do not provide a ground-truth of people's exact response to behavioral measures, in this case in terms of distance. Therefore, we believe that research conducted with projective measures (including crowdsourcing) could complement lab research in the same way that lab studies complement field studies. Perhaps the most important message for researchers is that research for public space could very well start in a lab, though should be conducted in

public space early on in the development cycle to influence the technical development of the robot.

The implications of this work for developers of social robots in public spaces are twofold. Firstly, our interviews showed that the technical realization of a robot platform does not automatically make a specific technology useful (or desirable) for people to use. In addition to providing guiding services, which were deemed useful for particular user groups, people indicated they would have liked to see a robot platform with additional functionalities. Therefore, even when deploying a technically advanced robot platform it seems important to equip this robot platform with a variety of functions targeting different user groups. The second implication for developers of social robots is the need to conduct longitudinal field studies to assess the acceptance of robot technology within its intended context of use.

10.3 The future is bright for social robots and people

Insofar as we can predict the future, we can provide some thoughts on the future of robots in our society. This thesis concerned the behaviors of a social robot in a public space. We conducted studies with robots which are currently not as capable as people, in terms of (e.g.) movement, perception, social intelligence and emotional intelligence. Therefore, robots are not people, robots are machines. Like computers and smartphones, though in this case larger and heavier. With technological advancement not all jobs will be the same, and as such there could be jobs which change as robots start taking over parts of people's jobs. When projecting this view to public spaces, it could be that in a few years time guide robots direct us at busy airports, and instead of walking to a transfer desk, a mobile transfer desk awaits us when we have arrived at an airport, and missed our connecting flight. As well as changing jobs, social robots can lead to the creation of jobs for both lower and higher-skilled workers.

In the end, people will always be people; social beings with continuously evolving needs and norms. In order for social robots to become and remain useful as intelligent tools, people will have to continually investigate people's preferences, perceptions and behaviors of these social robots, through complementary lab and field studies. What a fascinating modern age we live in!

Coding scheme Study 1

Based upon our observations of the video data we constructed a coding scheme. Using this coding scheme we coded actions which took place in (or in the vicinity of) specific locations, such as the narrow and wide hallways.

category	subcategory	#
Narrow Hallway (NH)	1. Pair/Triad exceptions; deviations from normal behavior	2
	2. People walking together with >3 people next to each other	5
	3. People search for the way	27
	4. People re-packing luggage	8
	5. Group waiting in the hallway	18
	6. Individuals waiting in the hallway	2
	7. Passengers evading other passengers who move in the opposing direction	4
	8. Pairs walk in 1 + 1 behind each other	5
	9. Passengers explaining the way to other passengers	1
	10. People are running	3
	11. People overtaking others	2
	12. Passenger sitting on the floor	1
	13. People waving to others	1
	14. Groups organize themselves in pairs	3
Wide Hallway (WH)	1. Pair / Triad exceptions; deviations from normal behavior	n/a
	2. People search for the way	5
	3. People re-packing luggage	7
	4. Groups waiting in the hallway	11
	5. Individuals waiting in the hallway	7
	6. People pass through other groups	3
	7. Groups organize themselves into pairs	7
	8. People are running	20
	9. Pairs walk in 1 + 1 behind each other	13
	10. Passengers watch their luggage while standing still	3

category	subcategory	#
Encounters with staff (ES)	1. People queuing for staff member	1
	2. Staff member looking around to assist passengers	2
	3. Staff member assisting passengers	17
	4. Passengers are getting information from staff at the SSTD	4
	5. Passenger showing documents to staff	7
Using information monitors (IM)	1. Passenger standing next to (1) other passenger (strangers)	4
	2. Passengers form a 2nd row	2
	3. Passenger looking at boarding pass	2
	Information monitors in wide hallway	
	1. Passenger standing next to (1) other passenger (strangers)	30
	2. Passengers form a 2nd row	21
	3. People take the place of others	2
Using moving walkway (CB)	1. Walking people overtake those who are standing still	12
	2. Passengers do not step on moving walkways currently not in use	1
	3. Passengers walking on the moving walkway	72
	4. Group stands still at the right; leaving left side open for passengers to overtake.	1
	5. Standing still on the moving walkway	11
Using SSTM (TD)	1. Passengers approaching machines	8
	2. Passengers using machines	9
	3. Passenger looking in bag to find documents	2
Encounters with vehicles (EV)	1. Airport vehicles driving around passengers	14
	2. People making way for airport vehicles	2
	3. No interference	13
	(undecided)	9

Questionnaire Study 2

Both questionnaires in this study were administered on a tablet. The questionnaires were administered in English.

Questionnaire I

This questionnaire was administered twice as the participants walked two times three minutes. Questions 1-5 were based upon the NASA-TLX questionnaire (Hart and Staveland, 1988), and question 7 was based upon Borg (1982).

Instructions: You have just walked for three minutes. We would know like to ask you to answer 7 questions regarding your experience.

1	tlx1	How physically demanding was the task?	Very low - Very high [7pLik]																																											
2	tlx2	How hurried or rushed was the pace of the task	Very low - Very high [7pLik]																																											
3	tlx3	How hard did you have to work to accomplish your level of performance?	Very low - Very high [7pLik]																																											
4	tlx4	How insecure, discouraged, irritated, stressed, and annoyed were you?	Very low - Very high [7pLik]																																											
5	tlx5	How successful were you in accomplishing what you were asked to do?	Failure - Perfect [7pLik]																																											
6		Would you have been able to walk further if you had been alone	Absolutely not - Absolutely [7pLik]																																											
7		We want you to rate your perception of exertion. This feeling should reflect how heavy and strenuous the exercise feels to you. Look at the rating scale below; it ranges from 1 to 15, where 1 means "no exertion at all" and 15 means "maximal exertion". Try to appraise your feeling of exertion as honestly as possible, without thinking about what the actual physical load is. Your own feeling of effort and exertion is important, not how it compares to other people's.																																												
		<table><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td><td>10</td><td>11</td><td>12</td><td>13</td><td>14</td><td>15</td></tr><tr><td>No exertion at all</td><td></td><td>Extremely light</td><td>Very light</td><td></td><td>Light</td><td></td><td>Somewhat hard</td><td></td><td>Hard (Heavy)</td><td></td><td>Very hard</td><td></td><td>Extremely hard</td><td>Maximal exertion</td></tr></table>															1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	No exertion at all		Extremely light	Very light		Light		Somewhat hard		Hard (Heavy)		Very hard		Extremely hard	Maximal exertion
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15																																
No exertion at all		Extremely light	Very light		Light		Somewhat hard		Hard (Heavy)		Very hard		Extremely hard	Maximal exertion																																

Questionnaire II

This questionnaire was administered as post-experiment questionnaire. Question 3 was based upon the IOS scale (Aron et al., 1991).

1	You have walked with- and without luggage.	☐ We walked farther with luggage ☐ We walked farther without luggage ☐ There was no difference in the distance we walked
2 demo1	During the past 7 days, on how many days were you physically active for a total of at least 60 minutes per day? Add up all the time you spent in any kind of physical activity each day.	(open question)
3 demo2	My estimated physical condition is	☐ Very bad ☐ Fairly bad ☐ Average ☐ Fairly good ☐ Very good
4 demo3	<p>The relationship between two or more people can be represented pictorially by two circles. The more the circles overlap, the closer the person feels towards the other persons. We would like you to select the picture that most accurately reflects how close you would feel towards your fellow group members.</p>	
5 ios		
6 demo4	What is your age?	(open question)
7 demo5	What is your gender?	☐ Male ☐ Female
8 demo6	How tall are you?	(open question)
9 demo7	How much do you weigh?	(open question)
10 fly1	Have you ever flown before?	☐ Yes ☐ No
11 fly2	Did you ever have to transfer at an international airport?	☐ Yes ☐ No
12 fly3	Did you ever miss your connecting flight while transferring?	☐ Yes ☐ No

Questionnaire Study 3

This questionnaire was administered online. The first part of the questionnaire contained the 19 social situations, of which situations 1-10 were used in the analysis reported in Chapter 4. These various situations differed based on *group size*, *time pressure*, *applicability of the situation*, and *involvement of the participant*, which led to a total of 8 conditions (numbered A to H) as described in Section 4.4.2. The questionnaire was administered in English.

Part I. Social situations - descriptions

	Situation	Condition A and E High (A) or low (E) time pressure; situation applies to minority of a small group	Condition B and F High (B) or low (F) time pressure; situation applies to majority of a small group
1a	Fall behind	One of the other group members falls behind and can't keep up with the speed of the robot guide and the rest of the group.	Two of the other group members fall behind and can't keep up with the speed of the robot guide and the rest of the group.
1b	Fall behind (you)	You fall behind and can't keep up with the speed of the robot guide and the rest of the group.	You and one other group member fall behind and can't keep up with the speed of the robot guide and the rest of the group.
2a	Passport	One of the other group members forgot their passport. Without it this person cannot travel and therefore, this person needs to go back to retrieve it.	Two of the other group members forgot their passport. Without it they cannot travel and therefore, they need to go back to retrieve them.
2b	Passport (you)	You forgot your passport. Without it you cannot travel and therefore, you need to go back to retrieve it.	You and one other group member forgot your passports. Without it you cannot travel and therefore, you need to go back to retrieve them.
3a	Store	One of the other group members enters a store, because this person wants to buy a present.	Two of the other group members enter a store, because they want to buy a present.
3b	Store (you)	You enter a store, because you want to buy a present.	You and one other group member enter a store, because you want to buy a present.
4a	Bathroom	One of the other group members urgently needs to go to the bathroom.	Two of the other group members urgently need to go to the bathroom.
4b	Bathroom (you)	You urgently need to go to the bathroom.	You and one other group member urgently need to go to the bathroom.
5	Elderly passes	An elderly person who does not belong to your group is crossing the way of your group. This person walks very slowly and causes the robot to slow down.	A large group of elderly persons, who do not belong to your group, is crossing the way of your group. They walk very slowly and cause the robot to slow down.
6a	Shoelace	One of the other group members has a loose shoelace. The person stops and begins to tie the shoelace.	One of the other group members has a loose shoelace. The person stops and begins to tie the shoelace. One other group member stops because of this.
6b	Shoelace (you)	You have a loose shoelace. You stop and begin to tie your shoelace.	You have a loose shoelace. You stop and begin to tie your shoelace. One other group member stops because of this.
7a	Lost	One of the other group members can't find the robot guide and other group members anymore.	Two of the other group members can't find the robot guide and other group members anymore.
7b	Lost (you)	You can't find the robot guide and other group members anymore.	You and one other group member can't find the robot guide and other group members anymore.
8a	Selfie	One of the other group members stand in the way of the robot to take a selfie with the robot.	Two of the other group members stand in the way of the robot to take a selfie with the robot.
8b	Selfie (you)	You stand in the way of the robot to take a selfie with the robot.	You and one other group member stand in the way of the robot to take a selfie with the robot.
9a	Smoke break	One of the group members wants to have a smoke before getting on the connecting flight. This person therefore goes to one of the smoking areas.	Two of the group members want to have a smoke before getting on the connecting flight. They therefore go to one of the smoking areas.
9b	Smoke break (you)	You want to have a smoke before getting on the connecting flight. You therefore go to one of the smoking areas.	You and one other group member want to have a smoke before getting on the connecting flight. You therefore go to one of the smoking areas.
10a	Meet familiar	One of the other group members meets someone familiar. This person stops and starts a conversation.	Two of the other group members meet someone familiar. They stop and start a conversation.
10b	Meet familiar (you)	You meet someone familiar. You stop and start a conversation.	You and one other group member meet someone familiar. You stop and start a conversation.

	Situation	Condition C and G High (C) or low (G) time pressure; situation applies to minority of a larger group	Condition D and H High (D) or low (H) time pressure; situation applies to majority of a larger group
1a	Fall behind	The minority of other group members falls behind and can't keep up with the speed of the robot guide and the rest of the group.	The majority of other group members falls behind and can't keep up with the speed of the robot guide and the rest of the group.
1b	Fall behind (you)	You and the minority of other group members fall behind and can't keep up with the speed of the robot guide and the rest of the group.	You and the majority of other group members fall behind and can't keep up with the speed of the robot guide and the rest of the group.
2a	Forgot passport	The minority of other group members forgot their passport. Without it they cannot travel and therefore, they need to go back to retrieve them.	The majority of other group members forgot their passport. Without it they cannot travel and therefore, they need to go back to retrieve them.
2b	Forgot passport (you)	You and the minority of other group members forgot your passports. Without it you cannot travel and therefore, you need to go back to retrieve them.	You and the majority of other group members forgot your passports. Without it you cannot travel and therefore, you need to go back to retrieve them.
3a	Store	The minority of other group members enters a store, because they want to buy a present.	The majority of other group members enters a store, because they want to buy a present.
3b	Store (you)	You and the minority of other group members enter a store, because you want to buy a present.	You and the majority of other group members enter a store, because you want to buy a present.
4a	Bathroom	The minority of other group members urgently needs to go to the bathroom.	The majority of other group members urgently needs to go to the bathroom.
4b	Bathroom (you)	You and the minority of other group members urgently need to go to the bathroom.	You and the majority of other group members urgently need to go to the bathroom.
5	Elderly passes in front	A small group of elderly persons, who do not belong to your group, is crossing the way of your group. They walk very slowly and cause the robot to slow down.	A large group of elderly persons, who do not belong to your group, is crossing the way of your group. They walk very slowly and cause the robot to slow down.
6a	Shoelace	One of the other group members has a loose shoelace. The person stops and begins to tie the shoelace. The minority of other group members stops because of this.	One of the other group members has a loose shoelace. The person stops and begins to tie the shoelace. The majority of other group members stops because of this.
6b	Shoelace (you)	You have a loose shoelace. You stop and begin to tie your shoelace. The minority of other group members stops because of this.	You have a loose shoelace. You stop and begin to tie your shoelace. The majority of other group members stops because of this.
7a	Lost	The minority of other group members can't find the robot guide and other group members anymore.	The majority of other group members can't find the robot guide and other group members anymore.
7b	Lost (you)	You and the minority of other group members can't find the robot guide and other group members anymore.	You and the majority of other group members can't find the robot guide and other group members anymore.
8a	Selfie	The minority of other group members stands in the way of the robot to take a selfie with the robot.	The majority of other group members stands in the way of the robot to take a selfie with the robot.
8b	Selfie (you)	You and the minority of other group members stand in the way of the robot to take a selfie with the robot.	You and the majority of other group members stand in the way of the robot to take a selfie.
9a	Smoke break	The minority of group members wants to have a smoke before getting on the connecting flight. They therefore go to one of the smoking areas.	The majority of group members wants to have a smoke before getting on the connecting flight. They therefore go to one of the smoking areas.
9b	Smoke break (you)	You and the minority of other group members want to have a smoke before getting on the connecting flight. You therefore go to one of the smoking areas.	You and the majority of other group members want to have a smoke before getting on the connecting flight. You therefore go to one of the smoking areas.
10a	Meet familiar	The minority of other group members meets someone familiar. They stop and start a conversation.	The majority of other group members meets someone familiar. They stop and start a conversation.
10b	Meet familiar (you)	You and the minority of other group members meet someone familiar. You stop and start a conversation.	You and the majority of other group members meet someone familiar. You stop and start a conversation.

Part II. Demographic information and experience with robots

Following completing of the first part, the second part of the questionnaire contained 49 questions, with which we collected among others demographic information (*demo*), participants recollections of the manipulations (*mc*), flying experience (*fly*), personality (*pers*, (Rammstedt and John, 2007)), experience with robots (*exp*) and participants' horizontal individualistic (*hi*), vertical individualistic (*vi*), horizontal collectivistic (*hc*) and vertical collectivistic (*vc*) orientations (Triandis and Gelfand, 1998).

1	mc1	How many people were following the robot guide (including yourself)?	(open question)
2	mc2	How much time did you have to get to your connecting flight?	Little time - A lot of time [7pLik]
3	fly1	Have you ever flown before?	☐ Yes ☐ No
4	fly2	How often do you fly?	☐ Less then once a year ☐ Once a year ☐ 2-3 times a year ☐ 4-6 times a year ☐ 7-12 times a year ☐ More than once per month
5	fly3	In general, what is the purpose of your flight?	☐ Private ☐ Business
6	fly4	Did you ever have to transfer at an international airport?	☐ Yes ☐ No
7	fly5	Did you ever miss your connecting flight while transferring?	☐ Yes ☐ No

Please, indicate your experience with the following robot categories

8	exp1	Manufacturing robot	No experience with this robot - I have and use this robot [5pLik]
9	exp2	Lawn mowing robot	No experience with this robot - I have and use this robot [5pLik]
11	exp3	Mopping / vacuum cleaning robot	No experience with this robot - I have and use this robot [5pLik]
12	exp4	Guarding robot	No experience with this robot - I have and use this robot [5pLik]
13	exp5	Entertainment robot	No experience with this robot - I have and use this robot [5pLik]

Please, indicate your experience with the following robot categories

14	exp6	Research robot	No experience with this robot - I have and use this robot [5pLik]
15	exp7	UAV	No experience with this robot - I have and use this robot [5pLik]
16	exp8	Using the robot guide is a ... idea?	☐ Good ☐ Bad
17	exp9	Using the robot guide is a ... idea?	☐ Wise ☐ Foolish
18	exp10	Using the robot guide is a ... idea?	☐ Pleasant ☐ Unpleasant
19	exp11	Using the robot guide is a ... idea?	☐ Positive ☐ Negative

20	demo1	What is your age?	(open question)
21	demo2	What is your gender?	☐ Male ☐ Female
22	demo3	What is your nationality?	(open question)
23	demo4	Which of the following categories best describes your primarily area of employment?	(open question)

How well do the following statements describe your personality?

24	pers1	I see myself as someone who is reserved	Disagree strongly - Strongly agree [5pLik]
25	pers2	I see myself as someone who is generally trusting	Disagree strongly - Strongly agree [5pLik]
26	pers3	I see myself as someone who tends to be lazy	Disagree strongly - Strongly agree [5pLik]
27	pers4	I see myself as someone who is relaxed, handles stress well	Disagree strongly - Strongly agree [5pLik]
28	pers5	I see myself as someone who has few artistic interests	Disagree strongly - Strongly agree [5pLik]
29	pers6	I see myself as someone who is outgoing, sociable	Disagree strongly - Strongly agree [5pLik]
30	pers7	I see myself as someone tends to find fault with others	Disagree strongly - Strongly agree [5pLik]
31	pers8	I see myself as someone who does a thorough job	Disagree strongly - Strongly agree [5pLik]
32	pers9	I see myself as someone who gets nervous easily	Disagree strongly - Strongly agree [5pLik]
33	pers10	I see myself as someone has an active imagination	Disagree strongly - Strongly agree [5pLik]

Please, rate the following statements on the given scale.

34	Hi1	I'd rather depend on myself than others	Never / definitely no - Always / definitely yes [9pLik]
35	Hi2	I rely on myself most of the time; I rarely rely on others	Never / definitely no - Always / definitely yes [9pLik]
36	Hi3	I often do "my own thing"	Never / definitely no - Always / definitely yes [9pLik]
37	Hi4	My personal identity, independent of others, is very important to me.	Never / definitely no - Always / definitely yes [9pLik]
38	Vi1	It is important that I do my job better than others	Never / definitely no - Always / definitely yes [9pLik]
39	Vi2	Winning is everything	Never / definitely no - Always / definitely yes [9pLik]
40	Vi3	Competition is the law of nature.	Never / definitely no - Always / definitely yes [9pLik]
41	Vi4	When another person does better than I do, I get tense and aroused. Horizontal	Never / definitely no - Always / definitely yes [9pLik]
42	Hc1	If a coworker gets a prize, I would feel proud.	Never / definitely no - Always / definitely yes [9pLik]
43	Hc2	The well-being of my coworkers is important to me.	Never / definitely no - Always / definitely yes [9pLik]
44	Hc3	To me, pleasure is spending time with others.	Never / definitely no - Always / definitely yes [9pLik]
45	Hc4	I feel good when I cooperate with others.	Never / definitely no - Always / definitely yes [9pLik]
46	Vc1	Parents and children must stay together as much as possible	Never / definitely no - Always / definitely yes [9pLik]
47	Vc2	It is my duty to take care of my family, even when I have to sacrifice what I want.	Never / definitely no - Always / definitely yes [9pLik]
48	Vc3	Family members should stick together, no matter what sacrifices are required.	Never / definitely no - Always / definitely yes [9pLik]
49	Vc4	It is important to me that I respect the decisions made by my groups	Never / definitely no - Always / definitely yes [9pLik]

Coding scheme Study 3

#	Code	Category
1	slow down	slow down
2	speed up	continue
3	ignore	continue
4	continue (guiding to gate)	continue
5	wait / stop - until event is over	wait
6	wait for a certain time - with time limit	wait
7	wait until event is over/continue	wait
8	wait for a certain time/continue	wait
9	provide route information / directions	provide / request information
10	provide flight information (gate etc.)	provide / request information
11	provide time information	provide / request information
12	provide other information (shopping)	provide / request information
13	refer to / get other person (including airline security)	involve other
14	refer to / get other robot	involve other
15	locate a person or place	assist passenger
16	guide people	go with group
17	ask person to get out of the way	continue
18	move around people (implies continuing)	continue
19	leave person /group behind	continue
20	ask person to hurry	continue
21	go back (to solve problem, also if this includes guiding)	assist passenger
22	attract attention (make sound, light)	attract attention
23	follow the users	go with group
24	language unclear	(language unclear)
25	other	(other)
26	attract the attention of others	attract attention
27	ask person to continue / walk on	continue

#	Code	Category
28	participant uncomfortable answering this question	(participant uncomfortable)
29	Offer assistance	Offer assistance
30	Urge people to continue / not do action / deny action	provide / request information
31	Speed up	Other
32	Allow action	Wait
33	User undecided	(other)
34	Ask group for input / what to do	Provide / request information

Questionnaire Study 4

The questionnaire consisted of 49 questions, and was administered on a tablet. The questionnaire was administered in English. Measures included two manipulation checks (*mc*), a measure of copresence (*Copr*), attentional engagement (*eng*) (Biocca and Harms, 2003), person realism (*rea*, (Lombard et al., 2009)) and attitude towards robots (*att*) based upon Heerink et al. (2006). Task load index (*tlx*) was measured using 5 items based upon the NASA-TLX (Hart and Staveland, 1988). Further measures included measures of user experience (*ux*) and demographic questions (*demo*) including personality (*pers*, (Gosling et al., 2003)) and Inclusion of Other in the Self (*ios*, (Aron et al., 1991)).

* indicates this item was reversely coded during analysis.

[5pLik] indicates 5-point Likert-scale.

[7pLik] indicates 7-point Likert-scale.

[7pSem] indicates 7-point semantic differential scale.

Instructions: You have just participated in an experiment in which you have solved a murder together with your group. We would like to ask you a few questions about your experience during this experiment. Please remember that your data can not be linked to your personal identity. If you have any questions, please ask the experimenter.

1	mc1	<p>In the picture above you see four pictures. Which of these pictures best represents how you were standing?</p>	
2	mc2	How many clues were provided during the experiment to solve the murder?	(Open question)

Please indicate the extent to which you agree or disagree with the following statements.

3	Copr1	I felt that <i>the other participants and I / the participant who collected the clues and we were in the same place.</i>	Not at all - to a very high degree [7pLik]
4	Copr2	I believe that <i>the other participants / the participant who collected the clues felt as if we were in the same place.</i>	Not at all - to a very high degree [7pLik]
5	Copr3	I was aware of that <i>the other participants were / the participant who collected the clues was there.</i>	Not at all - to a very high degree [7pLik]
6	Copr4	<i>The other participants were / participant who collected the clues was aware that I was there.</i>	Not at all - to a very high degree [7pLik]
7	Copr5	I hardly noticed <i>the other participants / the participant who collected the clues.</i>	I do not agree at all - I fully agree [7pLik]
8	Copr6	<i>The other participants / participant who collected the clues hardly noticed me.</i>	I do not agree at all - I fully agree [7pLik]
9	Copr7	I felt as if we were in different places rather than in the same place.	I do not agree at all - I fully agree [7pLik]
10	Copr8	I think that <i>the other participants / participant who collected the clues felt as if we were in different places rather than in the same place</i>	I do not agree at all - I fully agree [7pLik]
11	Eng1	I paid attention to <i>the other participants / participant who collected the clues.</i>	Not at all - to a very high degree [7pLik]
12	Eng2	<i>The other participants / participant who collected the clues paid attention to me.</i>	Not at all - to a very high degree [7pLik]

Please indicate the extent to which you agree or disagree with the following statements.

13	Eng3	I tended to ignore the other participants / the participant who collected the clues.	I do not agree at all - I fully agree [7pLik]
14	Eng4	The other participants / participant who collected the clues tended to ignore me.	I do not agree at all - I fully agree [7pLik]
15	Rea1	The other participants / participant who collected the clues looked like they would have done in reality	Not at all - to a very high degree [7pLik]
16	Rea2	The other participants / participant who collected the clues sounded like they would have done in reality	Not at all - to a very high degree [7pLik]
17	Att1	Did you feel uncomfortable talking by means of this telepresence robot?	Not at all - to a very high degree [7pLik]
18	Att2	Do you think it is a good idea to use a telepresence robot?	I do not agree at all - I fully agree [7pLik]
19	Att3	Do you feel at ease with the telepresence robot?	I do not agree at all - I fully agree [7pLik]
20	Att4	If you were to drive this robot / use this robot more often, would you be afraid to make mistakes or break something?	I do not agree at all - I fully agree [7pLik]
21	tlx1	How much physical activity was required? Was the task easy or demanding?	Low - High [10pLik]
22	tlx2	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?	Low - High [10pLik]
23	tlx3	How hard did you try to complete the task successfully?	Not at all - Very hard [10pLik]
24	tlx4*	How successful do you think you were in accomplishing the goals of the task set by the experimenter?	Very successful - Very unsuccessful [10pLik]
25	tlx5	How discouraged, irritated and/or stressed did you feel during the task?	Low - High [10pLik]
26	ux1	Together you have just worked on a task. Could you indicate how much effort you put in, in comparison with your peers?	I contributed significantly less compared with my peers - I contributed significantly more compared with my peers [7pLik]
27	ux2	How well did you operate the robot? / How well did the participant who collected the clues operate the robot?	(open question)
28	ux3	Who committed the murder? You are allowed to have an opinion different from the group discussion	(open question)
29	demo1	What is your gender?	O Male O Female
30	demo2	What is your age?	(open question)
31	demo3	What is your nationality?	(open question)
32	demo4	What is the primary language you speak?	(open question)
33	demo5	What is your age?	(open question)

34	demo6	What is the highest level of education you have completed?	☐ Primary school ☐ High school or equivalent ☐ Vocational / technical school ☐ Bachelor's degree ☐ Master's degree ☐ (Other)
----	-------	--	---

Here are a number of personality traits that may or may not apply to you. Please rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

35	pers1	Extraverted, enthusiastic	Disagree strongly - Agree strongly [7pLik]
36	pers2*	Reserved, quiet	Disagree strongly - Agree strongly [7pLik]
37	pers3*	Critical, quarrelsome	Disagree strongly - Agree strongly [7pLik]
38	pers4	Sympathetic, warm	Disagree strongly - Agree strongly [7pLik]
39	pers5	Dependable, self-disciplined	Disagree strongly - Agree strongly [7pLik]
40	pers6*	Disorganized, careless	Disagree strongly - Agree strongly [7pLik]
41	pers7*	Anxious, easily upset	Disagree strongly - Agree strongly [7pLik]
42	pers8	Calm, emotionally stable	Disagree strongly - Agree strongly [7pLik]
43	pers9	Open to new experiences, complex	Disagree strongly - Agree strongly [7pLik]
44	pers10*	Conventional, uncreative	Disagree strongly - Agree strongly [7pLik]

45	demo7	In which study programma are you enrolled?	(open question)
46	demo8	Do you have experience with robots in general?	☐ I have no experience with robots ☐ I have seen robots before ☐ I have played with robots before ☐ I have worked with robots before ☐ I have built robots myself
47	demo9	A telepresence robot is a robot like the one you have seen in this experiment; which allows you to have a conversation over distance through a robot. How much experience do you have with telepresence robots?	☐ I have never seen telepresence robots before ☐ I have seen a conversation through a telepresence robot, but not in real life ☐ I have seen a conversation through a telepresence robot in real life ☐ I have operated a telepresence robot myself
48	demo10	On average, how often do you play computer games every week?	☐ Never ☐ Once or twice a week ☐ Three - five times a week ☐ Daily ☐ Multiple times a day
49	ios	<p>The relationship between two or more people can be represented pictorially by two circles. The more the circles overlap, the closer the person feels towards the other persons. We would like you to select the picture that most accurately reflects how close you would feel towards your fellow group members.</p> <div style="display: flex; flex-wrap: wrap; justify-content: space-around;"> <div style="text-align: center;"> <p>1</p> </div> <div style="text-align: center;"> <p>2</p> </div> <div style="text-align: center;"> <p>3</p> </div> <div style="text-align: center;"> <p>4</p> </div> <div style="text-align: center;"> <p>5</p> </div> <div style="text-align: center;"> <p>6</p> </div> </div>

Questionnaire Study 5

The questionnaire consisted of 98 questions, and was administered online. The questionnaire was administered in English. Measures included the 36 situations and manipulation checks (*mc*). Negative Attitude Towards Robots was measured using the “Negative attitude towards interaction with robots” (*NarsI*), “Negative attitude towards the social influence of robots” (*NarsS*) and “Negative attitude towards emotional interactions with robots” (*NarsE*) scales (Nomura et al., 2008). Individualism and collectivism was assessed using the 3-item Individual Self-Representation (*ind*) and 4-item Individual Group-Representation (*col*) scales (Brewer and Chen, 2007). Five items were used to measure to what extent participants belonged to a low- or high-contact culture (*bench*) (Høgh-Olesen, 2008). Personality was assessed using the 20-item Mini-IPIP scale which measures the Big Five personality constructs (Donnellan et al., 2006). Finally a number of demographic questions (*demo*) were used to get information about participants’ background.

* indicates this item was reversely coded during analysis.

[5pLik] indicates 5-point Likert-scale.

[7pLik] indicates 7-point Likert-scale.

[7pSem] indicates 7-point semantic differential scale.

Each participant was shown the following 18 situations twice (in a random order). Participants were provided the following instructions: "The robot approached the family and has come to a halt between particular family members at a particular distance. Now it will interact with them." For each situation, participants answered the following question:

1-37	I believe the position of the robot is	Inappropriate - Appropriate [7pLik]
------	--	-------------------------------------

1 / 19	2 / 20	3 / 21	4 / 22
5 / 23	6 / 24	7 / 25	8 / 26
9 / 27	10 / 28	11 / 29	12 / 30
13 / 31	14 / 32	15 / 33	16 / 34
17 / 35	18 / 36		

37	mc1	Please indicate which of the following statements is true.	☐ The robot generally approached from the same direction ☐ The robot generally approached in between the same two people ☐ The robot generally approached from different directions
38	mc2	Please indicate which of the following statements is true.	☐ The robot generally stopped at the same distance ☐ The robot generally stopped at different distances

39

In the picture above, you see 15 numbered tiles. Imagine you know these three people very well. If you were to approach this group of people in order to provide information to them, where would you go and stand? Please choose the number of the tile.

(open question)

To what extent do you agree with the following statements?

40	narsI1	I would feel uneasy if I was given a job where I had to use robots	Strongly disagree - strongly agree [7pSem]
41	narsI2	The word "robot" means nothing to me	Strongly disagree - strongly agree [7pSem]
42	narsI3	I would feel nervous operating a robot in front of other people.	Strongly disagree - strongly agree [7pSem]
43	narsI4	I would hate the idea that robots or artificial intelligences were making judgments about things	Strongly disagree - strongly agree [7pSem]
44	narsI5	I would feel very nervous just standing in front of a robot	Strongly disagree - strongly agree [7pSem]
45	narsI6	I would feel paranoid talking with a robot	Strongly disagree - strongly agree [7pSem]
46	narsS1	I would feel uneasy if robots really had emotions	Strongly disagree - strongly agree [7pSem]

To what extend do you agree with the following statements?

47	narsS2	Something bad might happen if robots de-veloped into living beings	Strongly disagree - strongly agree [7pSem]
48	narsS3	I feel that if I depend on too much, something bad might happen	Strongly disagree - strongly agree [7pSem]
49	narsS4	I am concerned that robots would be a bad influence on children	Strongly disagree - strongly agree [7pSem]
50	narsS5	I feel that in the future, society will be dominated by robots	Strongly disagree - strongly agree [7pSem]
51	narsE1*	I would feel relaxed talking to robots	Strongly disagree - strongly agree [7pSem]
52	narsE2*	If robots had emotions, I would be able to make friends with them	Strongly disagree - strongly agree [7pSem]
53	narsE3*	I feel comforted being with robots that have emotions	Strongly disagree - strongly agree [7pSem]

Please indicate to what extend you agree with the following statements.

54	ind1	I enjoy being unique and different from others in many ways.	Not at all - Very much so [7pSem]
55	ind2	I often do "my own thing".	Not at all - Very much so [7pSem]
56	ind3	I am an unique individual.	Not at all - Very much so [7pSem]
57	col1	Overall, my group memberships have very little to do with how I feel about myself.	Not at all - Very much so [7pSem]
58	col2	The social groups I belong to are an important reflection of who I am.	Not at all - Very much so [7pSem]
59	col3	In general, belonging to social groups is an important part of my self-image.	Not at all - Very much so [7pSem]
60	col4	The social groups I belong to are unimportant to my sense of what kind of a person I am.	Not at all - Very much so [7pSem]

61 bench1

Here is a bench in a city. Someone you don't know is already sitting there. You arrive. On which of these spots would you go and sit?

62 bench2

Here is a bench in a city. Two people you don't know are already sitting there. You arrive. On which of these spots would you go and sit?

63 bench3

Here are two benches in a city. Someone you don't know is already sitting on one of the benches. You arrive. On which of these spots would you go and sit?

64 bench4

You are sitting on a bench in a city. Somebody you don't know arrives. On which of these spots would you prefer the stranger to sit?

65 bench5

You are sitting on one of two benches in a city. Somebody you don't know arrives. On which of these spots would you prefer the stranger to sit?

Please use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age.

66	ipip01	I am the life of the party	Very inaccurate - Very accurate [5pLik]
67	ipip02	I sympathize with others' feelings	Very inaccurate - Very accurate [5pLik]
68	ipip03	I get chores done right away	Very inaccurate - Very accurate [5pLik]
69	ipip04	I have frequent mood swings	Very inaccurate - Very accurate [5pLik]
70	ipip05	I have a vivid imagination	Very inaccurate - Very accurate [5pLik]
71	ipip06	I don't talk a lot	Very inaccurate - Very accurate [5pLik]
72	ipip07	I am not interested in other people's problems	Very inaccurate - Very accurate [5pLik]
73	ipip08	I often forget to put things back in their proper place	Very inaccurate - Very accurate [5pLik]
74	ipip09	I am relaxed most of the time	Very inaccurate - Very accurate [5pLik]

Please use the rating scale below to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age.

75	ipip10	I am not interested in abstract ideas	Very inaccurate - Very accurate [5pLik]
76	ipip11	I talk to a lot of different people at parties	Very inaccurate - Very accurate [5pLik]
77	ipip12	I feel others' emotions	Very inaccurate - Very accurate [5pLik]
78	ipip13	I like order	Very inaccurate - Very accurate [5pLik]
79	ipip14	I get upset easily	Very inaccurate - Very accurate [5pLik]
80	ipip15	I have difficulty understanding abstract ideas	Very inaccurate - Very accurate [5pLik]
81	ipip16	I keep in the background	Very inaccurate - Very accurate [5pLik]
82	ipip17	I am not really interested in others	Very inaccurate - Very accurate [5pLik]
83	ipip18	I make a mess of things	Very inaccurate - Very accurate [5pLik]
84	ipip19	I seldom feel blue	Very inaccurate - Very accurate [5pLik]
85	ipip20	I do not have a good imagination	Very inaccurate - Very accurate [5pLik]

86	demo1	What is your gender?	O Male O Female
87	demo2	What is your birthyear?	(open question)
88	demo3	In which country where you born?	(open question)
89	demo4	What is your primary language?	(open question)
90	demo5	Have you lived abroad for an extended period of time?	O No, or less than 1 year O Between 1 and 2 years O Between 3 and 6 years O More than 6 years
91	demo6	Could you please describe, briefly, with which culture you identify yourself most?	(open question)
92	demo7	Who of your ancestors where also born in [Q88]?	O Many generations of my family O Both my grandparents and parents O Only my parents O One of my parents O None of my ancestors

93	demo8	What is your marital status?	☐ Single ☐ Living with another ☐ Married ☐ Widowed ☐ Separated ☐ Divorced ☐ Would rather not say
94	demo9	How many children do you have?	(open question)
95	demo10	What is the highest level of education you have completed?	☐ Primary school ☐ High school or equivalent ☐ Vocational/technical school ☐ Bachelor's degree (B.Sc., B.A. etc.) ☐ Master's degree (M.Sc., M.A. etc.) ☐ Professional or doctoral degree (MD, JD, PhD etc)
96	demo11	What is the highest level of education your mother has completed?	☐ Primary school ☐ High school or equivalent ☐ Vocational/technical school ☐ Bachelor's degree (B.Sc., B.A. etc.) ☐ Master's degree (M.Sc., M.A. etc.) ☐ Professional or doctoral degree (MD, JD, PhD etc)
97	demo12	What is the highest level of education your father has completed?	☐ Primary school ☐ High school or equivalent ☐ Vocational/technical school ☐ Bachelor's degree (B.Sc., B.A. etc.) ☐ Master's degree (M.Sc., M.A. etc.) ☐ Professional or doctoral degree (MD, JD, PhD etc)
98	demo13	Which of the following categories best describes your primarily area of employment?	(open question)

Questionnaire Study 6

This questionnaire was administered on a laptop after the participants completed all distancing tasks. The questionnaire was administered in English.

The questionnaire consisted of 42 questions. Negative Attitude Towards Robots was measured using the “Negative attitude towards interaction with robots” (*NarsI*), “Negative attitude towards the social influence of robots” (*NarsS*) and “Negative attitude towards emotional interactions with robots” (*NarsE*) scales (Nomura et al., 2008). The perceived relationship between the robot and participant was measured using the Inclusion of Others in Self (*ios*) measure (Aron et al., 1991). Three manipulation checks (*mc*) were used to check if we successfully manipulated group membership and if participants correctly recalled how often they repeated the tasks. Individualism and collectivism was assessed using the 3-item Individual Self-Representation (*ind*) and 4-item Individual Group-Representation (*col*) scales (Brewer and Chen, 2007). Five items were used to measure to what extent participants belonged to a low- or high-contact culture (*bench*) (Høgh-Olesen, 2008). Finally a number of demographic questions (*demo*) were used to get information about participants’ background and their experience with robots (*exp*).

* indicates this item was reversely coded during analysis.

[5pLik] indicates 5-point Likert-scale.

[7pLik] indicates 7-point Likert-scale.

[7pSem] indicates 7-point semantic differential scale.

[10pSem] indicates 10-point semantic differential scale.

To what extend do you agree with the following statements?

1	narsI1	I would feel uneasy if I was given a job where I had to use robots	Strongly disagree - strongly agree [7pSem]
2	narsI2	The word "robot" means nothing to me	Strongly disagree - strongly agree [7pSem]
3	narsI3	I would feel nervous operating a robot in front of other people.	Strongly disagree - strongly agree [7pSem]
4	narsI4	I would hate the idea that robots or artificial intelligences were making judgments about things	Strongly disagree - strongly agree [7pSem]
5	narsI5	I would feel very nervous just standing in front of a robot	Strongly disagree - strongly agree [7pSem]
6	narsI6	I would feel paranoid talking with a robot	Strongly disagree - strongly agree [7pSem]
7	narsS1	I would feel uneasy if robots really had emotions	Strongly disagree - strongly agree [7pSem]
8	narsS2	Something bad might happen if robots developed into living beings	Strongly disagree - strongly agree [7pSem]
9	narsS3	I feel that if I depend on too much, something bad might happen	Strongly disagree - strongly agree [7pSem]
10	narsS4	I am concerned that robots would be a bad influence on children	Strongly disagree - strongly agree [7pSem]
11	narsS5	I feel that in the future, society will be dominated by robots	Strongly disagree - strongly agree [7pSem]
12	narsE1*	I would feel relaxed talking to robots	Strongly disagree - strongly agree [7pSem]
13	narsE2*	If robots had emotions, I would be able to make friends with them	Strongly disagree - strongly agree [7pSem]
14	narsE3*	I feel comforted being with robots that have emotions	Strongly disagree - strongly agree [7pSem]

15 ios

1

Group

You

2

Group

You

3

Group

You

4

Group

You

5

Group

You

6

Group

You

In the figure above you see six pictures. The relationship between two or more people can be represented pictorially by two circles. The more the circles overlap, the closer the person feels towards the other persons. We would like you to select the picture that most accurately reflects how close you would feel towards the robot in general.

16	mc1	Which of the following statements is true?	☐ I wore a blue cap ☐ I wore a red cap ☐ I wore a yellow cap ☐ I wore a green cap ☐ I wore a white cap ☐ I wore a purple cap
17	mc2	How often did you place the 3D figure representing yourself in relation to the robot?	(open question)
18	mc3	Which of the following statements is true?	☐ The robot was part of the blue team ☐ The robot was part of the green team ☐ The robot was part of the yellow team ☐ The robot was part of the red team ☐ The robot was part of the white team ☐ The robot was part of the purple team

Please indicate to what extent you agree with the following statements.

19	ind1	I enjoy being unique and different from others in many ways.	Not at all - Very much so [7pSem]
20	ind2	I often do "my own thing".	Not at all - Very much so [7pSem]
21	ind3	I am an unique individual.	Not at all - Very much so [7pSem]
22	col1	Overall, my group memberships have very little to do with how I feel about myself.	Not at all - Very much so [7pSem]
23	col2	The social groups I belong to are an important reflection of who I am.	Not at all - Very much so [7pSem]
24	col3	In general, belonging to social groups is an important part of my self-image.	Not at all - Very much so [7pSem]
25	col4	The social groups I belong to are unimportant to my sense of what kind of a person I am.	Not at all - Very much so [7pSem]

26	bench1	<p>Here is a bench in a city. Someone you don't know is already sitting there. You arrive. On which of these spots would you go and sit?</p>
----	--------	--

27	bench2	<p>Here is a bench in a city. Two people you don't know are already sitting there. You arrive. On which of these spots would you go and sit?</p>	
28	bench3	<p>Here are two benches in a city. Someone you don't know is already sitting on one of the benches. You arrive. On which of these spots would you go and sit?</p>	
29	bench4	<p>You are sitting on a bench in a city. Somebody you don't know arrives. On which of these spots would you prefer the stranger to sit?</p>	
30	bench5	<p>You are sitting on one of two benches in a city. Somebody you don't know arrives. On which of these spots would you prefer the stranger to sit?</p> <p>During the experiment you approached (a virtual version of) the robot several times.</p>	
31		Do you think this task was intimate?	Not at all - Very much so [10pSem]
32		Do you think this is a pleasant task?	Not at all - Very much so [10pSem]
33	exp1	Do you have experience with robots?	<ul style="list-style-type: none">☐ I have no experience with robots☐ I have seen robots before☐ I have played with robots before☐ I have worked with robots before☐ I have built robots myself

1

2

3

4

34 exp2

Above you see 4 robots. Which of these robots have you seen in real life prior to participating in this experiment?

35	demo1	What is your gender?	☐ Male ☐ Female
36	demo2	What is your age?	(open question)
37	demo3	In which country where you born?	(open question)
38	demo4	What is your primary (mother) language?	(open question)
39	demo5	Have you lived abroad for an extended period of time?	☐ No, or less than 1 year ☐ Between 1 and 2 years ☐ Between 3 and 6 years ☐ More than 6 years
40	demo6	What is your favorite color?	(open question)
41	demo7	What is the highest level of education you have completed?	☐ High school or equivalent ☐ Vocational/technical school ☐ Bachelor's degree (B.Sc., B.A. etc.) ☐ Master's degree (M.Sc., M.A. etc.) ☐ Professional or doctoral degree (MD, JD, PhD etc)
42	demo8	Which of the following categories best describes your primarily area of employment?	(open question)

Questionnaire Study 7

This questionnaire was administered as a pen-and-paper questionnaire after the participants had followed the robot. The questionnaire was administered in English.

The experiment contained three conditions (between-subjects): Eyes_{forward} (E_F), Eyes_{backward} (E_B) and Distance_{backward} (D_B).

The number of questions varied by condition as particular questions were not applicable due to the specific manipulation. The questionnaire assessed participants' general experience (V1 - V12), perception of the physical design (A1 - A5) and usability of the system (U1 - U7). A manipulation check assessed if people correctly identified the direction at which the robot head faced (MC1) and demographic questions (D1 - D8) provided us with background information of the sample. All questions were formulated as 5-point semantic differential scale.

Unless answer options are provided between brackets, the scale was formulated as "I disagree - I agree".

	Question	E _F	E _B	D _B
V1	The robot moved at a comfortable speed	x	x	x
V2	I was at a comfortable distance to the robot	x	x	x
V3	I felt that the distance prediction was accurate			x
V4	I felt it was clear at all times where the robot was heading	x	x	x
V5	I trusted that the robot knew where it was going	x	x	x
V6	I felt that the robot was helpful in guiding me to the destination	x	x	x
V7	I felt that the display was a useful addition to the robot	x	x	x
V8	I would like to see some information during the guiding process	x		
V9	I felt that the direction in which the screen was facing was natural	x	x	x
V10	I felt that the amount of information the robot gave at the waypoints was... [Too little - too much]	x	x	x
V11	I felt that the robot was speaking... [too soft - too loud]	x	x	x
V12	I felt that the robot was speaking... [too slow - too fast]	x	x	x
MC1	The tablet display was on the... [front of the robot - back of the robot]	x	x	x

	Question	E _F	E _B	D _B
A1	I feel that the robot's body design matches its function	x	x	x
A2	The robot's design was intimidating	x	x	x
A3	I felt that the robot's height was appropriate	x	x	x
A4	I feel that the robot's screen design matches its function	x	x	x
A5	I feel that the robot's height was [too short - too tall]	x	x	x
U1	I felt comfortable using this system	x	x	x
U2	The on-screen information provided by this system was clear			x
U3	The on-screen display was aesthetically pleasing		x	x
U4	Overall I am satisfied by this robots performance	x	x	x
U5	This system has all the functions and capabilities I expect it to have	x	x	x
U6	I enjoyed the guiding process	x	x	x
U7	I would use robots like these to guide me in public places	x	x	x
D1	What is your gender? [male / female]	x	x	x
D2	What is your age? [open question]	x	x	x
D3	What is your nationality? [open question]	x	x	x
D4	How tall are you? [open question]	x	x	x
D5	What is the highest level of education you have completed? [open question]	x	x	x
D6	How much experience do you have with robots? [none / I've seen robots / I've interacted with robots / I've programmed robots]	x	x	x
D7	How much experience do you have with social robots? [none - a lot]	x	x	x
D8	How familiar are you with the DesignLab? [not at all - a lot]	x	x	x
	Total number of questions	30	30	32

Questionnaire Study 8

This questionnaire was administered as a pen-and-paper questionnaire after the participants had followed the robot. The questionnaire was administered in English.

The questionnaire consisted of 49 questions and included constructs from the Source Credibility Scale: *Comp* and *Gdwl* (McCroskey and Teven, 1999) and the Anthropomorphism (*Ant*) and Likeability (*like*) measures from the Godspeed questionnaire (Bartneck et al., 2009). Negative Attitude Towards Robots was measured using the “Negative attitude towards interaction with robots” (*NarsI*) and “Negative attitude towards the social influence of robots” (*NarsS*) scales (Nomura et al., 2008). Satisfaction with the robot was measured through 3 items based on Lee et al. (2006b) (*enjoy*). The perceived relationship between the participants was measured using the Inclusion of Others in Self (*ios*) measure (Aron et al., 1991). Additional measures, such as participants’ self-reported engagement (*eng*), flying experience (*fly*) and demographics (*demo*) were not based upon validated questionnaires. Items from the Source Credibility scale, Godspeed questionnaire and Negative Attitude Towards Robots were randomized * indicates this item was reversely coded during analysis.

* indicates this item was reversely coded during analysis.

[5pLik] indicates 5-point Likert-scale.

[7pLik] indicates 7-point Likert-scale.

[7pSem] indicates 7-point semantic differential scale.

[10pSem] indicates 10-point semantic differential scale.

Instructions: You have just participated in an experiment in which you followed the SPENCER robot. We would like to ask you a few questions about your experience during this experiment. Please keep in mind that there are no right or wrong answers, and that your data can not be linked to your personal identity. If you have any questions, please ask the experimenter. This questionnaire consists of 50 questions. After completion of the questionnaire, a short debriefing will take place.

1	Comp1	Intelligent - Unintelligent	[7pSem]
2	Comp2*	Untrained - Trained	[7pSem]
3	Comp3	Inexpert - Expert	[7pSem]
4	Comp4*	Uninformed - Informed	[7pSem]
5	Comp5	Incompetent - Competent	[7pSem]
6	Comp6	Stupid - Bright	[7pSem]
7	Gdwl1*	Doesn't care about me - Cares about me	[7pSem]
8	Gdwl2	Doesn't have my interests at heart - Has my interests at heart	[7pSem]
9	Gdwl3	Self-centered - Not self-centered	[7pSem]
10	Gdwl4	Unconcerned with me - Concerned with me	[7pSem]
11	Gdwl5	Insensitive - Sensitive	[7pSem]
12	Gdwl6	Not understanding - Understanding	[7pSem]
13	Ant1	Fake - Natural	[7pSem]
14	Ant2	Machinelike - Humanlike	[7pSem]
15	Ant3	Unconscious - Conscious	[7pSem]
16	Ant4	Artificial - Lifelike	[7pSem]
17	Ant5	Moving rigidly - Moving elegantly	[7pSem]
18	Like1	Dislike - Like	[7pSem]
19	Like2	Unfriendly - Friendly	[7pSem]
20	Like3	Unkind - Kind	[7pSem]
21	Like4	Unpleasant - Pleasant	[7pSem]
22	Like5	Awful - Nice	[7pSem]

To what extend do you agree with the following statements?

23	enjoy1	I thought following the robot was fun	Not at all - absolutely [7pSem]
24	enjoy2	I thought following the robot was enjoyable	Not at all - absolutely [7pSem]
25	enjoy3	I thought following the robot was entertaining	Not at all - absolutely [7pSem]
26	eng1	If I had to follow this robot for another five minutes, I would do so	Not at all - absolutely [7pSem]

To what extent do you agree with the following statements?

27	narsI1	I would feel uneasy if I was given a job where I had to use robots	Strongly disagree - strongly agree [7pSem]
28	narsI2	The word "robot" means nothing to me	Strongly disagree - strongly agree [7pSem]
29	narsI3	I would feel nervous operating a robot in front of other people.	Strongly disagree - strongly agree [7pSem]
30	narsI4	I would hate the idea that robots or artificial intelligences were making judgments about things	Strongly disagree - strongly agree [7pSem]
31	narsI5	I would feel very nervous just standing in front of a robot	Strongly disagree - strongly agree [7pSem]
32	narsI6	I would feel paranoid talking with a robot	Strongly disagree - strongly agree [7pSem]
33	narsS1	I would feel uneasy if robots really had emotions	Strongly disagree - strongly agree [7pSem]
34	narsS2	Something bad might happen if robots developed into living beings	Strongly disagree - strongly agree [7pSem]
35	narsS3	I feel that if I depend on too much, something bad might happen	Strongly disagree - strongly agree [7pSem]
36	narsS4	I am concerned that robots would be a bad influence on children	Strongly disagree - strongly agree [7pSem]
37	narsS5	I feel that in the future, society will be dominated by robots	Strongly disagree - strongly agree [7pSem]

38	demo1	What is your gender?	☐ Male
			☐ Female
39	demo2	What is your age?	(open question)
40	demo3	What is your nationality?	(open question)
41	demo4	In which study program are you enrolled?	(open question)
42	fly1	Have you ever flown before?	☐ Yes
			☐ No
43	fly2	Did you ever have to transfer at an international airport?	☐ Yes
			☐ No
44	fly3	Did you ever miss your connecting flight while transferring?	☐ Yes
			☐ No

45 ios

<p>In the figure above you see six pictures. The relationship between two or more people can be represented pictorially by two circles. The more the circles overlap, the closer the person feels towards the other persons. We would like you to select the picture that most accurately reflects how close you would feel towards the robot.</p> <div><div>1</div><div> </div></div> <div><div>2</div><div> </div></div> <div><div>3</div><div> </div></div> <div><div>4</div><div> </div></div> <div><div>5</div><div> </div></div> <div><div>6</div><div> </div></div>		
46 open1	What did you like about the robot's behavior?	(open question)
47 open2	Which aspect of the guiding do you think should be improved before deploying this robot at an airport?	(open question)
48 open3	Would you consider the robot, as you saw it right now, to be a socially aware robot?	(open question)
49 open4	Do you have any comments and/or suggestions regarding the behavior of the robot if it were to guide you at an airport?	(open question)

Questionnaires Study 9

Data was collected through a semi-structured interview and a self-reported questionnaire. Participants on the first test day (groups 1-3) completed Questionnaire I. At the end of this test day we made two adjustments to the questionnaire, which are reflected in Questionnaire II administered to participants on the second test day (groups 4-9). Both questionnaires were administered on paper in English.

Semi-structured interview

The semi-structured interview was conducted by one of the experimenters - in Dutch for groups consisting solely of Dutch-speaking participants, and in English for all other groups. The questions can be divided into four parts, which address (a) participants' first impression, (b) points for improvement, positive and negative aspects, (c) circumstances under which a service like SPENCER would be attractive to use, and (d) the impact of a guide robot on participants' customer satisfaction. The interview has been coded using the coding scheme in Appendix K.

1	A	What is the first thing that comes to mind when you describe the experience you had with the Spencer robot?
2	B	What, in your opinion, was positive about the Spencer robot?
3	B	What did you not like about the Spencer robot?
4	B	What is your opinion of the reliability of the robot?
5	B	To what extent was the robot intuitive to use, and what could be improved?
6	B	What is our opinion of the speed of the robot?
7	B	Were there moment that you found the robot rude of polite? Please detail?
8	C	If available, would you use this service again in the future?
9	C	Can you tell in which situations such a service would be attractive to you?
10	D	Does this service impact our satisfaction with KLM, in what way?
11	B	Is there anything else you would like to give us feedback about concerning improving the Spencer robot?

Questionnaire I

This questionnaire was used on the first test day with groups 1-3. Unless answer options are provided, questions were administered as open questions.

[7pSem] indicates 7-point semantic differential scale.

Instruction: Thank you for your interest in the SPENCER robot. SPENCER is a KLM project in collaboration with the European Commission. The goal is to develop a robot that can guide passengers from A to B. With this survey we'd like to find out how this robot is being perceived, and which aspects of the robot could be improved. Thank you in advance for completing this questionnaire.

1	At which airport did you start your journey today?	(open question)
2	Which airport are you traveling to?	(open question)
	Including yourself, how many people are traveling in your immediate group?	(open question)
3	Is your trip for business or pleasure?	☐ Business ☐ Pleasure
4	What is your age?	(open question)
5	What is your gender?	☐ Male ☐ Female

To what extend do you agree with the following statements?

6	I think the robot is very likeable	Not at all - Very much so [7pSem]
7	I think the robot has a very appealing design	Not at all - Very much so [7pSem]
8	I think the robot the robot is intelligent	Not at all - Very much so [7pSem]
9	It was easy to use the robot	Not at all - Very much so [7pSem]
10	It was easy to understand the robot's intentions	Not at all - Very much so [7pSem]
11	My experience with the robot was very positive	Not at all - Very much so [7pSem]
12	The robot acted appropriate	Not at all - Very much so [7pSem]
13	This service improves my customer satisfaction	Not at all - Very much so [7pSem]
14	If you have any further comments, please write them in this box	(open question)

Questionnaire II

This questionnaire was used on the second test day with groups 4-9.

* indicates this item was reversely coded during analysis.

[7pSem] indicates 7-point semantic differential scale.

Instruction: Thank you for your interest in the SPENCER robot. SPENCER is a KLM project in collaboration with the European Commission. The goal is to develop a robot that can guide passengers from A to B. With this survey we'd like to find out how this robot is being perceived, and which aspects of the robot could be improved. Thank you in advance for completing this questionnaire.

1	At which airport did you start your journey today?	(open question)
2	Which airport are you traveling to?	(open question)
	Including yourself, how many people are traveling in your immediate group?	(open question)
3	Is your trip for business or pleasure?	☐ Business ☐ Pleasure
4	What is your age?	(open question)
5	What is your gender?	☐ Male ☐ Female

To what extend do you agree with the following statements?

6	I think the robot is very likeable	Not at all - Very much so [7pSem]
7	I think the robot has a very appealing design	Not at all - Very much so [7pSem]
8	I think the robot the robot is intelligent	Not at all - Very much so [7pSem]
9	It was easy to understand the robot's intentions	Not at all - Very much so [7pSem]
10	My experience with the robot was very positive	Not at all - Very much so [7pSem]
11*	The robot is unreliable	Not at all - Very much so [7pSem]
12	The robot acted appropriately	Not at all - Very much so [7pSem]
13	This service improves my customer satisfaction	Not at all - Very much so [7pSem]
14	The robot functioned well	Not at all - Very much so [7pSem]
15*	It was difficult to use the robot	Not at all - Very much so [7pSem]
16	What is your opinion of robots in general?	Very negative - Very positive [7pSem]
17	If you have any further comments, please write them in this box	(open question)

Coding scheme Study 9

After transcribing the contents of the semi-structured interviews (Appendix J) were coded by two coders (70.1% overlap) on the nature ($\kappa=0.688$) and topic of the statement ($\kappa=0.734$). The coding scheme was designed in a way to distinguish between the four topics of the interview: (a) participants' first impression, (b) points for improvement, positive and negative aspects, (c) circumstances under which a service like SPENCER would be attractive to use, and (d) the impact of a guide robot on participants' customer satisfaction.

1	Instructions scheme statements A: For each of the sentences, please indicate if the tone of the answer is positive, negative or neutral.	
pos	Positive tone	
neut	Neutral tone	
neg	Negative tone	
1	Instructions scheme statements B: For each of the sentences, please indicate if the answer is a positive remark, negative remark, improvement for the robot, or neutral remark.	
pos	Positive tone	
neut	Neutral tone	
neg	Negative tone	
2	For each of the sentences, please choose the category which best describes the content of the sentence. Please write down the letter or the name of the category	
A	Interaction - feedback provided by the robot on process (f.e. progress or use of boarding card reader)	
B	Interaction - (type of) content & information provided to the user	
C	Interaction - channels through which the robot communicates to the user	
D	Motion behavior - examples of rude and polite behavior	
E	Motion behavior - speed	
F	Motion behavior - stopping: avoidance of objects & person. Also: re-planning after stopping	

G	Appearance of the robot shell
H	Trustworthiness of the robot
I	Reliability of the robot
J	Ease of use of the robot in general
K	General comment about operating environment

1	<u>Instructions scheme statements C:</u> For each of the sentences, please choose the category which best describes the content of the sentence. Please write down the letter of the category
A	For other people
B	For other people - specifically elderly people or with mobility challenges
C	For other people - specifically who don't speak English very well
D	During disruptions of normal operations (e.g. delays due to bad weather, strikes)
E	Gates areas, as there's not so many people around
F	For people (including myself) - at an airport I'm unfamiliar with
G	For people (including myself) - when I get older or experiencing mobility challenges
H	Myself in general

1	<u>Instructions scheme statements D:</u> For each of the sentences, please indicate if the tone of the answer is positive, negative or neutral.
pos	Positive tone
neut	Neutral tone
neg	Negative tone

Bibliography

- Aarts, H. and Dijksterhuis, A. (2003). The Silence of the Library: Environment, Situational Norm, and Social Behavior. *Journal of Personality and Social Psychology*, 84(1):18–28.
- Aiello, J. R. (1987). Human spatial behavior. volume 1, pages 389–504. New York: John Wiley & Sons.
- Argyle, M. (1973). *Social interaction*, volume 103. Transaction Publishers. ISBN: 978-0202309125.
- Argyle, M. (1988). *Bodily Communication*. Routledge. ISBN: 978-0415051149.
- Argyle, M. and Dean, J. (1965). Eye-contact, distance and affiliation. *Sociometry*, pages 289–304.
- Aron, A., Aron, E. N., Tudor, M., and Nelson, G. (1991). Close Relationships as Including Other in the Self. *Journal of Personality and Social Psychology*, 60(2):241–253.
- Axelrod, R. (1986). An Evolutionary Approach to Norms. *American Political Science Review*, 80(4):1095–1111.
- Bailenson, J. N., Blascovich, J., Beall, A. C., and Loomis, J. M. (2001). Equilibrium Theory Revisited: Mutual Gaze and Personal Space in Virtual Environments. *Presence: Teleoperators and Virtual Environments*, 10(6):583–598.
- Bainbridge, W., Hart, J., Kim, E. S., Scassellati, B., et al. (2008). The Effect of Presence on Human-Robot Interaction. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, pages 701–706. IEEE.
- Baldassare, M. and Feller, S. (1975). Cultural Variations in Personal Space: Theory, Methods, and Evidence. *Ethos*, 3(4):481–503.
- Ball, A., Rye, D., Silvera-Tawil, D., and Velonaki, M. (2015). Group Vs. Individual Comfort When a Robot Approaches. In *Proceedings of the 7th International Conference on Social Robotics (ICSR 2015)*, pages 41–50. Springer.
- Bandini, S., Crociani, L., Gorrini, A., and Vizzari, G. (2014). Empirical Investigation on Pedestrian Dynamics in Presence of Groups: a Real World Case Study. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2014)*, pages 1545–1546. International Foundation for Autonomous Agents and Multiagent Systems.
- Bartneck, C. and Forlizzi, J. (2004). A Design-Centred Framework for Social Human-Robot Interaction. In *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2004)*, pages 591–594.
- Bartneck, C., Kulić, D., Croft, E., and Zoghbi, S. (2009). Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. *International Journal of Social Robotics*, 1(1):71–81.

- Bazire, M. and Brézillon, P. (2005). Understanding Context Before Using It. In *Proceedings of the International and Interdisciplinary Conference on Modeling and Using Context (CON-TEXT2005)*, pages 29–40. Springer.
- Belpaeme, T., Baxter, P., De Greeff, J., Kennedy, J., Read, R., Looije, R., Neerincx, M., Baroni, I., and Zelati, M. C. (2013). Child-Robot Interaction: Perspectives and Challenges. In *Proceedings of the 5th International Conference on Social Robotics (ICSR 2015)*, pages 452–459. Springer.
- Beyer, L., Hermans, A., and Leibe, B. (2015). Biternion Nets: Continuous Head Pose Regression from Discrete Training Labels. In *Proceedings of the 37th German Conference on Pattern Recognition (GCPR 2015)*, pages 157–168. Springer.
- Biocca, F. and Harms, C. (2003). Guide to the Networked Minds Social Presence Inventory v. 1.2.
- Bohannon, R. W. (1997). Comfortable and maximum walking speed of adults aged 20-79 years: reference values and determinants. *Age and Ageing*, 26(1):15–19.
- Borg, G. A. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise*, 14(5):377–381.
- Brauer, M. and Chaurand, N. (2010). Descriptive Norms, Prescriptive Norms, and Social Control: An intercultural Comparison of People's Reactions to Uncivil Behaviors. *European Journal of Social Psychology*, 40(3):490–499.
- Breazeal, C. (2003). Emotion and Sociable Humanoid Robots. *International Journal of Human-Computer Studies*, 59(1):119–155.
- Breazeal, C., DePalma, N., Orkin, J., Chernova, S., and Jung, M. (2013). Crowdsourcing Human-Robot Interaction: New Methods and System Evaluation in a Public Environment. *Journal of Human-Robot Interaction*, 2(1):82–111.
- Breazeal, C. and Velasquez, J. (1999). Robot in Society: Friend or Appliance. In *Proceedings of the 1999 Autonomous Agents Workshop on Emotion-Based Agent Architectures*, pages 18–26.
- Brewer, M. B. and Chen, Y.-R. (2007). Where (Who) Are Collectives in Collectivism? Toward Conceptual Clarification of Individualism and Collectivism. *Psychological Review*, 114(1):133–151.
- Bueno, D. R., Viruete, E., and Montano, L. (2011). An autonomous tour guide robot in a next generation smart museum. In *Proceedings of the 5th International Symposium on Ubiquitous Computing and Ambient Intelligence (UCAmI 2011)*. Citeseer.
- Burgard, W., Cremers, A. B., Fox, D., Hähnel, D., Lakemeyer, G., Schulz, D., Steiner, W., and Thrun, S. (1998). The Interactive Museum Tour-Guide Robot. In *Proceedings of the Fifteenth National/Tenth Conference on Artificial Intelligence/Innovative Applications of Artificial Intelligence (AAAI '98/IAAI '98)*, pages 11–18.
- Burgoon, J. K. (1978). A Communication Model of Personal Space Violations: Explication and an Initial Test. *Human Communication Research*, 4(2):129–142.
- Butland, R., Pang, J., Gross, E., Woodcock, A., and Geddes, D. (1982). Two-, six-, and 12-minute walking tests in respiratory disease. *British Medical Journal (Clinical Research Edition)*, 284(6329):1607–1608.
- Butler, J. T. and Agah, A. (2001). Psychological Effects of Behavior Patterns of a Mobile Personal Robot. *Autonomous Robots*, 10(2):185–202.

- Casper, J. and Murphy, R. R. (2003). Human-Robot Interactions during the Robot-Assisted Urban Search and Rescue Response at the World Trade Center. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 33(3):367–385.
- Chidambaram, V., Chiang, Y.-H., and Mutlu, B. (2012). Designing Persuasive Robots: How Robots Might Persuade People Using Vocal and Nonverbal Cues. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, pages 293–300. ACM.
- Cialdini, R. B. and Trost, M. R. (1998). Social Influence: Social Norms, Conformity and Compliance. In Fiske, S. T. and Lindzey, G., editors, *The Handbook of Social Psychology*, chapter 4, pages 151–192. McGraw-Hill. New York, NY, US, 4th edition.
- Ciolek, M. T. (1983). The Proxemics Lexicon: A First Approximation. *Journal of Nonverbal Behavior*, 8(1):55–79.
- Creswell, J. W. and Plano Clark, V. L. (2007). *Designing and Conducting Mixed Methods Research*. Thousand Oaks, Calif. : SAGE Publications. ISBN: 978-1-4129-2792-5.
- Cruz-Maya, A. and Tapus, A. (2016). Influence of User’s Personality on Task Execution When Reminded by a Robot. In *Proceedings of the 8th International Conference on Social Robotics (ICSR 2016)*, pages 829–838. Springer.
- Daamen, W. and Hoogendoorn, S. P. (2006). Free Speed Distributions for Pedestrian Traffic. In *Proceedings of the Transportation Research Board 85th Annual Meeting*.
- Dautenhahn, K. (2007). Methodology & Themes of Human-Robot Interaction: A Growing Research Field. *International Journal of Advanced Robotic Systems*, 4(1):103–108.
- Dautenhahn, K., Walters, M. L., Woods, S., Koay, K. L., Nehaniv, C. L., Sisbot, A., Alami, R., and Siméon, T. (2006). How May I Serve You?: A Robot Companion Approaching a Seated Person in a Helping Context. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI 2006)*, pages 172–179. ACM.
- Dey, A. K. and Abowd, G. D. (2000). Towards a Better Understanding of Context and Context-Awareness. In *Proceedings of the CHI’ 2000 Workshop on the What, Who, Where, When, Why and How of Context-Awareness*.
- Dillman, D. A., Sinclair, M. D., and Clark, J. R. (1993). Effects of questionnaire length, respondent-friendly design, and a difficult question on response rates for occupant-addressed census mail surveys. *Public Opinion Quarterly*, 57(3):289–304.
- DiSalvo, C. F., Gemperle, F., Forlizzi, J., and Kiesler, S. (2002). All Robots Are Not Created Equal: The Design and Perception of. Humanoid Robot Heads. In *Proceedings of the 4th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques (DIS 2002)*, pages 321–326. ACM.
- Donnellan, M. B., Oswald, F. L., Baird, B. M., and Lucas, R. E. (2006). The Mini-IPIP Scales: Tiny-yet-Effective Measures of the Big Five Factors of Personality. *Psychological Assessment*, 18(2):192–203.
- Dosey, M. A. and Meisels, M. (1969). Personal Space and Self-protection. *Journal of Personality and Social Psychology*, 11(2):93–97.
- Duffy, B. R. (2003). Anthropomorphism and the Social Robot. *Robotics and Autonomous Systems*, 42(3):177–190.
- Duke, M. P. and Nowicki, S. (1972). A New Measure and Social-Learning Model for Interpersonal Distance. *Journal of Experimental Research in Personality*, 6(2-3):119–132.

- Eisenhardt, K. M. (1989). Building Theories from Case Study Research. *The Academy of Management Review*, 14(4):532–550.
- Elster, J. (2000). Social Norms and Economic Theory. In Crothers, L. and Lockhart, C., editors, *Culture and Politics*, chapter 20, pages 363–380. Springer.
- Emery, N. J. (2000). The Eyes Have It: The Neuroethology, Function and Evolution of Social Gaze. *Neuroscience & Biobehavioral Reviews*, 24(6):581–604.
- Enright, P. L. and Sherrill, D. L. (1998). Reference Equations for the Six-Minute Walk Distance in Healthy Adults. *American Journal of Respiratory and Critical Care Medicine*, 158(5):1384–1387.
- Epley, N., Waytz, A., and Cacioppo, J. T. (2007). On Seeing Human: A Three-Factor Theory of Anthropomorphism. *Psychological Review*, 114(4):864–886.
- Eresha, G., Haring, M., Endrass, B., Andre, E., and Obaid, M. (2013). Investigating the Influence of Culture on Proxemic Behaviors for Humanoid Robots. In *Proceedings of the 2013 IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2013)*, pages 430–435. IEEE.
- Erickson, F. and Schultz, J. (1997). When is a context? some issues and methods in the analysis of social competence. In Cole, M., Engestrom, Y., and Vasquez, O., editors, *Mind, culture, and activity: Seminal papers from the laboratory of comparative human cognition*, chapter 2, pages 22–31. Cambridge University Press New York, NY.
- Evers, V., Menezes, N., Merino, L., Gavrilu, D., Nabais, F., Pantic, M., Alvito, P., and Karreman, D. (2014). The Development and Real-World Deployment of FROG, the Fun Robotic Outdoor Guide. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2014)*, pages 100–100. ACM.
- Eyssel, F. and Kuchenbrandt, D. (2012). Social Categorization of Social Robots: Anthropomorphism as a Function of Robot Group Membership. *British Journal of Social Psychology*, 51(4):724–731.
- Ezer, N., Fisk, A., and Rogers, W. (2009). Attitudinal and Intentional Acceptance of Domestic Robots by Younger and Older Adults. *Universal Access in Human-Computer Interaction*, pages 39–48.
- Faber, F., Bennewitz, M., Eppner, C., Görög, A., Gonsior, C., Joho, D., Schreiber, M., and Behnke, S. (2009). The Humanoid Museum Tour Guide Robotinho. In *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*, pages 891–896. IEEE.
- Federici, M. L., Gorrini, A., Manenti, L., and Vizzari, G. (2014). An Innovative Scenario for Pedestrian Data Collection: The Observation of an Admission Test at the University of Milano-Bicocca. In *Pedestrian and Evacuation Dynamics 2012*, pages 143–150. Springer.
- Feil-Seifer, D. and Matarić, M. J. (2005). Defining Socially Assistive Robotics. In *Proceedings of the 9th International Conference on Rehabilitation Robotics (ICORR 2005)*, pages 465–468. IEEE.
- Field, A. (2009). *Discovering Statistics Using SPSS*. Sage Publications Ltd, London, UK, 3rd edition. ISBN: 978-1-84787-3.
- Field, A. (2016). *An Adventure in Statistics: The Reality Enigma*. Sage Publications Ltd, London, UK, 1st edition. ISBN: 978-1-4462-1044-4.
- Finnis, K. and Walton, D. (2008). Field observations to determine the influence of population size, location and individual factors on pedestrian walking speeds. *Ergonomics*, 51(6):827–842.

- Fiore, M., Khambhaita, H., Milliez, G., and Alami, R. (2015). An Adaptive and Proactive Human-Aware Robot Guide. In *Proceedings of the 8th International Conference on Social Robotics (ICSR 2016)*, pages 194–203. Springer.
- Fong, T., Nourbakhsh, I., and Dautenhahn, K. (2003). A Survey of Socially Interactive Robots. *Robotics and Autonomous Systems*, 42(3-4):143–166.
- Forston, R. F. (1975). Proxemic Research: A Check on the Validity of its Techniques. In Williams, T. R., editor, *Socialization and Communication in Primary Groups*, chapter 18, pages 339–346. De Gruyter Mouton.
- Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., et al. (2011). Differences Between Tight and Loose Cultures: A 33-Nation Study. *Science*, 332(6033):1100–1104.
- George, D. and Mallery, P. (2016). *IBM SPSS Statistics 23 Step by Step: A Simple Guide and Reference*. Boston : Pearson Education, 14th edition. ISBN: 978-0134320250.
- Gibbs, J. P. (1965). Norms: The Problem of Definition and Classification. *American Journal of Sociology*, pages 586–594.
- Glick, P., DeMorest, J. A., and Hotze, C. A. (1988). Keeping Your Distance: Group Membership, Personal Space, and Requests for Small Favors. *Journal of Applied Social Psychology*, 18(4):315–330.
- Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A. C., et al. (2005). Designing Robots for Long-Term Social Interaction. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005)*, pages 1338–1343. IEEE.
- Gockley, R. and Matarić, M. J. (2006). Encouraging Physical Therapy Compliance With a Hands-off Mobile Robot. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI 2006)*, pages 150–155. ACM.
- Goetz, J., Kiesler, S., and Powers, A. (2003). Matching Robot Appearance and Behavior to Tasks to Improve Human-Robot Cooperation. In *Proceedings of the 12th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2003)*, pages 55–60. IEEE.
- González-Jiménez, J., Galindo, C., and Ruiz-Sarmiento, J. (2012). Technical Improvements of the Giraff Telepresence Robot Based on Users' Evaluation. In *Proceedings of the 21th IEEE International Workshop on Robot and Human Interactive Communication (ROMAN 2012)*, pages 827–832. IEEE.
- Goodrich, M. A. and Schultz, A. C. (2007). Human–Robot Interaction: A Survey. *Foundations and Trends in Human-Computer Interaction*, 1(3):203–275.
- Gosling, S. D., Rentfrow, P. J., and Swann Jr, W. B. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality*, 37(6):504–528.
- Gray, D. E. (2014). *Doing Research in the Real World*. Sage Publications, 3rd edition. ISBN: 9781446260197.
- Greenberg, C. I., Strube, M. J., and Myers, R. A. (1980). A Multitrait-Multimethod Investigation of Interpersonal Distance. *Journal of Nonverbal Behavior*, 5(2):104–114.
- Gupta, V., Hanges, P. J., and Dorfman, P. (2002). Cultural Clusters: Methodology and findings. *Journal of World Business*, 37(1):11–15.
- Haase, R. F. and Markey, M. J. (1973). A methodological note on the study of personal space. *Journal of Consulting and Clinical Psychology*, 40(1):122.

- Hale, J. L. and Burgoon, J. K. (1984). Models of reactions to changes in nonverbal immediacy. *Journal of Nonverbal Behavior*, 8(4):287–314.
- Hall, E. T. (1966). *The Hidden Dimension*. New York, NY: Anchor Books. ISBN: 0-385-08476-5.
- Ham, J., van Esch, M., Limpens, Y., de Pee, J., Cabibihan, J.-J., and Ge, S. S. (2012). The Automaticity of Social Behavior towards Robots: The Influence of Cognitive Load on Interpersonal Distance to Approachable versus Less Approachable Robots. In *Proceedings of the 4th International Conference on Social Robotics (ICSR 2012)*, pages 15–25. Springer.
- Hamamura, T. (2012). Are cultures becoming individualistic? A cross-temporal comparison of individualism–collectivism in the United States and Japan. *Personality and Social Psychology Review*, 16(1):3–24.
- Hart, S. G. and Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, 52:139–183.
- Hayashi, K., Sakamoto, D., Kanda, T., Shiomi, M., Koizumi, S., Ishiguro, H., Ogasawara, T., and Hagita, N. (2007). Humanoid Robots as a Passive-Social Medium - A Field Experiment at a Train Station. In *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI 2007)*, pages 137–144. IEEE.
- Hayduk, L. A. (1978). Personal Space: An Evaluative and Orienting Overview. *Psychological Bulletin*, 85(1):117–134.
- Hayduk, L. A. (1983). Personal Space: Where We Now Stand. *Psychological Bulletin*, 94(2):293–335.
- Hediger, H. (1950). *Wild Animals in Captivity: An Outline of the Biology of Zoological Gardens*. London: Butterworths Scientific Publications. ISBN:.
- Heerink, M., Kröse, B., Wielinga, B., and Evers, V. (2006). Studying the acceptance of a robotic agent by elderly users. *International Journal of Assistive Robotics and Mechatronics*, 7(3):33–43.
- Henrich, J., Heine, S. J., and Norenzayan, A. (2010). The Weirdest People in the World? *Behavioral and Brain Sciences*, 33(2-3):61–83.
- Henry, P. J. (2008). College Sophomores in the Laboratory Redux: Influences of a Narrow Data Base on Social Psychology’s View of the Nature of Prejudice. *Psychological Inquiry*, 19(2):49–71.
- Hicheur, H., Vieilledent, S., and Berthoz, A. (2005). Head motion in humans alternating between straight and curved walking path: Combination of stabilizing and anticipatory orienting mechanisms. *Neuroscience Letters*, 383(1):87–92.
- Hofstede, G. H. and Hofstede, G. (2001). *Culture’s Consequences: Comparing values, Behaviors, Institutions and Organizations across Nations*. Sage. ISBN: 90-254-2681-6.
- Høgh-Olesen, H. (2008). Human Spatial Behaviour: The Spacing of People, Objects and Animals in Six Cross-Cultural Samples. *Journal of Cognition and Culture*, 8(3):245–280.
- Hollands, M. A., Patla, A., and Vickers, J. (2002). “Look where you’re going!”: gaze behaviour associated with maintaining and changing the direction of locomotion. *Experimental Brain Research*, 143(2):221–230.
- Hudlicka, E., Payr, S., Ventura, R., Becker-Asano, C., Fischer, K., Leite, I., and Von, C. (2009). Social interaction with robots and agents: Where do we stand, Where do we go? In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII 2009)*, pages 1–6. IEEE.

- Hüttenrauch, H., Eklundh, K. S., Green, A., Topp, E., et al. (2006). Investigating spatial relationships in human-robot interaction. In *Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006)*, pages 5052–5059. IEEE.
- Jones, S. E. and Aiello, J. R. (1979). A Test of the Validity of Projective and Quasi-Projective Measures of Interpersonal Distance. *Western Journal of Communication*, 43(2):143–152.
- Joosse, M. P. and Evers, V. (2017). A Guide Robot at the Airport: First Impressions. In *Proceedings of the 11th ACM International Conference on Human-Robot Interaction (HRI 2017)*, pages 149–150. IEEE.
- Joosse, M. P., Knuppe, R. A., Pinggen, G. L. J., Varkevisser, R. A., Vukoja, J., Lohse, M., and Evers, V. (2015a). Robots Guiding Small Groups: The Effect of Appearance Change on the User Experience. In *Proceedings of the Fourth International Symposium on New Frontiers in Human-Robot Interaction*. University of Kent.
- Joosse, M. P., Lohse, M., and Evers, V. (2015b). Crowdsourcing Culture in HRI: Lessons Learned From Quantitative and Qualitative Data Collections. Paper presented at the 3rd International Workshop on Culture Aware Robotics at ICSR'15, Paris, 30 October 2015.
- Joosse, M. P., Lohse, M., and Evers, V. (2015c). How a Guide Robot Should Behave at an Airport - Insights Based on Observing Passengers. Technical report.
- Joosse, M. P., Poppe, R. W., Lohse, M., and Evers, V. (2014). Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People. In *Proceedings of the 5th ACM International Conference on Collaboration Across Boundaries: Culture, Distance & Technology (CABS 2014)*, pages 121–130. ACM.
- Joosse, M. P., Sardar, A. H., and Evers, V. (2011). Behave: a Set of Measures to Assess Users' Attitudinal and Non-Verbal Behavioral Responses to a Robot's Social Behaviors. In *Proceedings of the 3rd International Conference on Social Robotics (ICSR 2011)*, pages 84–94. Springer.
- Joosse, M. P., Van Waveren, S., Zaga, C., and Evers, V. (2017). Groups in Conflict at the Airport: How Everyday People Think a Robot Should Act. Paper presented at the CSCW'17 Workshop on Robots in Groups and Teams.
- Kamide, H., Mae, Y., Takubo, T., Ohara, K., and Arai, T. (2014). Direct Comparison of Psychological Evaluation Between vRtual and Real Humanoids: Personal Space and Subjective Impressions. *International Journal of Human-Computer Studies*, 72(5):451–459.
- Kanda, T., Hirano, T., Eaton, D., and Ishiguro, H. (2004). Interactive Robots as Social Partners and Peer Tutors for Children: A Field Trial. *Journal of Human-Computer Interaction*, 19(1):61–84.
- Kanda, T., Sato, R., Saiwaki, N., and Ishiguro, H. (2007). A Two-Month Field Trial in an Elementary School for Long-Term Human-Robot Interaction. *IEEE Transactions on Robotics*, 23(5):962–971.
- Kanda, T., Shiomi, M., Miyashita, Z., Ishiguro, H., and Hagita, N. (2009). An Affective Guide Robot in a Shopping Mall. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2009)*, pages 173–180. ACM.
- Karahanna, E., Evaristo, J. R., and Srite, M. (2006). Levels of Culture and Individual Behavior: An Integrative Perspective. *Advanced Topics in Global Information Management*, 5(1):30–50.
- Karremen, D., Ludden, G., and Evers, V. (2015). Visiting Cultural Heritage with a Tour Guide Robot: A User Evaluation Study in-the-Wild. In *Proceedings of the 7th International Conference on Social Robotics (ICSR 2015)*, pages 317–326. Springer.

- Karreman, D. E., Bradford, G. U. S., van Dijk, E. M. A. G., Lohse, M., Evers, V., et al. (2013). Picking Favorites: The Influence of Robot Eye-Gaze on Interactions with Multiple Users. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013)*, pages 123–128. IEEE.
- Karreman, D. E., Utama, L., Joosse, M. P., Lohse, M., van Dijk, E. M. A. G., and Evers, V. (2014). Robot Etiquette: How to Approach a Pair of People? In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2014)*, pages 196–197. ACM.
- Kelly, J. R., McCarty, M. K., and Iannone, N. E. (2013). Interaction in Small Groups. In DeLamater, J. and Ward, A., editors, *Handbook of Social Psychology*, chapter 14, pages 413–438. Springer Netherlands, Dordrecht, the Netherlands.
- Kendon, A. (1990). *Conducting interaction: Patterns of behavior in focused encounters*, volume 7. Cambridge University Press, Cambridge, UK. ISBN: 978-0-521-38938-9.
- Kendon, A. (2010). Spacing and Orientation in Co-present Interaction. In *Development of Multimodal Interfaces: Active Listening and Synchrony*, pages 1–15. Springer.
- Kidd, C. D. and Breazeal, C. (2008). Robots at Home: Understanding Long-Term Human-Robot Interaction. In *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2008)*, pages 3230–3235. IEEE.
- Kim, Y., Kwak, S. S., and Kim, M.-s. (2013). Am I acceptable to you? Effect of a robot's verbal language forms on people's social distance from robots. *Computers in Human Behavior*, 29(3):1091–1101.
- Kim, Y. and Mutlu, B. (2014). How social distance shapes human–robot interaction. *International Journal of Human-Computer Studies*, 72(12):783–795.
- Klamer, T., Allouch, S. B., and Heylen, D. K. J. (2010). “Adventures of Harvey” - Use, Acceptance of and Relationship Building with a Social Robot in a Domestic Environment. In *Proceedings of the International Conference on Human-Robot Personal Relationship*, pages 74–82. Springer.
- Knapp, M. L., Hall, J. A., and Horgan, T. G. (2013). *Nonverbal Communication in Human Interaction*. Boston, MA : Wadsworth Cengage Learning. ISBN: 9781285083513.
- Knight, H. and Simmons, R. (2013). Estimating Human Interest and Attention Via Gaze Analysis. In *Proceedings of the 2013 IEEE International Conference on Robotics and Automation (ICRA 2013)*, pages 4350–4355. IEEE.
- Knowles, E. S. (1980). Convergent Validity of Personal Space Measures: Consistent Results with Low Intercorrelations . *Journal of Nonverbal Behavior*, 4(4):240–248.
- Koay, K. L., Syrdal, D. S., Ashgari-Oskoei, M., Walters, M. L., and Dautenhahn, K. (2014). Social roles and baseline proxemic preferences for a domestic service robot. *International Journal of Social Robotics*, 6(4):469–488.
- Koay, K. L., Syrdal, D. S., Walters, M. L., and Dautenhahn, K. (2007). Living with Robots: Investigating the Habituation Effect in Participants' Preferences During a Longitudinal Human-Robot Interaction Study. In *Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2007)*, pages 564–569. IEEE.
- Koide, Y., Kanda, T., Sumi, Y., Kogure, K., and Ishiguro, H. (2004). An Approach to Integrating an Interactive Guide Robot with Ubiquitous Sensors. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2004)*, volume 3, pages 2500–2505. IEEE.

- Kosiński, T., Obaid, M., Woźniak, P. W., Fjeld, M., and Kucharski, J. (2016). A fuzzy data-based model for Human-Robot Proxemics. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2016)*, pages 335–340. IEEE.
- Kristoffersson, A., Coradeschi, S., Eklundh, K. S., and Loutfi, A. (2013a). Towards Measuring Quality of Interaction in Mobile Robotic Telepresence using Sociometric Badges. *Paladyn, Journal of Behavioral Robotics*, 4(1):34–48.
- Kristoffersson, A., Eklundh, K. S., and Loutfi, A. (2013b). Measuring the Quality of Interaction in Mobile Robotic Telepresence: A Pilot's Perspective. *International Journal of Social Robotics*, 5(1):89–101.
- Kruse, T., Kirsch, A., Khambhaita, H., and Alami, R. (2014). Evaluating Directional Cost Models in Navigation. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2014)*, pages 350–357. ACM.
- Kucner, T., Saarinen, J., Magnusson, M., and Lilienthal, A. J. (2013). Conditional Transition Maps: Learning Motion Patterns in Dynamic Environments. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013)*, pages 1196–1201. IEEE.
- Kuno, Y., Sadazuka, K., Kawashima, M., Tsuruta, S., Yamazaki, K., and Yamazaki, A. (2007a). Effective Head Gestures for Museum Guide Robots in Interaction with Humans. In *Proceedings of the 16th IEEE International Symposium on Robot and Human interactive Communication (RO-MAN 2007)*, pages 151–156. IEEE.
- Kuno, Y., Sadazuka, K., Kawashima, M., Yamazaki, K., Yamazaki, A., and Kuzuoka, H. (2007b). Museum Guide Robot Based on Sociological Interaction Analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2007)*, pages 1191–1194. ACM.
- Kuno, Y., Sekiguchi, H., Tsubota, T., Moriyama, S., Yamazaki, K., and Yamazaki, A. (2006). Museum Guide Robot with Communicative Head Motion. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2006)*, pages 33–38. IEEE.
- Lauckner, M., Kobiela, F., and Manzey, D. (2014). “Hey robot, please step back!”-exploration of a spatial threshold of comfort for human-mechanoid spatial interaction in a hallway scenario. In *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2014)*, pages 780–787. IEEE.
- Laxman, K. K., Rastogi, R., and Chandra, S. (2010). Pedestrian Flow Characteristics in Mixed Traffic Conditions. *Journal of Urban Planning and Development*, 136(1):23–33.
- Lazar, J., Feng, J. H., and Hochheiser, H. (2010). *Research methods in human-computer interaction*. Chichester, UK: John Wiley & Sons. ISBN: 978-0-470-72337-1.
- Lee, K. M., Jung, Y., Kim, J., and Kim, S. R. (2006a). Are Physically Embodied Social Agents Better Than Disembodied Social Agents?: The Effects of Physical Embodiment, Tactile Interaction, and People's Loneliness in Human-Robot Interaction. *International Journal of Human-Computer Studies*, 64(10):962–973.
- Lee, K. M., Peng, W., Jin, S.-A., and Yan, C. (2006b). Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human-Robot Interaction. *Journal of Communication*, 56(4):754–772.

- Lee, M. K. and Takayama, L. (2011). "now, i have a body": Uses and social norms for mobile remote presence in the workplace. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2011)*, pages 33–42. ACM.
- Lee, S.-I., Lau, I. Y.-m., Kiesler, S., and Chiu, C.-Y. (2005). Human Mental Models of Humanoid Robots. In *Proceedings of the 2005 IEEE International Conference on Robotics and Automation (ICRA 2005)*, pages 2767–2772. IEEE.
- Levine, R. V. and Norenzayan, A. (1999). The Pace of Life in 31 Countries. *Journal of Cross-Cultural Psychology*, 30(2):178–205.
- Levinson, D. J. et al. (1986). A Conception of Adult Development. *American Psychologist*, 41(1):3–13.
- Li, D., Rau, P. L. P., and Li, Y. (2010). A Cross-cultural Study: Effect of Robot Appearance and Task. *International Journal of Social Robotics*, 2(2):175–186.
- Lichtenthäler, C. (2014). *Legibility of Robot Behavior: Investigating Legibility of Robot Navigation in Human-Robot Path Crossing Scenarios*. PhD thesis, München, Technische Universität München, Dissertation.
- Lichtenthäler, C., Peters, A., Griffiths, S., and Kirsch, A. (2013). Social navigation - identifying robot navigation patterns in a path crossing scenario. In *Proceedings of the 5th International Conference on Social Robotics (ICSR 2013)*, pages 84–93. Springer.
- Linder, T. and Arras, K. O. (2015). Real-Time Full-Body Human Attribute Classification in RGB-D Using a Tessellation Boosting Approach. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2015)*, pages 1335–1341. IEEE.
- Linder, T., Breuers, S., Leibe, B., and Arras, K. O. (2016). On Multi-Modal People Tracking from Mobile Platforms in Very Crowded and Dynamic Environments. In *Proceedings of the 2016 IEEE International Conference on Robotics and Automation (ICRA 2016)*, pages 5512–5519.
- Little, K. B. (1965). Personal Space. *Journal of Experimental Social Psychology*, 1(3):237–247.
- Little, K. B. (1968). Cultural Variations in Social Schemata. *Journal of Personality and Social Psychology*, 10(1):1.
- Lofland, J., Snow, D., Anderson, L., and Lofland, L. H. (2006). *Analyzing social settings: a guide to qualitative observation and analysis*.
- Lohse, M. (2011). The Role of Expectations and Situations in Human-Robot Interaction. In Dautenhahn, K. and Sanders, J., editors, *New Frontiers in Human-Robot Interaction*, pages 35–56.
- Lombard, M., Ditton, T. B., and Weinstein, L. (2009). Measuring Presence: The Temple Presence Inventory. In *Proceedings of the 12th Annual International Workshop on Presence*, pages 1–15.
- Lomranz, J. (1976). Cultural Variations in Personal Space. *The Journal of Social Psychology*, 99(1):21–27.
- Love, K. D. and Aiello, J. R. (1980). Using Projective Techniques to Measure Interaction Distance: A Methodological Note. *Personality and Social Psychology Bulletin*, 6(1):102–104.

- Ludewig, Y., Döring, N., and Exner, N. (2012). Design and Evaluation of the Personality Trait Extraversion of a Shopping Robot. In *Proceedings of the 21st IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2012)*, pages 372–379. IEEE.
- Maguire, M. (2001). Context of Use within usability activities. *International Journal of Human-Computer Studies*, 55(4):453–483.
- Mahoney, E. R. (1974). Compensatory Reactions to Spatial Immediacy. *Sociometry*, 37(3):423–431.
- Marshall, P., Rogers, Y., and Pantidi, N. (2011). Using F-formations to analyse spatial patterns of interaction in physical environments. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW 2011)*, pages 445–454. ACM.
- McCroskey, J. C. and Teven, J. J. (1999). Goodwill: A Reexamination of the Construct and Its Measurement. *Communications Monographs*, 66(1):90–103.
- McGavin, C., Gupta, S., and McHardy, G. (1976). Twelve-Minute Walking Test for Assessing Disability in Chronic Bronchitis. *British Medical Journal*, 1(6013):822–823.
- Mead, R. and Mataric, M. (2016). Robots Have Needs Too: How and Why People Adapt Their Proxemic Behavior to Improve Robot Social Signal Understanding. *Journal of Human-Robot Interaction*, 5(2):48–68.
- Merriam-Webster (2015a). Norm. <http://www.merriam-webster.com/dictionary/norm>. Accessed: 2015-06-07.
- Merriam-Webster (2015b). Robot. <http://www.merriam-webster.com/dictionary/robot>. Accessed: 2015-05-27.
- Merriam-Webster (2017a). Context. <https://www.merriam-webster.com/dictionary/context>. Accessed: 2017-03-21.
- Merriam-Webster (2017b). Group. <https://www.merriam-webster.com/dictionary/group>. Accessed: 2017-04-13.
- Michalowski, M. P., Šabanović, S., and Simmons, R. (2006). A Spatial Model of Engagement for a Social Robot. In *Proceedings of the 9th IEEE International Workshop on Advanced Motion Control*, pages 762–767. IEEE.
- Moreland, R. L. (2010). Are Dyads Really Groups? *Small Group Research*, 41(2):251–267.
- Moussaïd, M., Perozo, N., Garnier, S., Helbing, D., and Theraulaz, G. (2010). The Walking Behaviour of Pedestrian Social Groups and Its Impact on Crowd Dynamics. *PLoS one*, 5(4):e10047.
- Mumm, J. and Mutlu, B. (2011). Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI 2011)*, pages 331–338. ACM.
- Nass, C., Fogg, B., and Moon, Y. (1996). Can Computers be Teammates? *International Journal of Human-Computer Studies*, 45(6):669–678.
- Nass, C. and Moon, Y. (2000). Machines and Mindlessness: Social Responses to Computers. *Journal of Social Issues*, 56(1):81–103.
- Nass, C., Steuer, J., and Tauber, E. R. (1994). Computers are Social Actors. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI 1994)*, pages 72–78. ACM.

- Nass, C. I. and Yen, C. (2010). *The Man who Lied to His Laptop: What Machines Teach Us about Human Relationships*. New York, NY: Penguin Group. ISBN: 978-1-61723-004-2.
- Nomura, T., Kanda, T., Suzuki, T., and Kato, K. (2008). Prediction of Human Behavior in Human–Robot Interaction Using Psychological Scales for Anxiety and Negative Attitudes Toward Robots. *IEEE Transactions on Robotics*, 24(2):442–451.
- Norris, S. (2004). *Analyzing Multimodal Interaction: A Methodological Framework*. New York, NY : Routledge, 2004. ISBN: 978-0415328555.
- Novelli, D., Drury, J., and Reicher, S. (2010). Come together: Two studies concerning the impact of group relations on personal space. *British Journal of Social Psychology*, 49(2):223–236.
- Nummenmaa, L., Hyönä, J., and Hietanen, J. K. (2009). I'll walk this way: Eyes reveal the direction of locomotion and make passersby look and go the other way. *Psychological Science*, 20(12):1454–1458.
- Obaïd, M., Sandoval, E. B., Złotowski, J., Moltchanova, E., Basedow, C. A., and Bartneck, C. (2016). Stop! That is Close Enough. How Body Postures Influence Human-Robot Proximity. In *Proceedings of the 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2016)*, pages 354–361. IEEE.
- Ohara, K., Negi, S., Takubo, T., Mae, Y., and Arai, T. (2009). Evaluation of virtual and real robot based on human impression. In *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*, pages 873–878. IEEE.
- Okal, B. and Arras, K. O. (2016). Learning Socially Normative Robot Navigation Behaviors with Bayesian Inverse Reinforcement Learning.
- Okuno, Y., Kanda, T., Imai, M., Ishiguro, H., and Hagita, N. (2009). Providing Route Directions: Design of Robot's Utterance, Gesture, and Timing. In *Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2009)*, pages 53–60. IEEE.
- Oulasvirta, A. (2009). Field Experiments in HCI: Promises and Challenges. In Isomäki, H. and Saariluoma, P., editors, *Future Interaction Design II*, pages 87–116. Springer.
- Pacchierotti, E., Christensen, H., Jensfelt, P., et al. (2006). Evaluation of Passing Distance for Social Robots. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2006)*, pages 315–320. IEEE.
- Palmieri, L., Koenig, S., and Arras, K. O. (2016). RRT-Based Nonholonomic Motion Planning Using Any-Angle Path Biasing.
- Pandey, A. K. and Alami, R. (2009). A Framework for Adapting Social Conventions in a Mobile Robot Motion in Human-Centered Environment. In *Proceedings of the International Conference on Advanced Robotics (ICAR 2009)*, pages 1–8. IEEE.
- Patterson, M. L. (2014). Reflections on Historical Trends and Prospects in Contemporary Nonverbal Research. *Journal of Nonverbal Behavior*, 38(2):171–180.
- Pedersen, D. M. (1973). Development of a Personal Space Measure. *Psychological Reports*, 32(2):527–535.
- Pelto, P. J. (1968). The differences between “tight” and “loose” societies. *Society*, 5(5):37–40.

- Pitsch, K., Kuzuoka, H., Suzuki, Y., Süssenbach, L., Luff, P., and Heath, C. (2009). "The First Five Seconds": Contingent Stepwise Entry into an Interaction as a Means to Secure Sustained Engagement in HRI. In *Proceedings of the 18th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2009)*, pages 985–991. IEEE.
- Pitsch, K. and Wrede, S. (2014). When a robot orients visitors to an exhibit. Referential practices and interactional dynamics in real world HRI. In *Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2014)*, pages 36–42. IEEE.
- Pitsch, K., Wrede, S., Seele, J.-C., and Süssenbach, L. (2011). Attitude of German Museum Visitors towards an Interactive Art Guide Robot. In *Proceedings of the 6th International Conference on Human-Robot Interaction (HRI 2011)*, pages 227–228. ACM.
- Posner, R. A. (1997). Social Norms and the Law: An Economic Approach. *The American Economic Review*, 87(2):365–369.
- Prentice, D. A. (2012). The Psychology of Social Norms and the Promotion of Human Rights. In Goodman, R., Jinks, D., and Woods, A. K., editors, *Understanding Social Action, Promoting Human Rights*, pages 23–46. Oxford University Press, USA. ISBN: 9780195371895.
- Qian, K., Ma, X., Dai, X., and Fang, F. (2010). Socially Acceptable Pre-collision Safety Strategies for Human-Compliant Navigation of Service Robots. *Advanced Robotics*, 24(13):1813–1840.
- Rammstedt, B. and John, O. P. (2007). Measuring Personality in One Minute or Less: A 10-Item Short Version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1):203–212.
- Read, R. and Belpaeme, T. (2014). Situational Context Directs How People Affectively Interpret Robotic Non-linguistic Utterances. In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2014)*, pages 41–48. ACM.
- Reeves, B. and Nass, C. (1996). *The Media Equation: How People Treat Computers, Television, and New Media Like Real People and Places*. CSLI Publications, Stanford, CA. ISBN: 1-57586-053-8.
- Rehm, M., André, E., and Nischt, M. (2005). Let's Come Together - Social Navigation Behaviors of Virtual and Real Humans. In *International Conference on Intelligent Technologies for Interactive Entertainment (INTETAIN 2005)*, pages 124–133. Springer.
- Reja, U., Manfreda, K. L., Hlebec, V., and Vehovar, V. (2003). Open-ended vs. Close-ended Questions in Web Questionnaires. In Ferligo, A. and Mrvar, A., editors, *Developments in Applied Statistics*, pages 160–117.
- Remland, M. S., Jones, T. S., and Brinkman, H. (1995). Interpersonal Distance, Body Orientation, and Touch: Effects of Culture, Gender, and Age. *The Journal of Social Psychology*, 135(3):281–297.
- Riccio, F., Vanzo, A., Mirabella, V., Catarci, T., and Nardi, D. (2016). Enabling Symbiotic Autonomy in Short-Term Interactions: A User Study. In *Proceedings of the 8th International Conference on Social Robotics (ICSR 2016)*, pages 796–807. Springer.
- Rios-Martinez, J., Spalanzani, A., and Laugier, C. (2014). From Proxemics Theory to Socially-Aware Navigation: A Survey. *International Journal of Social Robotics*, 7(2):137–153.
- Rodriguez-Lizundia, E., Marcos, S., Zalama, E., Gómez-García-Bermejo, J., and Gordaliza, A. (2015). A Bellboy Robot: Study of the Effects of Robot Behaviour on User Engagement and Comfort. *International Journal of Human-Computer Studies*, 82:83–95.

- Rothenbücher, D., Li, J., Sirkin, D., Mok, B., and Ju, W. (2015). Ghost Driver: A Platform for Investigating Interactions Between Pedestrians and Driverless Vehicles. In *Adjunct Proceedings of the 7th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pages 44–49. ACM.
- Šabanović, S., Bennett, C. C., Chang, W.-L., and Huber, L. (2013). PARO robot affects diverse interaction modalities in group sensory therapy for older adults with dementia. In *Proceedings of the IEEE International Conference on Rehabilitation Robotics (ICORR 2013)*, pages 1–6. IEEE.
- Šabanović, S., Reeder, S., and Kechavarzi, B. (2014). Designing Robots in the Wild: In situ Prototype Evaluation for a Break Management Robot. *Journal of Human-Robot Interaction*, 3(1):70–88.
- Sakamoto, D., Kanda, T., Ono, T., Kamashima, M., Imai, M., and Ishiguro, H. (2004). Cooperative embodied communication emerged by interactive humanoid robots. In *Proceedings of the 13th IEEE International Workshop on Robot and Human Interactive Communication (RO-MAN 2004)*, pages 443–448. IEEE.
- Sardar, A. H., Joosse, M. P., Weiss, A., and Evers, V. (2012). Don't stand so Close to Me: Users' Attitudinal and Behavioral Responses to Personal Space Invasion by Robots. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI 2012)*, pages 229–230. ACM.
- Saunders, M., Lewis, P., and Thornhill, A. (2007). *Research Methods for Business Students*. Essex, U.K.: Pearson Education Ltd.
- Shen, J. and Hu, H. (2006). Visual Navigation of a Museum Guide Robot. In *Proceedings of the Sixth World Congress on Intelligent Control and Automation (WCICA 2006)*, volume 2, pages 9169–9173. IEEE.
- Shiomi, M., Kanda, T., Glas, D. F., Satake, S., Ishiguro, H., and Hagita, N. (2009). Field Trial of Networked Social Robots in a Shopping Mall. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, pages 2846–2853. IEEE.
- Shiomi, M., Kanda, T., Ishiguro, H., and Hagita, N. (2006). Interactive Humanoid Robots for a Science Museum. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction (HRI 2006)*, pages 305–312. ACM.
- Shiomi, M., Kanda, T., Ishiguro, H., and Hagita, N. (2010). A larger audience, please!: Encouraging people to listen to a guide robot. In *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010)*, pages 31–38. IEEE Press.
- Shiomi, M., Sakamoto, D., Kanda, T., Ishi, C. T., Ishiguro, H., and Hagita, N. (2008). A Semi-autonomous Communication Robot - A Field Trial at a Train Station. In *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI 2008)*, pages 303–310. ACM.
- Shiomi, M., Sakamoto, D., Kanda, T., Ishi, C. T., Ishiguro, H., and Hagita, N. (2011). Field Trial of a Networked Robot at a Train Station. *International Journal of Social Robotics*, 3(1):27–40.
- Siek, K. A., Hayes, G. R., Newman, M. W., and Tang, J. C. (2014). Field Deployments: Knowing from Using in Context. In Olson, J. S. and Kellogg, W. A., editors, *Ways of Knowing in HCI*, pages 119–142. Springer.
- Silvera-Tawil, D., Velonaki, M., and Rye, D. (2015). Human-Robot Interaction with Humanoid Diamandini Using an Open Experimentation Method. In *Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2015)*, pages 425–430. IEEE.

- Sommer, R. (1959). Studies in Personal Space. *Sociometry*, 22(3):247–260.
- Sommer, R. (1969). *Personal Space. The Behavioral Basis of Design*. Englewood Cliffs, NJ: Prentice-Hall.
- Sommer, R. (2002). From Personal Space to Cyberspace. *Handbook of Environmental Psychology*, 2.
- Sorokowska, A., Sorokowski, P., Hilpert, P., Cantarero, K., Frackowiak, T., Ahmadi, K., Alghraibeh, A. M., Aryeetey, R., Bertoni, A., Bettache, K., et al. (2017). Preferred Interpersonal Distances: A Global Comparison. *Journal of Cross-Cultural Psychology*, pages 577–592.
- Srinivasan, V. and Takayama, L. (2016). Help Me Please: Robot Politeness Strategies for Soliciting Help From People. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems (CHI 2016)*, pages 4945–4955. ACM.
- Steffen, T. M., Hacker, T. A., and Mollinger, L. (2002). Age- and Gender-Related Test Performance in Community-Dwelling Elderly People: Six-Minute Walk Test, Berg Balance Scale, Timed Up & Go Test, and Gait Speeds. *Physical Therapy*, 82(2):128–137.
- Strauss, A. L. (1987). *Qualitative analysis for social scientists*. Cambridge, UK: Cambridge University Press. ISBN: 978-0521338066.
- Sundstrom, E. and Altman, I. (1976). Interpersonal relationships and personal space: Research review and theoretical model. *Human Ecology*, 4(1):47–67.
- Sussman, N. M. and Rosenfeld, H. M. (1982). Influence of Culture, Language, and Sex on Conversational Distance. *Journal of Personality and Social Psychology*, 42(1):66.
- Syrdal, D. S., Dautenhahn, K., Koay, K. L., Walters, M. L., and Ho, W. C. (2013). Sharing Spaces, Sharing Lives - The Impact of Robot Mobility on User Perception of a Home Companion Robot. In *Proceedings of the 5th International Conference on Social Robotics (ICSR 2013)*, pages 321–330. Springer.
- Syrdal, D. S., Dautenhahn, K., Woods, S., Walters, M. L., and Koay, K. L. (2006). 'doing the right thing wrong'-personality and tolerance to uncomfortable robot approaches. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (ROMAN 2006)*, pages 183–188. IEEE.
- Tajfel, H. (1970). Experiments in Intergroup Discrimination. *Scientific American*, 223(5):96–102.
- Takayama, L., Ju, W., and Nass, C. (2008). Beyond Dirty, Dangerous and Dull: What Everyday People Think Robots Should Do. In *Proceedings of the 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI 2008)*, pages 25–32. ACM.
- Takayama, L. and Pantofaru, C. (2009). Influences on Proxemic Behaviors in Human-Robot Interaction. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2009)*, pages 5495–5502. IEEE.
- Tanaka, F. and Matsuzoe, S. (2012). Children Teach a Care-Receiving Robot to Promote Their Learning: Field Experiments in a Classroom for Vocabulary Learning. *Journal of Human-Robot Interaction*, 1(1).
- Tedesco, J. F. and Fromme, D. K. (1974). Cooperation, Competition and Personal Space. *Sociometry*, 37(1):116–121.
- Teng, C.-I., Tseng, F.-C., Chen, Y.-S., and Wu, S. (2012). Online gaming misbehaviours and their adverse impact on other gamers. *Online Information Review*, 36(3):342–358.

- Thrun, S., Bennewitz, M., Burgard, W., Cremers, A. B., Dellaert, F., Fox, D., Hähnel, D., Rosenberg, C., Roy, N., Schulte, J., et al. (1999). MINERVA: A Second-Generation Museum Tour-Guide Robot. In *Proceedings of the International Conference on Robotics and Automation (ICRA 1999)*, volume 3. IEEE.
- Tolor, A. and Salafia, W. R. (1971). The Social Schemata Technique as a Projective Device. *Psychological Reports*, 28(2):423–429.
- Tomatis, N., Philippsen, R., Jensen, B., Arras, K. O., Terrien, G., Piguët, R., and Siegwart, R. (2002). Building a Fully Autonomous Tour Guide Robot: Where Academic Research Meets Industry. In *Proceedings of the 33rd International Symposium on Robotics (ISR)*, number LSA-CONF-2002-023.
- Torta, E., Cuijpers, R. H., and Juola, J. F. (2013). Design of a Parametric Model of Personal Space for Robotic Social Navigation. *International Journal of Social Robotics*, 5(3):357–365.
- Torta, E., Cuijpers, R. H., Juola, J. F., and van der Pol, D. (2011). Design of Robust Robotic Proxemic Behaviour. In *Proceedings of the 3rd International Conference on Social Robotics (ICSR 2011)*, pages 21–30. Springer.
- Triandis, H. C. (1972). *The analysis of subjective culture*. New York, NY: Wiley-Interscience. ISBN: 9780471889052.
- Triandis, H. C. and Gelfand, M. J. (1998). Converging Measurement of Horizontal and Vertical Individualism and Collectivism. *Journal of Personality and Social Psychology*, 74(1):118–128.
- Triebel, R., Arras, K., Alami, R., Beyer, L., Breuers, S., Chatila, R., Chetouani, M., Cremers, D., Evers, V., Fiore, M., et al. (2016). Spencer: A socially aware service robot for passenger guidance and help in busy airports. In *Field and Service Robotics: Results of the 10th International Conference*, pages 607–622. Springer.
- Trochim, W. M. The Research Methods Knowledge Base, 2nd Edition. <http://www.socialresearchmethods.net/kb/convdisc.php>. Version current as of October 20, 2006. Lat accessed: 2017-06-24.
- Tsang, R. C. (2005). Reference Values for 6-Minute Walk Test and Hand-Grip Strength in Healthy Hong Kong Chinese Adults. *Hong Kong Physiotherapy Journal*, 23(1):6–12.
- von Grünau, M. and Anston, C. (1995). The Detection of Gaze Direction: A Stare-In-The-Crowd Effect. *Perception*, 24(11):1297–1313.
- Vroon, J. H., Joosse, M. P., Lohse, M., Kolkmeier, J., Kim, J., Truong, K., Englebienne, G., Heylen, D. K. J., and Evers, V. (2015). Dynamics of Social Positioning Patterns in Group-Robot Interactions. In *Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2015)*, pages 394–399. IEEE.
- Walkey, F. H. and Gilmour, D. R. (1979). Comparative evaluation of a videotaped measure of interpersonal distance. *Journal of Consulting and Clinical Psychology*, 47(3):575.
- Walters, M. L., Dautenhahn, K., Te Boekhorst, R., Koay, K. L., Kaouri, C., Woods, S., Nehaniv, C., Lee, D., and Werry, I. (2005a). The influence of subjects' personality traits on personal spatial zones in a human-robot interaction experiment. In *Proceedings of the 14th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2005)*, pages 347–352. IEEE.

- Walters, M. L., Dautenhahn, K., Woods, S. N., and Koay, K. L. (2007). Robotic Etiquette: Results from User Studies Involving a Fetch and Carry Task. In *Proceedings of the 2nd ACM/IEEE International Conference on Human-Robot Interaction (HRI 2007)*, pages 317–324. IEEE.
- Walters, M. L., Lohse, M., Hanheide, M., Wrede, B., Syrdal, D. S., Koay, K. L., Green, A., Hüttenrauch, H., Dautenhahn, K., Sagerer, G., et al. (2011a). Evaluating the Robot Personality and Verbal Behavior of Domestic Robots Using Video-Based Studies. *Advanced Robotics*, 25(18):2233–2254.
- Walters, M. L., Oskoei, M., Syrdal, D. S., Dautenhahn, K., et al. (2011b). A long-term Human-Robot Proxemic study. In *Proceedings of the 20th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2011)*, pages 137–142. IEEE.
- Walters, M. L., Syrdal, D. S., Dautenhahn, K., Te Boekhorst, R., and Koay, K. L. (2008a). Avoiding the uncanny valley: robot appearance, personality and consistency of behavior in an attention-seeking home scenario for a robot companion. *Autonomous Robots*, 24(2):159–178.
- Walters, M. L., Syrdal, D. S., Koay, K. L., Dautenhahn, K., and Boekhorst, R. T. (2008b). Human approach distances to a mechanical-looking robot with different robot voice styles. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2008)*, pages 707–712. IEEE.
- Walters, M. L., Woods, S., Koay, K. L., and Dautenhahn, K. (2005b). Practical and Methodological Challenges in Designing and Conducting Human-Robot Interaction Studies. In *Proceedings of the AISB 2005 Symposium on Robot Companions : Hard Problems and Open Challenges in Robot-Human Interaction*, volume 5, pages 110–120.
- Wang, L., Rau, P.-L. P., Evers, V., Robinson, B. K., and Hinds, P. (2010). When in Rome: the role of culture & context in adherence to robot recommendations. In *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010)*, pages 359–366. IEEE Press.
- Weiss, A., Igelsböck, J., Tscheligi, M., Bauer, A., Kühnlenz, K., Wollherr, D., and Buss, M. (2010). Robots Asking for Directions - The Willingness of Passers-by to Support Robots. In *Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI 2010)*, pages 23–30. IEEE Press.
- Weiss, A., Mirnig, N., Bruckenberg, U., Strasser, E., Tscheligi, M., Wollherr, D., Stanczyk, B., et al. (2015). The Interactive Urban Robot: User-centered development and final field trial of a direction requesting robot. *Paladyn, Journal of Behavioral Robotics*, 6(1):42–56.
- Weiss, A., Mirnig, N., Buchner, R., Förster, F., and Tscheligi, M. (2011). Transferring Human-Human Interaction Studies to HRI Scenarios in Public Space. In *Proceedings of the IFIP Conference on Human-Computer Interaction (INTERACT 2011)*, pages 230–247. Springer.
- Willemse, C. J., Huisman, G., Jung, M. M., van Erp, J. B., and Heylen, D. K. (2016). Observing Touch from Video: The Influence of Social Cues on Pleasantness Perceptions. In *Proceedings of the International Conference on Human Haptic Sensing and Touch Enabled Computer Applications (EuroHaptics 2016)*, pages 196–205. Springer.
- Williams, K. D. (2010). Dyads can be groups (and often are). *Small Group Research*, 41(2):268–274.
- Willis, A., Gjersøe, N., Havard, C., Kerridge, J., and Kukla, R. (2004). Human Movement Behaviour in Urban Spaces: Implications for the Design and Modelling of Effective Pedestrian Environments. *Environment and Planning B Planning and Design*, 31(6):805–828.

- Woods, S., Walters, M., Koay, K. L., and Dautenhahn, K. (2006a). Comparing human robot interaction scenarios using live and video based methods: towards a novel methodological approach. In *Proceedings of the 9th IEEE International Workshop on Advanced Motion Control*, pages 750–755. IEEE.
- Woods, S. N., Walters, M. L., Koay, K. L., and Dautenhahn, K. (2006b). Methodological issues in hri: A comparison of live and video-based methods in robot to human approach direction trials. In *Proceedings of the 15th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2006)*, pages 51–58. IEEE.
- Yamazaki, K., Yamazaki, A., Okada, M., Kuno, Y., Kobayashi, Y., Hoshi, Y., Pitsch, K., Luff, P., vom Lehn, D., and Heath, C. (2009). Revealing gauguin: Engaging visitors in robot guide’s explanation in an art museum. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI 2009)*, pages 1437–1446. ACM.
- Ye, J., Chen, X., and Jian, N. (2012). Impact analysis of human factors on pedestrian traffic characteristics. *Fire Safety Journal*, 52:46–54.
- Yohanan, S. and MacLean, K. E. (2012). The Role of Affective Touch in Human-Robot Interaction: Human Intent and Expectations in Touching the Haptic Creature. *International Journal of Social Robotics*, 4(2):163–180.
- Young, S. B. (1999). Evaluation of Pedestrian Walking Speeds in Airport Terminals. *Transportation Research Record: Journal of the Transportation Research Board*, 1674(1):20–26.

List of publications

Relevant for this thesis

Peer-reviewed conference publications- and journal articles

- **Joosse, M.P.** and Evers, V. (2017). A Guide Robot at the Airport: First Impressions. In *Proceedings of the 11th ACM International Conference on Human-Robot Interaction (HRI 2017)*, pages 149-150. IEEE.
- Triebel, R., Arras, K., Alami, R., Beyer, L., Breuers, S., Chatila, R., Chetouani, M., Cremers, D., Evers, V., Fiore, M., Hung, H., Islas Ramirez, O.A., **Joosse, M.P.**, Khambhaita, H., Kucner, T., Leibe, B., Lilienthal, A., Linder, T., Lohse, M., Magnusson, M., Okal, B., Palmieri, L., Rafi, U., van Rooij, M.M.J.W., and Zhang, L. (2016). SPENCER: A Socially Aware Service Robot for Passenger Guidance and Help in Busy Airports. In *Field and Service Robotics: Results of the 10th International Conference*, pages 607-622. Springer
- **Joosse, M.P.**, Knuppe, R.A., Pinggen, G.L.J., Varkevisser, R.A., Vukoja, J., Lohse, M., and Evers, V. (2015). Robots Guiding Small Groups: The Effect of Appearance Change on the User Experience. In *Proceedings of the Fourth International Symposium on New Frontiers in Human-Robot Interaction*. University of Kent.
- Vroon, J.H., **Joosse, M.P.**, Lohse, M., Kolkmeier, J., Kim, J., Truong, K.P., Englebienne, G., Heylen, D.K.J., and Evers, V. (2015). Dynamics of Social Positioning Patterns in Group-Robot Interactions. In *Proceedings of the 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN 2015)*, pages 394-399. IEEE.
- **Joosse, M.P.**, Poppe, R.W., Lohse, M., and Evers, V. (2014). Cultural Differences in how an Engagement-Seeking Robot should Approach a Group of People. In *Proceedings of the 5th ACM International Conference on Collaboration Across Boundaries: Culture, Distance & Technology (CABS 2014)*, pages 121-130. ACM.

Technical reports and peer-reviewed workshop papers

- **Joosse, M.P.**, Van Waveren, S., Zaga, C., and Evers, V. (2017). Groups in Conflict at the Airport: How Everyday People Think a Robot Should Act. Paper presented at the CSCW'17 Workshop on Robots in Groups and Teams.
- **Joosse, M.P.**, Lohse, M., and Evers, V. (2015). How a Guide Robot Should Behave at an Airport - Insights Based on Observing Passengers. Technical report TR-CTIT-15-01.
- **Joosse, M.P.**, Lohse, M., and Evers, V. (2015). Crowdsourcing Culture in HRI: Lessons Learned From Quantitative and Qualitative Data Collections. Paper presented at the 3rd International Workshop on Culture Aware Robotics at ICSR'15, Paris, 30 October 2015.

Additional publications

- **Joosse, M.P.**, Lohse, M., and Evers, V. (2014) Sound over Matter: The Effects of Functional Noise, Robot Size and Approach Velocity in Human-Robot Encounters. In *Proceed-*

- ings of the 2014 ACM/IEEE Conference on Human-Robot Interaction (HRI 2014), pages 184-185. ACM.
- Karreman, D.E., Utama, L., **Joosse, M.P.**, Lohse, M., van Dijk, E.M.A.G., and Evers, V. (2014). Robot Etiquette: How to Approach a Pair of People? In *Proceedings of the 2014 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2014)*, pages 196-197. ACM.
 - **Joosse, M.P.**, Sardar, A.H., Lohse, M., and Evers, V. (2013) BEHAVE-II: The Revised Set of Measures to Assess Users' Attitudinal and Behavioral Responses to a Social Robot. *International Journal of Social Robotics*, 5(3):379-388.
 - **Joosse, M.P.**, Lohse, M., Gallego Pérez, J., and Evers, V. (2013) What you do is who you are: The role of task context in perceived social robot personality. In *Proceedings of the 2013 IEEE International Conference on Robotics and Automation (ICRA 2013)*, pages 2134-2139. IEEE.
 - Lohse, M., van Berkel, N., van Dijk, E.M.A.G., **Joosse, M.P.**, Karreman, D.E., and Evers, V. (2013) The Influence of Approach Speed and Functional Noise on Users' Perception of a Robot. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2013)*, pages 1670-1675. IEEE.
 - Sardar, A.H., **Joosse, M.P.**, Weiss, A., and Evers, V. (2012) Don't stand so close to me: Users' attitudinal and behavioral responses to personal space invasion by a robot. In *Proceedings of the 7th Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI 2012)*, pages 229-230. ACM.
 - **Joosse, M.P.**, Sardar, A.H., and Evers, V. (2011). Behave: a Set of Measures to Assess Users' Attitudinal and Non-Verbal Behavioral Responses to a Robot's Social Behaviors. In *Proceedings of the 3rd International Conference on Social Robotics (ICSR 2011)*, pages 84-94. Springer.

2011

- 1 Botond Cseke (RUN) *Variational Algorithms for Bayesian Inference in Latent Gaussian Models*
- 2 Nick Tinnemeier (UU) *Organizing Agent Organizations. Syntax and Operational Semantics of an Organization-Oriented Programming Language*
- 3 Jan Martijn van der Werf (TUE) *Compositional Design and Verification of Component-Based Information Systems*
- 4 Hado van Hasselt (UU) *Insights in Reinforcement Learning: Formal analysis and empirical evaluation of temporal-difference*
- 5 Base van der Raadt (VUA) *Enterprise Architecture Coming of Age: Increasing the Performance of an Emerging Discipline*
- 6 Yiwen Wang (TUE) *Semantically-Enhanced Recommendations in Cultural Heritage*
- 7 Yujia Cao (UT) *Multimodal Information Presentation for High Load Human Computer Interaction*
- 8 Nieske Vergunst (UU) *BDI-based Generation of Robust Task-Oriented Dialogues*
- 9 Tim de Jong (OU) *Contextualised Mobile Media for Learning*
- 10 Bart Bogaert (UvT) *Cloud Content Contention*
- 11 Dhaval Vyas (UT) *Designing for Awareness: An Experience-focused HCI Perspective*
- 12 Carmen Bratosin (TUE) *Grid Architecture for Distributed Process Mining*
- 13 Xiaoyu Mao (UvT) *Airport under Control. Multiagent Scheduling for Airport Ground Handling*
- 14 Milan Lovric (EUR) *Behavioral Finance and Agent-Based Artificial Markets*
- 15 Marijn Koolen (UvA) *The Meaning of Structure: the Value of Link Evidence for Information Retrieval*
- 16 Maarten Schadd (UM) *Selective Search in Games of Different Complexity*
- 17 Jiyin He (UvA) *Exploring Topic Structure: Coherence, Diversity and Relatedness*
- 18 Mark Ponsen (UM) *Strategic Decision-Making in complex games*
- 19 Ellen Rusman (OU) *The Mind 's Eye on Personal Profiles*
- 20 Qing Gu (VUA) *Guiding service-oriented software engineering: A view-based approach*
- 21 Linda Terlouw (TUD) *Modularization and Specification of Service-Oriented Systems*
- 22 Junte Zhang (UvA) *System Evaluation of Archival Description and Access*
- 23 Wouter Weerkamp (UvA) *Finding People and their Utterances in Social Media*
- 24 Herwin van Welbergen (UT) *Behavior Generation for Interpersonal Coordination with Virtual Humans On Specifying, Scheduling and Realizing Multimodal Virtual Human Behavior*
- 25 Syed Waqar ul Qounain Jaffry (VUA) *Analysis and Validation of Models for Trust Dynamics*
- 26 Matthijs Aart Pontier (VUA) *Virtual Agents for Human Communication: Emotion Regulation and Involvement-Distance Trade-Offs in Embodied Conversational Agents and Robots*
- 27 Aniel Bhulai (VUA) *Dynamic website optimization through autonomous management of design patterns*
- 28 Rianne Kaptein (UvA) *Effective Focused Retrieval by Exploiting Query Context and Document Structure*
- 29 Faisal Kamiran (TUE) *Discrimination-aware Classification*
- 30 Egon van den Broek (UT) *Affective Signal Processing (ASP): Unraveling the mystery of emotions*
- 31 Ludo Waltman (EUR) *Computational and Game-Theoretic Approaches for Modeling Bounded Rationality*
- 32 Nees-Jan van Eck (EUR) *Methodological Advances in Bibliometric Mapping of Science*
- 33 Tom van der Weide (UU) *Arguing to Motivate Decisions*
- 34 Paolo Turrini (UU) *Strategic Reasoning in Interdependence: Logical and Game-theoretical Investigations*

- 35 Maaïke Harbers (UU) *Explaining Agent Behavior in Virtual Training*
 - 36 Erik van der Spek (UU) *Experiments in serious game design: a cognitive approach*
 - 37 Adriana Burlutiu (RUN) *Machine Learning for Pairwise Data, Applications for Preference Learning and Supervised Network Inference*
 - 38 Nyree Lemmens (UM) *Bee-inspired Distributed Optimization*
 - 39 Joost Westra (UU) *Organizing Adaptation using Agents in Serious Games*
 - 40 Viktor Clerc (VUA) *Architectural Knowledge Management in Global Software Development*
 - 41 Luan Ibraimi (UT) *Cryptographically Enforced Distributed Data Access Control*
 - 42 Michal Sindlar (UU) *Explaining Behavior through Mental State Attribution*
 - 43 Henk van der Schuur (UU) *Process Improvement through Software Operation Knowledge*
 - 44 Boris Reuderink (UT) *Robust Brain-Computer Interfaces*
 - 45 Herman Stehouwer (UvT) *Statistical Language Models for Alternative Sequence Selection*
 - 46 Beibei Hu (TUD) *Towards Contextualized Information Delivery: A Rule-based Architecture for the Domain of Mobile Police Work*
 - 47 Azizi Bin Ab Aziz (VUA) *Exploring Computational Models for Intelligent Support of Persons with Depression*
 - 48 Mark Ter Maat (UT) *Response Selection and Turn-taking for a Sensitive Artificial Listening Agent*
 - 49 Andreea Niculescu (UT) *Conversational interfaces for task-oriented spoken dialogues: design aspects influencing interaction quality*
- 2012**
- 1 Terry Kakeeto (UvT) *Relationship Marketing for SMEs in Uganda*
 - 2 Muhammad Umair (VUA) *Adaptivity, emotion, and Rationality in Human and Ambient Agent Models*
 - 3 Adam Vanya (VUA) *Supporting Architecture Evolution by Mining Software Repositories*
 - 4 Jurriaan Souer (UU) *Development of Content Management System-based Web Applications*
 - 5 Marijn Plomp (UU) *Maturing Interorganisational Information Systems*
 - 6 Wolfgang Reinhardt (OU) *Awareness Support for Knowledge Workers in Research Networks*
 - 7 Rianne van Lambalgen (VUA) *When the Going Gets Tough: Exploring Agent-based Models of Human Performance under Demanding Conditions*
 - 8 Gerben de Vries (UvA) *Kernel Methods for Vessel Trajectories*
 - 9 Ricardo Neisse (UT) *Trust and Privacy Management Support for Context-Aware Service Platforms*
 - 10 David Smits (TUE) *Towards a Generic Distributed Adaptive Hypermedia Environment*
 - 11 J. C. B. Rantham Prabhakara (TUE) *Process Mining in the Large: Preprocessing, Discovery, and Diagnostics*
 - 12 Kees van der Sluijs (TUE) *Model Driven Design and Data Integration in Semantic Web Information Systems*
 - 13 Suleman Shahid (UvT) *Fun and Face: Exploring non-verbal expressions of emotion during playful interactions*
 - 14 Evgeny Knutov (TUE) *Generic Adaptation Framework for Unifying Adaptive Web-based Systems*
 - 15 Natalie van der Wal (VUA) *Social Agents. Agent-Based Modelling of Integrated Internal and Social Dynamics of Cognitive and Affective Processes*
 - 16 Fiemke Both (VUA) *Helping people by understanding them: Ambient Agents supporting task execution and depression treatment*
 - 17 Amal Elgammal (UvT) *Towards a Comprehensive Framework for Business Process Compliance*
 - 18 Eltjo Poort (VUA) *Improving Solution Architecting Practices*
 - 19 Helen Schonenberg (TUE) *What's Next? Operational Support for Business Process Execution*
 - 20 Ali Bahramisharif (RUN) *Covert Visual Spatial Attention, a Robust Paradigm for Brain-Computer Interfacing*
 - 21 Roberto Cornacchia (TUD) *Querying Sparse Matrices for Information Retrieval*
 - 22 Thijs Vis (UvT) *Intelligence, politie en veiligheidsdienst: verenigbare grootheden?*
 - 23 Christian Muehl (UT) *Toward Affective Brain-Computer Interfaces: Exploring the Neurophysiology of Affect during Human Media Interaction*
 - 24 Laurens van der Werff (UT) *Evaluation of Noisy Transcripts for Spoken Document Retrieval*
 - 25 Silja Eckartz (UT) *Managing the Business Case Development in Inter-Organizational IT Projects: A Methodology and its Application*
 - 26 Emile de Maat (UvA) *Making Sense of Legal Text*
 - 27 Hayrettin Gurkok (UT) *Mind the Sheep! User Experience Evaluation & Brain-Computer Interface Games*
 - 28 Nancy Pascall (UvT) *Engendering Technology Empowering Women*
 - 29 Almer Tigelaar (UT) *Peer-to-Peer Information Retrieval*
 - 30 Alina Pommeranz (TUD) *Designing Human-Centered Systems for Reflective Decision Making*

- 31 Emily Bagarukayo (RUN) *A Learning by Construction Approach for Higher Order Cognitive Skills Improvement, Building Capacity and Infrastructure*
- 32 Wietske Visser (TUD) *Qualitative multi-criteria preference representation and reasoning*
- 33 Rory Sie (OUN) *Coalitions in Cooperation Networks (COCOON)*
- 34 Pavol Jancura (RUN) *Evolutionary analysis in PPI networks and applications*
- 35 Evert Haasdijk (VUA) *Never Too Old To Learn: On-line Evolution of Controllers in Swarm- and Modular Robotics*
- 36 Denis Ssebugwawo (RUN) *Analysis and Evaluation of Collaborative Modeling Processes*
- 37 Agnes Nakakawa (RUN) *A Collaboration Process for Enterprise Architecture Creation*
- 38 Selmar Smit (VUA) *Parameter Tuning and Scientific Testing in Evolutionary Algorithms*
- 39 Hassan Fatemi (UT) *Risk-aware design of value and coordination networks*
- 40 Agus Gunawan (UVT) *Information Access for SMEs in Indonesia*
- 41 Sebastian Kelle (OU) *Game Design Patterns for Learning*
- 42 Dominique Verpoorten (OU) *Reflection Amplifiers in self-regulated Learning*
- 43 Anna Tordai (VUA) *On Combining Alignment Techniques*
- 44 Benedikt Kratz (UVT) *A Model and Language for Business-aware Transactions*
- 45 Simon Carter (UVa) *Exploration and Exploitation of Multilingual Data for Statistical Machine Translation*
- 46 Manos Tsagkias (UVa) *Mining Social Media: Tracking Content and Predicting Behavior*
- 47 Jorn Bakker (TUE) *Handling Abrupt Changes in Evolving Time-series Data*
- 48 Michael Kaisers (UM) *Learning against Learning: Evolutionary dynamics of reinforcement learning algorithms in strategic interactions*
- 49 Steven van Kervel (TUD) *Ontology driven Enterprise Information Systems Engineering*
- 50 Jeroen de Jong (TUD) *Heuristics in Dynamic Scheduling: a practical framework with a case study in elevator dispatching*
- 5 Dulce Pumareja (UT) *Groupware Requirements Evolutions Patterns*
- 6 Romulo Goncalves (CWI) *The Data Cyclotron: Juggling Data and Queries for a Data Warehouse Audience*
- 7 Giel van Lankveld (UvT) *Quantifying Individual Player Differences*
- 8 Robbert-Jan Merk (VUA) *Making enemies: cognitive modeling for opponent agents in fighter pilot simulators*
- 9 Fabio Gori (RUN) *Metagenomic Data Analysis: Computational Methods and Applications*
- 10 Jeewanie Jayasinghe Arachchige (UvT) *A Unified Modeling Framework for Service Design*
- 11 Evangelos Pournaras (TUD) *Multi-level Reconfigurable Self-organization in Overlay Services*
- 12 Marian Razavian (VUA) *Knowledge-driven Migration to Services*
- 13 Mohammad Safiri (UT) *Service Tailoring: User-centric creation of integrated IT-based homecare services to support independent living of elderly*
- 14 Jafar Tanha (UVa) *Ensemble Approaches to Semi-Supervised Learning*
- 15 Daniel Hennes (UM) *Multiagent Learning: Dynamic Games and Applications*
- 16 Eric Kok (UU) *Exploring the practical benefits of argumentation in multi-agent deliberation*
- 17 Koen Kok (VUA) *The PowerMatcher: Smart Coordination for the Smart Electricity Grid*
- 18 Jeroen Janssens (UVT) *Outlier Selection and One-Class Classification*
- 19 Renze Steenhuisen (TUD) *Coordinated Multi-Agent Planning and Scheduling*
- 20 Katja Hofmann (UVa) *Fast and Reliable Online Learning to Rank for Information Retrieval*
- 21 Sander Wubben (UVT) *Text-to-text generation by monolingual machine translation*
- 22 Tom Claassen (RUN) *Causal Discovery and Logic*
- 23 Patricio de Alencar Silva (UVT) *Value Activity Monitoring*
- 24 Haitham Bou Ammar (UM) *Automated Transfer in Reinforcement Learning*
- 25 Agnieszka Anna Latoszek-Berendsen (UM) *Intention-based Decision Support. A new way of representing and implementing clinical guidelines in a Decision Support System*
- 26 Alireza Zarghami (UT) *Architectural Support for Dynamic Homecare Service Provisioning*
- 27 Mohammad Huq (UT) *Inference-based Framework Managing Data Provenance*
- 28 Frans van der Sluis (UT) *When Complexity becomes Interesting: An Inquiry into the Information eXperience*
- 29 Iwan de Kok (UT) *Listening Heads*
- 30 Joyce Nakatumba (TUE) *Resource-Aware Business Process Management: Analysis and Support*
- 31 Dinh Khoa Nguyen (UVT) *Blueprint Model and Language for Engineering Cloud Applications*

2013

- 1 Viorel Milea (EUR) *News Analytics for Financial Decision Support*
- 2 Erietta Liarou (CWI) *MonetDB/DataCell: Leveraging the Column-store Database Technology for Efficient and Scalable Stream Processing*
- 3 Szymon Klarman (VUA) *Reasoning with Contexts in Description Logics*
- 4 Chetan Yadati (TUD) *Coordinating autonomous planning and scheduling*

- 32 Kamakshi Rajagopal (OUN) *Networking For Learning: The role of Networking in a Lifelong Learner's Professional Development*
 - 33 Qi Gao (TUD) *User Modeling and Personalization in the Microblogging Sphere*
 - 34 Kien Tjin-Kam-Jet (UT) *Distributed Deep Web Search*
 - 35 Abdallah El Ali (UvA) *Minimal Mobile Human Computer Interaction*
 - 36 Than Lam Hoang (TUE) *Pattern Mining in Data Streams*
 - 37 Dirk Börner (OUN) *Ambient Learning Displays*
 - 38 Eelco den Heijer (VUA) *Autonomous Evolutionary Art*
 - 39 Joop de Jong (TUD) *A Method for Enterprise Ontology based Design of Enterprise Information Systems*
 - 40 Pim Nijssen (UM) *Monte-Carlo Tree Search for Multi-Player Games*
 - 41 Jochem Liem (UvA) *Supporting the Conceptual Modelling of Dynamic Systems: A Knowledge Engineering Perspective on Qualitative Reasoning*
 - 42 Léon Planken (TUD) *Algorithms for Simple Temporal Reasoning*
 - 43 Marc Bron (UvA) *Exploration and Contextualization through Interaction and Concepts*
- 2014**
- 1 Nicola Barile (UU) *Studies in Learning Monotone Models from Data*
 - 2 Fiona Tulyano (RUN) *Combining System Dynamics with a Domain Modeling Method*
 - 3 Sergio Raul Duarte Torres (UT) *Information Retrieval for Children: Search Behavior and Solutions*
 - 4 Hanna Jochmann-Mannak (UT) *Websites for children: search strategies and interface design - Three studies on children's search performance and evaluation*
 - 5 Jurriaan van Reijssen (UU) *Knowledge Perspectives on Advancing Dynamic Capability*
 - 6 Damian Tamburri (VUA) *Supporting Networked Software Development*
 - 7 Arya Adriansyah (TUE) *Aligning Observed and Modeled Behavior*
 - 8 Samur Araujo (TUD) *Data Integration over Distributed and Heterogeneous Data Endpoints*
 - 9 Philip Jackson (UvT) *Toward Human-Level Artificial Intelligence: Representation and Computation of Meaning in Natural Language*
 - 10 Ivan Salvador Razo Zapata (VUA) *Service Value Networks*
 - 11 Janneke van der Zwaan (TUD) *An Empathic Virtual Buddy for Social Support*
 - 12 Willem van Willigen (VUA) *Look Ma, No Hands: Aspects of Autonomous Vehicle Control*
 - 13 Arlette van Wissen (VUA) *Agent-Based Support for Behavior Change: Models and Applications in Health and Safety Domains*
 - 14 Yangyang Shi (TUD) *Language Models With Meta-information*
 - 15 Natalya Mogles (VUA) *Agent-Based Analysis and Support of Human Functioning in Complex Socio-Technical Systems: Applications in Safety and Healthcare*
 - 16 Krystyna Milian (VUA) *Supporting trial recruitment and design by automatically interpreting eligibility criteria*
 - 17 Kathrin Dentler (VUA) *Computing healthcare quality indicators automatically: Secondary Use of Patient Data and Semantic Interoperability*
 - 18 Mattijs Ghijsen (UvA) *Methods and Models for the Design and Study of Dynamic Agent Organizations*
 - 19 Vinicius Ramos (TUE) *Adaptive Hypermedia Courses: Qualitative and Quantitative Evaluation and Tool Support*
 - 20 Mena Habib (UT) *Named Entity Extraction and Disambiguation for Informal Text: The Missing Link*
 - 21 Cassidy Clark (TUD) *Negotiation and Monitoring in Open Environments*
 - 22 Marieke Peeters (UU) *Personalized Educational Games: Developing agent-supported scenario-based training*
 - 23 Eleftherios Sidirourgos (UvA/CWI) *Space Efficient Indexes for the Big Data Era*
 - 24 Davide Ceolin (VUA) *Trusting Semi-structured Web Data*
 - 25 Martijn Lappenschaar (RUN) *New network models for the analysis of disease interaction*
 - 26 Tim Baarslag (TUD) *What to Bid and When to Stop*
 - 27 Rui Jorge Almeida (EUR) *Conditional Density Models Integrating Fuzzy and Probabilistic Representations of Uncertainty*
 - 28 Anna Chmielowiec (VUA) *Decentralized k-Clique Matching*
 - 29 Jaap Kabbedijk (UU) *Variability in Multi-Tenant Enterprise Software*
 - 30 Peter de Cock (UvT) *Anticipating Criminal Behaviour*
 - 31 Leo van Moergestel (UU) *Agent Technology in Agile Multiparallel Manufacturing and Product Support*
 - 32 Naser Ayat (UvA) *On Entity Resolution in Probabilistic Data*
 - 33 Tesfa Tegegne (RUN) *Service Discovery in eHealth*
 - 34 Christina Manteli (VUA) *The Effect of Governance in Global Software Development: Analyzing Transactive Memory Systems*
 - 35 Joost van Ooijen (UU) *Cognitive Agents in Virtual Worlds: A Middleware Design Approach*

- 36 Joos Buijs (TUE) *Flexible Evolutionary Algorithms for Mining Structured Process Models*
- 37 Maral Dadvar (UT) *Experts and Machines United Against Cyberbullying*
- 38 Danny Plass-Oude Bos (UT) *Making brain-computer interfaces better: improving usability through post-processing*
- 39 Jasmina Maric (UvT) *Web Communities, Immigration, and Social Capital*
- 40 Walter Omona (RUN) *A Framework for Knowledge Management Using ICT in Higher Education*
- 41 Frederic Hogenboom (EUR) *Automated Detection of Financial Events in News Text*
- 42 Carsten Eijkhof (CWI/TUD) *Contextual Multi-dimensional Relevance Models*
- 43 Kevin Vlaanderen (UU) *Supporting Process Improvement using Method Increments*
- 44 Paulien Meesters (UvT) *Intelligent Blauw: Intelligence-gestuurde politiezorg in gebiedsgebonden eenheden*
- 45 Birgit Schmitz (OUN) *Mobile Games for Learning: A Pattern-Based Approach*
- 46 Ke Tao (TUD) *Social Web Data Analytics: Relevance, Redundancy, Diversity*
- 47 Shangsong Liang (UvA) *Fusion and Diversification in Information Retrieval*
- 13 Giuseppe Procaccianti (VUA) *Energy-Efficient Software*
- 14 Bart van Straalen (UT) *A cognitive approach to modeling bad news conversations*
- 15 Klaas Andries de Graaf (VUA) *Ontology-based Software Architecture Documentation*
- 16 Changyun Wei (UT) *Cognitive Coordination for Cooperative Multi-Robot Teamwork*
- 17 André van Cleeff (UT) *Physical and Digital Security Mechanisms: Properties, Combinations and Trade-offs*
- 18 Holger Pirk (CWI) *Waste Not, Want Not!: Managing Relational Data in Asymmetric Memories*
- 19 Bernardo Tabuenca (OUN) *Ubiquitous Technology for Lifelong Learners*
- 20 Lois Vanhée (UU) *Using Culture and Values to Support Flexible Coordination*
- 21 Sibren Fetter (OUN) *Using Peer-Support to Expand and Stabilize Online Learning*
- 22 Zheming Zhu (UT) *Co-occurrence Rate Networks*
- 23 Luit Gazendam (VUA) *Cataloguer Support in Cultural Heritage*
- 24 Richard Berendsen (UvA) *Finding People, Papers, and Posts: Vertical Search Algorithms and Evaluation*
- 25 Steven Woudenberg (UU) *Bayesian Tools for Early Disease Detection*
- 26 Alexander Hogenboom (EUR) *Sentiment Analysis of Text Guided by Semantics and Structure*
- 27 Sándor Héman (CWI) *Updating compressed column-stores*
- 28 Janet Bagorogoza (TiU) *Knowledge Management and High Performance: The Uganda Financial Institutions Model for HPO*
- 29 Hendrik Baier (UM) *Monte-Carlo Tree Search Enhancements for One-Player and Two-Player Domains*
- 30 Kiavash Bahreini (OUN) *Real-time Multimodal Emotion Recognition in E-Learning*
- 31 Yakup Koç (TUD) *On Robustness of Power Grids*
- 32 Jerome Gard (UL) *Corporate Venture Management in SMEs*
- 33 Frederik Schadd (UM) *Ontology Mapping with Auxiliary Resources*
- 34 Victor de Graaff (UT) *Geosocial Recommender Systems*
- 35 Junchao Xu (TUD) *Affective Body Language of Humanoid Robots: Perception and Effects in Human Robot Interaction*

2015

- 1 Niels Netten (UvA) *Machine Learning for Relevance of Information in Crisis Response*
- 2 Faiza Bukhsh (UvT) *Smart auditing: Innovative Compliance Checking in Customs Controls*
- 3 Twan van Laarhoven (RUN) *Machine learning for network data*
- 4 Howard Spoelstra (OUN) *Collaborations in Open Learning Environments*
- 5 Christoph Bösch (UT) *Cryptographically Enforced Search Pattern Hiding*
- 6 Farideh Heidari (TUD) *Business Process Quality Computation: Computing Non-Functional Requirements to Improve Business Processes*
- 7 Maria-Hendrike Peetz (UvA) *Time-Aware Online Reputation Analysis*
- 8 Jie Jiang (TUD) *Organizational Compliance: An agent-based model for designing and evaluating organizational interactions*
- 9 Randy Klaassen (UT) *HCI Perspectives on Behavior Change Support Systems*
- 10 Henry Hermans (OUN) *OpenU: design of an integrated system to support lifelong learning*
- 11 Yongming Luo (TUE) *Designing algorithms for big graph datasets: A study of computing bisimulation and joins*
- 12 Julie M. Birkholz (VUA) *Modi Operandi of Social Network Dynamics: The Effect of Context on Scientific Collaboration Networks*

2016

- 1 Syed Saiden Abbas (RUN) *Recognition of Shapes by Humans and Machines*
- 2 Michiel Christiaan Meulendijk (UU) *Optimizing medication reviews through decision support: prescribing a better pill to swallow*
- 3 Maya Sappelli (RUN) *Knowledge Work in Context: User Centered Knowledge Worker Support*

- 4 Laurens Rietveld (VUA) *Publishing and Consuming Linked Data*
- 5 Evgeny Sherkhonov (UvA) *Expanded Acyclic Queries: Containment and an Application in Explaining Missing Answers*
- 6 Michel Wilson (TUD) *Robust scheduling in an uncertain environment*
- 7 Jeroen de Man (VUA) *Measuring and modeling negative emotions for virtual training*
- 8 Matje van de Camp (TiU) *A Link to the Past: Constructing Historical Social Networks from Unstructured Data*
- 9 Archana Nottamkandath (VUA) *Trusting Crowdsourced Information on Cultural Artefacts*
- 10 George Karafotias (VUA) *Parameter Control for Evolutionary Algorithms*
- 11 Anne Schuth (UvA) *Search Engines that Learn from Their Users*
- 12 Max Knobbout (UU) *Logics for Modelling and Verifying Normative Multi-Agent Systems*
- 13 Nana Baah Gyan (VUA) *The Web, Speech Technologies and Rural Development in West Africa: An ICT4D Approach*
- 14 Ravi Khadka (UU) *Revisiting Legacy Software System Modernization*
- 15 Steffen Michels (RUN) *Hybrid Probabilistic Logics: Theoretical Aspects, Algorithms and Experiments*
- 16 Guangliang Li (UvA) *Socially Intelligent Autonomous Agents that Learn from Human Reward*
- 17 Berend Weel (VUA) *Towards Embodied Evolution of Robot Organisms*
- 18 Albert Meroño Peñuela (VUA) *Refining Statistical Data on the Web*
- 19 Julia Efremova (TUE) *Mining Social Structures from Genealogical Data*
- 20 Daan Odijk (UvA) *Context & Semantics in News & Web Search*
- 21 Alejandro Moreno Céleri (UT) *From Traditional to Interactive Playspaces: Automatic Analysis of Player Behavior in the Interactive Tag Playground*
- 22 Grace Lewis (VUA) *Software Architecture Strategies for Cyber-Foraging Systems*
- 23 Fei Cai (UvA) *Query Auto Completion in Information Retrieval*
- 24 Brend Wanders (UT) *Repurposing and Probabilistic Integration of Data; An Iterative and data model independent approach*
- 25 Julia Kiseleva (TUE) *Using Contextual Information to Understand Searching and Browsing Behavior*
- 26 Dilhan Thilakaratne (VUA) *In or Out of Control: Exploring Computational Models to Study the Role of Human Awareness and Control in Behavioural Choices, with Applications in Aviation and Energy Management Domains*
- 27 Wen Li (TUD) *Understanding Geo-spatial Information on Social Media*
- 28 Mingxin Zhang (TUD) *Large-scale Agent-based Social Simulation - A study on epidemic prediction and control*
- 29 Nicolas Höning (TUD) *Peak reduction in decentralised electricity systems -Markets and prices for flexible planning*
- 30 Ruud Mattheij (UvT) *The Eyes Have It*
- 31 Mohammad Khelghati (UT) *Deep web content monitoring*
- 32 Eelco Vriezekolk (UT) *Assessing Telecommunication Service Availability Risks for Crisis Organisations*
- 33 Peter Bloem (UvA) *Single Sample Statistics, exercises in learning from just one example*
- 34 Dennis Schunselaar (TUE) *Configurable Process Trees: Elicitation, Analysis, and Enactment*
- 35 Zhaochun Ren (UvA) *Monitoring Social Media: Summarization, Classification and Recommendation*
- 36 Daphne Karreman (UT) *Beyond R2D2: The design of nonverbal interaction behavior optimized for robot-specific morphologies*
- 37 Giovanni Sileno (UvA) *Aligning Law and Action - a conceptual and computational inquiry*
- 38 Andrea Minuto (UT) *MATERIALS THAT MATTER - Smart Materials meet Art & Interaction Design*
- 39 Merijn Bruijnes (UT) *Believable Suspect Agents; Response and Interpersonal Style Selection for an Artificial Suspect*
- 40 Christian Detweiler (TUD) *Accounting for Values in Design*
- 41 Thomas King (TUD) *Governing Governance: A Formal Framework for Analysing Institutional Design and Enactment Governance*
- 42 Spyros Martzoukos (UVA) *Combinatorial and Compositional Aspects of Bilingual Aligned Corpora*
- 43 Saskia Koldijk (RUN) *Context-Aware Support for Stress Self-Management: From Theory to Practice*
- 44 Thibault Sellam (UVA) *Automatic Assistants for Database Exploration*
- 45 Bram van de Laar (UT) *Experiencing Brain-Computer Interface Control*
- 46 Jorge Gallego Perez (UT) *Robots to Make you Happy*
- 47 Christina Weber (UL) *Real-time foresight - Preparedness for dynamic innovation networks*
- 48 Tanja Buttler (TUD) *Collecting Lessons Learned*
- 49 Gleb Polevoy (TUD) *Participation and Interaction in Projects. A Game-Theoretic Analysis*
- 50 Yan Wang (UvT) *The Bridge of Dreams: Towards a Method for Operational Performance Alignment in IT-enabled Service Supply Chains*

2017

- 1 Jan-Jaap Oerlemans (UL) *Investigating Cyber-crime*
- 2 Sjoerd Timmer (UU) *Designing and Understanding Forensic Bayesian Networks using Argumentation*
- 3 Daniël Harold Telgen (UU) *Grid Manufacturing; A Cyber-Physical Approach with Autonomous Products and Reconfigurable Manufacturing Machines*
- 4 Mrunal Gawade (CWI) *Multi-core Parallelism in a Column-store*
- 5 Mahdieh Shadi (UVA) *Collaboration Behavior*
- 6 Damir Vandic (EUR) *Intelligent Information Systems for Web Product Search*
- 7 Roel Bertens (UU) *Insight in Information: from Abstract to Anomaly*
- 8 Rob Konijn (VU) *Detecting Interesting Differences: Data Mining in Health Insurance Data using Outlier Detection and Subgroup Discovery*
- 9 Dong Nguyen (UT) *Text as Social and Cultural Data: A Computational Perspective on Variation in Text*
- 10 Robby van Delden (UT) *(Steering) Interactive Play Behavior*
- 11 Florian Kunneman (RUN) *Modelling patterns of time and emotion in Twitter #anticipointment*
- 12 Sander Leemans (TUE) *Robust Process Mining with Guarantees*
- 13 Gijs Huisman (UT) *Social Touch Technology - Extending the reach of social touch through haptic technology*
- 14 Shoshannah Tekofsky (UvT) *You Are Who You Play You Are: Modelling Player Traits from Video Game Behavior*
- 15 Peter Berck (RUN) *Memory-Based Text Correction*
- 16 Aleksandr Chuklin (UVA) *Understanding and Modeling Users of Modern Search Engines*
- 17 Daniel Dimov (UL) *Crowdsourced Online Dispute Resolution*
- 18 Ridho Reinanda (UVA) *Entity Associations for Search*
- 19 Jeroen Vuurens (UT) *Proximity of Terms, Texts and Semantic Vectors in Information Retrieval*
- 20 Mohammadbashir Sedighi (TUD) *Fostering Engagement in Knowledge Sharing: The Role of Perceived Benefits, Costs and Visibility*
- 21 Jeroen Linssen (UT) *Meta Matters in Interactive Storytelling and Serious Gaming (A Play on Worlds)*
- 22 Sara Magliacane (VU) *Logics for causal inference under uncertainty*
- 23 David Graus (UVA) *Entities of Interest — Discovery in Digital Traces*
- 24 Chang Wang (TUD) *Use of Affordances for Efficient Robot Learning*
- 25 Veruska Zamborlini (VU) *Knowledge Representation for Clinical Guidelines, with applications to Multimorbidity Analysis and Literature Search*
- 26 Merel Jung (UT) *Socially intelligent robots that understand and respond to human touch*
- 27 Michiel Joosse (UT) *Investigating Positioning and Gaze Behaviors of Social Robots: People's Preferences, Perceptions and Behaviors*

As technology advances, application areas for robots are no longer limited to the factories where they perform repetitive tasks behind fences. Robots are envisioned to provide services to us in everyday public spaces - in which they will encounter and interact with people. These types of robots can be seen as social robots, they should interact with people following the behavioral norms people expect of the robot within a specific context.

In this thesis we have investigated people's preferences for, perceptions of, and behaviors towards social robots through a series (lab and field) studies. The work presented provides empirical and methodological contributions to the development of social robots in semi-public spaces.

In order to design behaviors for future social robots it is necessary for HRI researchers to study people's interactions with robots in this context of use during the development cycle of a social robot.

CTIT

CTIT Ph.D. Thesis Series No. 17-440
DOI 10.3990/1.9789036543767

ISBN 978-90-365-4376-7

9 789036 543767